



HAL
open science

Tests optimaux pour détecter les signaux faibles dans les séries chronologiques

Marwa Ltaifa

► **To cite this version:**

Marwa Ltaifa. Tests optimaux pour détecter les signaux faibles dans les séries chronologiques. Mathématiques [math]. Université de Lorraine; Université de Sousse (Tunisie), 2021. Français. NNT : 2021LORR0189 . tel-03552957

HAL Id: tel-03552957

<https://hal.univ-lorraine.fr/tel-03552957>

Submitted on 2 Feb 2022

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.



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>



Thèse

présentée pour l'obtention du titre de

Docteur de l'Université de Lorraine et de l'Université de Sousse

Mention : Mathématiques Appliquées (Statistique)

par

Marwa Ltaifa Ep Zorgati

Tests optimaux pour détecter les signaux faibles dans les séries chronologiques

Soutenue le 22 octobre 2021 devant le jury composé de

Directeurs :	Joseph Ngatchou Wandji	Professeur à l'Université de Lorraine
	Khalifa El Mabrouk	Professeur à l'Université de Sousse
Rapporteuses :	Madalina Olteanu	Professeure à l'Université Paris Dauphine-PSL
	Sana Louhichi	Professeure à l'Université Grenoble Alpes
Président :	Clément Dombry	Professeur à l'Université de Franche Comté

Institut Élie Cartan de Lorraine, UMR CNRS 7502, Laboratoire de Mathématiques,
BP 70239, 54 506 Vandœuvre-lès-Nancy Cedex, France
ESSTHS, Laboratoire de Mathématiques Modélisation Déterministe et Aléatoire,
BP 285, Rue Lamine Abassi, 4011, Hammam Sousse, Tunisie

Remerciements

Le moment étant venu d'apposer un point final à ce manuscrit de thèse, je tiens à remercier les nombreuses personnes qui m'ont accompagnée, de près comme de loin, dans cette aventure.

Je tiens à remercier en tout premier lieu mon directeur de thèse, Monsieur Joseph Ngatchou-Wandji. Je le connais maintenant depuis près de 6 ans. Il m'a enseigné en master 2 et a été mon maître de stage. De plus en plus passionnée par tout ce qui avait trait à l'étude et l'analyse des données, j'ai décidé de faire ma thèse avec lui. Il a énormément contribué à ce que ma thèse se déroule dans d'excellentes conditions scientifiques, malgré les va et vient et les difficultés matérielles rencontrées au début. Il a toujours été disponible, à l'écoute de mes nombreuses questions, et s'est toujours intéressé à l'avancée de mes travaux. Les nombreuses discussions que nous avons eues ainsi que ses conseils sont pour beaucoup dans le résultat final de ce travail. Sa capacité d'analyse et son enthousiasme m'ont montré que le monde de la recherche pouvait être un univers passionnant. Enfin, ses nombreuses relectures et corrections de cette thèse ont été très appréciables. Cette thèse lui doit beaucoup. Pour tout cela, je lui dis merci.

Je remercie Monsieur Khalifa El Mabrouk, l'autre encadrant de cette thèse. C'est grâce à sa collaboration avec le laboratoire IECL que j'ai eu la chance de travailler avec Monsieur Joseph Ngatchou-Wandji, ce qui s'est avéré une expérience très enrichissante. Malgré ses nombreuses occupations, il s'est montré très disponible et a été à l'écoute. Ses encouragements m'ont beaucoup stimulée.

Je remercie Monsieur Hamdi Fathallah avec qui j'ai eu l'occasion de travailler au début de ma thèse. Je lui suis reconnaissante de m'avoir donné l'opportunité d'enseigner à l'École Supérieure des Sciences et de la Technologie de Hammam Sousse, où je suis toujours étudiante. Je lui dis merci pour ses encouragements.

J'adresse également mes remerciements aux membres du jury. Je remercie particu-

lièrement mesdames Madalina Olteanu et Sana Louhichi pour leur relecture attentive de mon manuscrit. Leurs rapports m'ont permis d'améliorer mon travail. Je tiens à remercier Monsieur Clément Dombry d'avoir accepté d'être président du jury

Je suis reconnaissante à tous les membres du Laboratoire IECL, à son ancien directeur Xavier Antoine et à sa nouvelle directrice Anne Gégout-Petit qui mettent tout en œuvre pour que chaque membre du laboratoire trouve sa place et réussisse sa mission. Je remercie aussi mesdames Paola Schneider, Laurence Quirot et Hala Hamed pour leur extrême gentillesse. Je souhaite particulièrement remercier Mme Élodie Cunat pour son aide précieuse.

Je ne suis pas en mesure de trouver une formule simple pour remercier très particulièrement mes amis Noura, Nahed, Rihab, Youssef, Dorsaf et Hanen pour leur soutien permanent. Je voudrais également remercier toutes les personnes extérieures au domaine universitaire qui m'ont, à leur façon, apporté leur aide.

Je dédie ce travail à l'âme de mon oncle Abdou. À toi mon cher oncle qui m'a encouragée à suivre ce chemin et accompagnée dans toutes les étapes de ma vie. Tu m'as appris à aller jusqu'au bout de mes rêves et voilà qu'aujourd'hui l'un d'eux devient réalité, mais tu n'es plus là. J'espère que là où tu es, tu seras fier de moi. Paix à ton âme.

Bien entendu, ces remerciements ne seraient complets sans la mention de ma famille. Merci à mes parents et mon frère qui ont su croire en moi et qui m'ont apporté toute leur aide quand j'en ai eu besoin. Ils m'ont permis de réaliser que la famille est sacrée. Ils ont été pour moi une vraie source d'inspiration, et ont toujours été à mes côtés durant les moments difficiles. Je les remercie pour m'avoir donné le goût des études, pour m'avoir encouragée et pour avoir contribué de près ou de loin à ce que je suis devenue.

Pour terminer, je tiens à remercier mon mari, Mehdi, pour sa présence, sa confiance, son soutien et sa patience. Je le remercie d'avoir su trouver les mots justes pendant les moments difficiles. Je le remercie pour son amour.

Bien entendu, cette liste n'est pas exhaustive et je remercie tous ceux et celles qui me connaissent et qui me permettent de me sentir exister...

Notation et abréviations

\odot	Produit matriciel d'Hadamard.
\mathbb{N}	Ensemble des entiers naturels.
\mathbb{Z}	Ensemble des entiers relatifs.
\mathbb{R}	Ensemble des nombres réels.
\mathbb{P}	Probabilité.
\mathbb{E}	Espérance.
$\mathbf{1}$	La fonction indicatrice.
<i>v.a.</i>	Variable aléatoire.
<i>p.s.</i>	Presque sûre.
iid	Indépendante et identiquement distribuée.
$g = o_P(1)$	g tend vers 0 en probabilité.
$\mathcal{C}(\mathbb{R}^{k+1}, \mathbb{R})$	L'espace des fonctions continue de \mathbb{R}^{k+1} dans \mathbb{R} .
$\mathcal{C}([0, 1], \mathbb{R})$	L'espace des fonctions continue à valeur réelle dans l'intervalle $[0, 1]$.
$\langle \cdot, \cdot \rangle$	Produit scalaire euclidien .
$\ \cdot\ $	La norme euclidienne sur $\mathbb{R}^d, d \in \mathbb{N}$.
$\ A\ _M = \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} $	La norme de la matrice carrée $A = (a_{ij})_{1 \leq i, j \leq n}$.
\xrightarrow{loi}	Convergence en loi.
$\xrightarrow{p.s.}$	Convergence presque sûre.
\xrightarrow{P}	Convergence en probabilité.
$\mathcal{B}(A)$	Un borélien de A .
$\mathcal{L}_2 = L_2(\eta)$	L'ensemble des fonctions de carrée intégrable par rapport à une mesure positive η .

Pour $h_1, h_2, h \in \mathcal{L}_2$, on pose $\langle h_1, h_2 \rangle = \int_{\mathbb{R}^{k+1}} h_1(\beta)h_2(\beta)d\eta(\beta)$ et $\|h\|_{\mathcal{L}_2} = \langle h, h \rangle^{\frac{1}{2}}$, respectivement le produit scalaire et la norme dans \mathcal{L}_2 .

Soit d, d_1, d_2 des entiers tels que $d = d_1 + d_2$, et \mathcal{G} une fonction continue et dérivable dans \mathbb{R}^d . On note pour $z = (z_1, \dots, z_d) = (z^{(1)}, z^{(2)}) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$,

Gradient de $\mathcal{G}(z)$:

$$\partial_z \mathcal{G}(z) := \left(\frac{\partial \mathcal{G}(z)}{\partial z_1}, \dots, \frac{\partial \mathcal{G}(z)}{\partial z_d} \right), \quad \partial_{z^{(k)}} \mathcal{G}(z) := \left(\frac{\partial \mathcal{G}(z)}{\partial z_1^{(k)}}, \dots, \frac{\partial \mathcal{G}(z)}{\partial z_{d_k}^{(k)}} \right), k = 1, 2$$

Hessien de $\mathcal{G}(z)$:

$$\partial_z^2 \mathcal{G}(z) := \left(\frac{\partial^2 \mathcal{G}(z)}{\partial z_i \partial z_j} : 1 \leq i, j \leq d \right), \quad \partial_{z^{(k)}}^2 \mathcal{G}(z) := \left(\frac{\partial^2 \mathcal{G}(z)}{\partial z_i^{(k)} \partial z_j^{(k)}} : 1 \leq i, j \leq d_k \right), k = 1, 2$$

$$\partial_{z^{(1)}z^{(2)}}^2 \mathcal{G}(z) := \left(\frac{\partial^2 \mathcal{G}(z)}{\partial z_i^{(1)} \partial z_j^{(2)}} : 1 \leq i \leq d_1, 1 \leq j \leq d_2 \right)$$

On désigne par H^\top la transposée d'une matrice H . Soient $U = (U_1, \dots, U_d)^\top$ et $D = (D_1, \dots, D_d)^\top$, où pour tout $i = 1, \dots, d$, $U_i \in \mathbb{R}$ et D_i est une matrice $1 \times d$. On note $D^\top U = U^\top D = (D_1 U_1, \dots, D_d U_d)^\top$.

Articles et publications

- i. "Détection de rupture dans des données climatiques et dans des données de santé". Biostatistiques et sciences de la santé. 2021. <http://www.openscience.fr/Detection-de-rupture-dans-des-donnees-climatiques-et-dans-des-donnees-de-sante>. (Avec Joseph Ngatchou-Wandji).
- ii. "On detecting weak changes in the mean of CHARN models." arXiv :2101.08597. 2021. (Avec Joseph-Ngatchou-Wandji).
- iii. "On detecting weak changes in the mean of CHARN models." 2021. Soumis. (Avec Joseph Ngatchou-Wandji).
- iv. "Detecting changes in some real financial time series". 2021. Soumis. (Avec Joseph Ngatchou-Wandji et Francis Biesmans).

Communications

- i. **Juillet 2021** : Exposé sur la "Détection de ruptures faibles dans la moyenne des modèles CHARN" aux journées des doctorants tenues les 29 et 30 juillet 2021 à l'École Supérieure des Sciences et de la Technologie de Hammam Sousse en Tunisie.
- ii. **Juin 2021** : Exposé sur la "Détection de ruptures faibles dans la moyenne des modèles CHARN" aux 52èmes journées de Statistique organisées par la Société Française de Statistique (SFdS). <https://hal.archives-ouvertes.fr/hal-03329969/document>
- iii. **Novembre 2019** : Exposé sur les "Tests de ruptures dans les modèles CHARN" à la journée des doctorants à Nancy.
- iv. **Juin 2019** : Exposé sur les "Tests de ruptures dans les modèles CHARN" aux 51èmes journées de Statistique organisées par la Société Française de Statistique (SFdS), du 3 à 7 juin 2019 à l'IECL de Nancy. <http://jds2019.sfds.asso.fr/program/Soumissions/subm250.pdf>

- v. **Avril 2019** : Exposé sur les "Tests de ruptures dans les modèles CHARN" aux Huitièmes Rencontres des jeunes statisticiens 2019. Ces journées ont eu lieu du 1er au 5 avril 2019 à Porquerolles.
- vi. **Avril 2019** : Exposé sur les "Tests optimaux pour détecter les ruptures dans la moyenne des modèles CHARN". Groupe de travail de probabilités et statistique à l'ESSTHS. Université de Sousse.
- vii. **Avril 2018** : Poster intitulé "Locally asymptotically optimal test for detecting small breaks in CHARN models". Journée d'automne de l'École Doctorale IAEM Lorraine.

Résumé de la thèse

Cette thèse s'intéresse à la construction de tests localement asymptotiquement optimaux pour détecter les ruptures dans la moyenne des modèles Conditional Heteroskedastic AutoRegressive Nonlinear (CHARN) décrits par l'équation stochastique suivante :

$$X_t = T(Z_{t-1}) + \gamma^\top \omega(t) + V(Z_{t-1})\varepsilon_t, \quad t \in \mathbb{Z}, \quad (1)$$

où $\gamma = (\gamma_1, \dots, \gamma_k, \gamma_{k+1})^\top \in \mathbb{R}^{k+1}$ et pour $t_1, \dots, t_k, 1 < t_1 < \dots < t_k < n$, $\omega(t) = (\mathbb{1}_{[1, t_1]}(t), \mathbb{1}_{[t_1, t_2]}(t), \dots, \mathbb{1}_{[t_k, n]}(t))^\top \in \{0, 1\}^{k+1}$, $(X_t)_{t \in \mathbb{Z}}$ est un processus stationnaire par morceaux et ergodique, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc centré réduit de densité f , pour tout $t \in \mathbb{Z}$, $Z_t = (X_t, \dots, X_{t-p+1})^\top$, $p \in \mathbb{N}$, et T et V sont des fonctions réelles telles que $V(x) > 0, \forall x \in \mathbb{R}^p$ et n le nombre des observations. Le modèle (1) contient une large classe de modèles de séries chronologiques comme les modèles AR, MA, ARMA, ARIMA, ARCH etc.

L'attention est portée aux petites ruptures. Celles difficilement observables à l'œil nu, contrairement à celles considérées dans la littérature. Une telle étude ne semble pas avoir déjà été réalisée dans le contexte des séries chronologiques. Le test étudié est celui du rapport de vraisemblance, pour tester

$$H_0 : \gamma = \gamma_0 \text{ contre } H_\beta^{(n)} : \gamma = \gamma_0 + \frac{\beta}{\sqrt{n}} = \gamma_n, \quad n > 1,$$

pour $\gamma_0 \in \mathbb{R}^{k+1}$ et $\beta \in \mathbb{R}^{k+1}$ caractérisant respectivement la situation où il n'y a pas de rupture, et celle où il y a au moins une rupture à trouver.

Ce document est organisé comme suit :

- Le chapitre 1 constitue l'introduction générale de la thèse. Y sont rappelés quelques notions et outils fondamentaux utiles.
- Le chapitre 2 fait l'état de l'art sur la détection des ruptures dans les séries chronologiques. Ce chapitre est divisé en deux parties. La première concerne l'estimation des ruptures et leurs localisations. La deuxième concerne les tests d'existence de points de ruptures.
- Le chapitre 3 traite le cas où les fonctions T et V sont connues et lorsqu'elles sont

connues mais dépendant de paramètres inconnus. Dans ce dernier cas, la situation où le paramètre γ_0 est connu et celle où il est inconnu sont étudiées. Dans le cas où il est inconnu, il est estimé par la méthode du maximum de vraisemblance. L'étude du test est basée essentiellement sur la propriété locale asymptotique (LAN) énoncée par exemple dans Dreesbeke et Fine (1996).

- Le chapitre 4 est une généralisation du chapitre 3. Ici, l'amplitude du saut est quelconque et inconnue. Il s'agit donc de tester

$$H_0 : \gamma = \gamma_0 \text{ contre } H^{(n)} = \bigcup_{\beta \in \mathbb{R}^{k+1}} H_{\beta}^{(n)}.$$

Un test de type Cramer-Von-Mises est construit. Les techniques de Ngatchou-Wandji (2009) sont utilisées pour trouver la loi asymptotique du test sous l'hypothèse alternative.

- Le chapitre 5 présente des résultats numériques faits en utilisant le logiciel R. Les résultats obtenus pour des données simulées sont d'abord présentés et commentés. Viennent ensuite ceux des applications à plusieurs jeux de données réelles.

Le chapitre 6 conclut la thèse et dégage les perspectives pour la suite.

Mots clés : Modèle CHARN, Ruptures, Contiguïté, stationnarité par morceaux, Tests optimaux, Normalité Asymptotique Locale.

Summary of the thesis

This thesis focuses on the construction of locally asymptotically optimal tests to detect breaks in the mean of Conditional Heteroskedastic AutoRegressive Nonlinear (CHARN) models described by the following stochastic equation :

$$X_t = T(Z_{t-1}) + \gamma^\top \omega(t) + V(Z_{t-1})\varepsilon_t, \quad t \in \mathbb{Z}, \quad (2)$$

where $\gamma = (\gamma_1, \dots, \gamma_k, \gamma_{k+1})^\top \in \mathbb{R}^{k+1}$ and for $t_1, \dots, t_k, 1 < t_1 < \dots < t_k < n$, $\omega(t) = (\mathbb{1}_{[1, t_1[}(t), \mathbb{1}_{[t_1, t_2[}(t), \dots, \mathbb{1}_{[t_k, n[}(t))^\top \in \{0, 1\}^{k+1}$, $(X_t)_{t \in \mathbb{Z}}$ is a piecewise, ergodic stationary process, $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a standard white noise with density f , for any $t \in \mathbb{Z}$, $Z_t = (X_t, \dots, X_{t-p+1})^\top$, $p \in \mathbb{N}$, and T and V are real functions such as $V(x) > 0, \forall x \in \mathbb{R}^p$ and n the number of observations. The model (2) contains a large class of time series models like AR, MA, ARMA, ARIMA, ARCH etc.

Attention is paid to small breaks. Those which are difficult to observe with the naked eyes, unlike those considered in the literature. Such a study does not appear to have already been carried out in the context of time series. The test studied is the likelihood ratio test to test

$$H_0 : \gamma = \gamma_0 \text{ against } H_\beta^{(n)} : \gamma = \gamma_0 + \frac{\beta}{\sqrt{n}} = \gamma_n, \quad n > 1,$$

for $\gamma_0 \in \mathbb{R}^{k+1}$ and $\beta \in \mathbb{R}^{k+1}$ characterizing respectively the situation where there is no break, and that where there is at least one break to be found.

This document is organized as follows :

- Chapter 1 constitutes the general introduction to the thesis. There, some useful basic concepts and tools are recalled.
- Chapter 2 reviews the state of the art on the detection of breaks in time series. This chapter is divided into two parts. The first concerns the estimation of breaks and their locations. The second concerns the tests for the existence of break-points.
- Chapter 3 deals with the case where the functions T and V are known, and when they are known but depend on unknown parameters. In the latter case, the situation where the parameter γ_0 is known and the one where it is unknown are studied. When it is unknown, it is estimated by the maximum likelihood method. The study of the test is

based essentially on the asymptotic local property (LAN) stated for example in Dreesbeke et Fine (1996).

- Chapter 4 is a generalization of chapter 3. Here, the magnitude of the jump is arbitrary and unknown. Therefore, one has to test

$$H_0 : \gamma = \gamma_0 \text{ against } H^{(n)} = \bigcup_{\beta \in \mathbb{R}^{k+1}} H_{\beta}^{(n)}.$$

A Cramer-Von-Mises type test is constructed. The techniques of Ngatchou-Wandji (2009) are used to find the asymptotic distribution of the test under the alternative hypothesis. Chapter 5 presents numerical results obtained using software R. The results obtained for simulated data are first presented and commented. This is followed by those for applications with several real datasets.

Chapter 6 concludes the thesis and sets out some prospects.

Keywords : CHARN model, Breaks, Contiguity, Piecewise stationarity, Optimal tests, Local Asymptotic Normality.

Table des matières

1	Introduction et outils fondamentaux	17
1.1	Introduction générale	17
1.1.1	Contributions	19
1.1.2	Structure de la thèse	19
1.2	Outils fondamentaux	20
1.2.1	Rappels sur les tests de vraisemblance	21
1.2.2	Quelques rappels sur les séries chronologiques	24
1.2.3	Normalité asymptotique locale et contiguïté	27
1.2.4	Puissance locale asymptotique des tests	31
1.2.5	Théorème limite centrale des martingales	32
1.2.6	Convergence faible dans l'espace $\mathbb{C}([0, 1], \mathbb{R})$	33
1.2.7	Espaces de Fréchet	35
2	État de l'art	37
2.1	Modèle CHARN	37
2.2	Estimation des points de ruptures	38
2.3	Tests d'existence de points de ruptures	43
2.3.1	Tests généraux	43
2.3.2	Le CUSUM test	50
2.3.3	La segmentation binaire	60
3	Détection des ruptures et estimation de leurs localisations : utilisation d'un test du rapport de vraisemblance	63
3.1	Introduction	63
3.1.1	Hypothèses générales	64
3.1.2	Détection des ruptures	65
3.2	Cas où les fonctions T , V et la densité du bruit sont connues	69
3.3	Cas où les fonctions T et V sont paramétriques et la densité du bruit est connue	78

3.3.1	Cas où γ_0 est connue	86
3.3.2	Cas où γ_0 est inconnu	109
3.4	Conclusion	117
4	Détection des ruptures et leurs localisations par un test du type Cramér-Von-Mises	119
4.1	Introduction	119
4.2	Détection des ruptures	120
4.3	Convergence faible du processus de statistique centrale	120
4.3.1	Cas où γ_0 est connue	121
4.3.2	Cas où γ_0 est inconnu	128
4.4	Conclusion	140
5	Applications	141
5.1	Introduction	141
5.2	Simulation	141
5.2.1	Étude de la puissance	141
5.2.2	Détection des ruptures	145
5.3	Application aux données réelles	149
5.3.1	Modélisation	150
5.3.2	Anomalie climatique annuelle	150
5.3.3	COVID-19 en France	151
5.4	Quelques comparaisons	154
5.4.1	Comparaison avec Horváth <i>et al.</i> (2020) et quelques méthodes de segmentation binaire	154
5.4.2	Comparaison avec Xiong <i>et al.</i> (2015)	155
5.4.3	Comparaison avec Bardet et Kengne (2014) et Fryzlewicz et Rao (2014)	161
5.4.4	Comparaison avec Fotopoulos <i>et al.</i> (2009)	165
6	Conclusion et perspectives	197
6.1	Conclusion générale	197
6.2	Perspectives	198
	Bibliographie	199

Table des figures

5.1	Chronogrammes des observations	168
5.2	Puissance locale du test pour $k = 1$ et $\alpha = 0.05$	169
5.3	Puissance locale du test pour $k = 1$ et $\alpha = 0.01$	170
5.4	Puissance locale du test pour $k = 1$ et $\alpha = 0.1$	171
5.5	Puissance locale du test pour $k = 1$ et $\alpha = 0.05$	172
5.6	Puissance locale du test pour $k = 1$ et $\alpha = 0.01$	173
5.7	Puissance locale du test pour $k = 1$ et $\alpha = 0.1$	174
5.8	le cas de non-changement pour $\alpha = 0.05$	175
5.9	Le cas de non-changement pour $\alpha = 0.01$	175
5.10	le cas de non-changement pour $\alpha = 0.1$	176
5.11	Détection des ruptures pour $k = 1$ et $\alpha = 0.05$	177
5.12	Détection des ruptures pour $k = 1$ et $\alpha = 0.01$	178
5.13	Détection des ruptures pour $k = 1$ et $\alpha = 0.1$	179
5.14	Détection des ruptures pour $k = 2$ et $\alpha = 0.05$	180
5.15	Détection des ruptures pour $k = 2$ et $\alpha = 0.01$	181
5.16	Détection des ruptures pour $k = 2$ et $\alpha = 0.1$	182
5.17	Détection des ruptures pour $k = 1$	183
5.18	Une rupture connue contre la présence d'une rupture faible supplémentaire pour $\alpha = 0.05$	183
5.19	Une rupture connue contre la présence d'une rupture faible supplémentaire pour $\alpha = 0.01$	184
5.20	Détection des ruptures pour $k = 1$ et $\alpha = 0.05$	185
5.21	Détection des ruptures pour $k = 1$ et $\alpha = 0.01$	186
5.22	Anomalies climatiques annuelles mondiales de 1880 à 2015.	187
5.23	Taux de mortalité du COVID-19 en France du 15/02/2020 au 10/07/2020.	188
5.24	Nombre de décès quotidiens du COVID-19 en France du 27/02/2020 au 10/07/2020.	189
5.25	Nouveaux cas quotidiens du COVID-19 en France du 25/02/2020 au 10/07/2020.	190

5.26	La rivière Han jiang de Chine.	191
5.27	Les données des crues univariées composées de la AMMDD (Q_1) de 1950 à 2011.	192
5.28	Les données du volume annuel maximal des crues sur 3 jours (V_3) et du volume maximal annuel des crues sur 15 jours (V_{15}) de 1950 à 2011.	193
5.29	Clôture quotidienne de l'indice boursier FTSE 100 du 02/01/2004 au 30/04/2013.	194
5.30	Les données journalières des rendements logarithmiques sur l'action Verizon communication (VZ) en 2002.	195

Chapitre 1

Introduction et outils fondamentaux

1.1 Introduction générale

Beaucoup de données sont issues des phénomènes susceptibles de varier avec le temps. C'est par exemple le cas des données financières, hydrologiques, climatiques, économiques... Il est souvent nécessaire de savoir si ces données évoluent de manière stationnaire ou pas. En d'autres termes, y en a-t-il parmi elles qui deviennent de la stationnarité? De telles données sont appelées ruptures. De façon mathématique, une rupture dans une série de variables aléatoires peut se définir comme un changement dans leur loi de probabilité. Il existe plusieurs types de ruptures moins forts que les changements dans la loi. On pense entre autres, aux changements dans la moyenne, dans la variance, dans la corrélation... Dans cette thèse, on s'intéresse à l'étude des faibles ruptures, celles que nous ne pouvons pas voir à l'œil nu (voir par exemple la Figure 5.3). En santé, celles-ci peuvent être des signaux annonciateurs de maladies. En finance, elles peuvent annoncer une crise financière. En climatologie, elles peuvent être des signes précurseurs d'une tempête, d'une sécheresse, d'une inondation ou encore d'une canicule.

Les premiers résultats sur la détection de rupture datent de Page (1955), lorsqu'il utilisait les sommes partielles pour tester les changements dans la moyenne des observations indépendantes. Depuis lors, les recherches sur ce sujet ont considérablement avancé et ont été étendues aux séries chronologiques. Une liste non-exhaustive d'articles sur ce sujet comprend Härdle et Tsybakov (1997), Härdle *et al.* (1998), Bardet *et al.* (2009), Döring (2010, 2011), Ciuperca (2011), Bardet et Kengne (2014), Amano (2012), Yang et Song (2014), Kengne (2012), Chen *et al.* (2011), Dehling *et al.* (2013), Dehling *et al.* (2017), Wang *et al.* (2012), Francq et Zakoïan (2012), Bardet *et al.* (2012), Zhou *et al.* (2014), Fotopoulos *et al.* (2009), Huh (2010), Enikeeva *et al.* (2018), Gombay et Serban (2009). On trouve aussi des méthodes basées sur l'optimisation comme Lavielle (1998), Hawkins

(2001) et Menne et Williams Jr (2005). Des détails sur ces articles sont présentés au chapitre 2.

L'objectif de cette thèse est de proposer une méthode pour détecter les changements dans la moyenne d'une série chronologique stationnaire par morceaux. La principale nouveauté de ce travail par rapport à ceux qui existent est que nous cherchons à détecter les faibles ruptures. On ne trouve que quelques travaux abordant l'étude de ces ruptures. Ce sont par exemple Fotopoulos *et al.* (2009), Khakhubia (1987), Bhattacharya et Zhou (2017) qui les étudient dans le contexte de données indépendantes.

Dans cette thèse, nous considérons une série d'observations X_1, X_2, \dots, X_n générée par le modèle CHARN(p) suivant

$$X_t = T(Z_{t-1}) + \gamma^\top \omega(t) + V(Z_{t-1})\varepsilon_t, \quad t \in \mathbb{Z}, \quad (1.1)$$

où $\gamma = (\gamma_1, \dots, \gamma_k, \gamma_{k+1})^\top \in \mathbb{R}^{k+1}$ et pour $t_1, \dots, t_k, 1 < t_1 < \dots < t_k < n$, $\omega(t) = (\mathbb{1}_{[1, t_1[}(t), \mathbb{1}_{[t_1, t_2[}(t), \dots, \mathbb{1}_{[t_k, n[}(t))^\top \in \{0, 1\}^{k+1}$, $(X_t)_{t \in \mathbb{Z}}$ est un processus stationnaire par morceaux et ergodique, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc centré réduit de densité f , pour tout $t \in \mathbb{Z}$, $Z_t = (X_t, \dots, X_{t-p+1})^\top$, $p \in \mathbb{N}$, et T et V sont des fonctions réelles telles que $\inf_{x \in \mathbb{R}^p} V(x) > 0$.

Nous construisons un test du rapport de vraisemblance de :

$$H_0 : \gamma = \gamma_0 \text{ contre } H_\beta^{(n)} : \gamma = \gamma_0 + \frac{\beta}{\sqrt{n}} = \gamma_n, \quad n > 1,$$

pour $\gamma_0 \in \mathbb{R}^{k+1}$ et $\beta \in \mathbb{R}^{k+1}$ dépend de t . Ces paramètres caractérisent respectivement la situation où il n'y a pas de changement et celle où il y a au moins un changement à déterminer.

L'un des avantages de ce test est qu'on peut prendre $\gamma_0 \neq 0$. Dans ce cas, on peut tester une hypothèse de changements connus contre des changements connus plus d'autres changements faibles inconnus. C'est-à-dire que sous l'hypothèse H_0 , la série présente des risques connus et sous $H_\beta^{(n)}$, la série présente au moins un risque supplémentaire inconnu qui peut être très faible (voir Figure 5.1 (f)). À notre connaissance, ceci n'est pas encore étudié dans la littérature. En effet, dans les travaux existants, les auteurs cherchent toujours à tester une hypothèse de non-changement contre une hypothèse de changement. Comme exemples, on peut citer Chen *et al.* (2011), Zhou *et al.* (2014), Sen et Srivastava (1975) et Gardner (1969). Par ailleurs, l'utilisation de test du rapport de vraisemblance est une nouvelle approche dans le contexte des séries chronologiques stationnaires par morceaux du type CHARN. Nous montrons dans cette thèse comment l'utiliser pour détecter les changements et leurs localisations dans les séries chronologiques.

Notre méthode s'appuie sur le chronogramme de la série duquel on peut obtenir des informations a priori comme par exemple les potentiels points de rupture. La suite de la

méthode consiste à déterminer la puissance locale du test et à en déterminer un estimateur. S'il existe au moins un point pour lequel l'estimation de la puissance est supérieure au seuil de signification du test, cela implique que la série comporte une rupture. Si cette estimation est maximale en des points, ceux-ci sont considérés comme localisations des ruptures.

Le chronogramme nous aide à localiser les plages où il y a plus de chance de trouver des ruptures. Cela évite une recherche exhaustive coûteuse comme c'est souvent le cas avec les méthodes classiques, qui, d'ailleurs, peuvent ne pas détecter les petites ruptures qui nous intéressent. Cependant, nous avons observé au cours des simulations que pour des ruptures de très petites amplitudes le graphique ne nous renseignait pas toujours sur les possibles points de rupture potentiels. Dans ce cas nous avons expérimenté la méthode comme si les ruptures étaient connues.

Il existe plusieurs situations où on peut être amené à rechercher des ruptures locales. C'est par exemple le cas des données liées à des événements qui se produisent à une date connue. Il est clair que l'approche et/ou la fin de cet événement peut avoir des incidences sur la dynamique des données. Il peut alors être très utile de savoir quand les changements débutent ou se terminent.

1.1.1 Contributions

Nos travaux de recherche contribuent à proposer des méthodes et stratégies pour détecter les faibles changements dans la moyenne des modèles CHARN. Notre première contribution est l'étude d'un test du rapport de vraisemblance pour détecter ces faibles ruptures. Nous établissons un résultat de normalité locale asymptotique (LAN) duquel nous déduisons l'optimalité asymptotique du test, et une expression explicite de sa puissance locale. Cette dernière est exprimée comme fonction des localisations des points de rupture et de leurs amplitudes. Les localisations des points de rupture estimés sont obtenues en maximisant une estimation de la puissance locale.

Une autre contribution est l'étude d'un test du type Cramer-Von-Mises pour tester un modèle CHARN à coefficients dépendants du temps, contre une classe contiguë de modèles CHARN. Ce travail qui comporte un aspect bayésien, est motivé par l'envie de généraliser notre première contribution.

1.1.2 Structure de la thèse

Cette thèse est structurée de la manière suivante : nous donnons dans le premier chapitre les définitions et les outils fondamentaux que nous utilisons. Dans le deuxième chapitre, nous présentons quelques travaux liés à la détection des ruptures dans les séries chrono-

logiques. Dans le chapitre 3, nous traitons les cas où les fonctions T et V sont connues et lorsqu'elles sont connues mais dépendant de paramètres inconnus. Dans la dernière situation, nous étudions le cas où le paramètre γ_0 est connu et celui où il est inconnu. Notons qu'en pratique les fonctions T et V et le paramètre γ_0 sont inconnus. Donc l'étude des cas où ils sont connus sont juste des outils à utiliser dans le cas où ils sont inconnus. La puissance locale du test joue un rôle déterminant dans notre méthode. Aussi, nous donnons son expression explicite. Lorsque T et V sont connues, nous montrons que le logarithme du rapport de vraisemblance vérifie la propriété de la normalité locale asymptotique (LAN) énoncée dans Dreesbeke et Fine (1996). Puis, à l'aide du troisième Lemme de Le Cam, nous montrons que les deux hypothèses du test sont contiguës. Ceci est une conséquence de la normalité locale asymptotique. Pour les autres scénarios considérés, nous montrons que la statistique du test dans chaque cas est équivalente en probabilité à la statistique du test du premier cas étudié. Donc asymptotiquement, elles ont les mêmes propriétés asymptotiques. Dans chacun de ces cas considérés, nous trouvons une expression explicite de la puissance locale du test. Nous signalons que de tels résultats sont rares dans la littérature, sauf peut-être Dehling *et al.* (2017) qui étudient la puissance locale des tests CUSUM et Wilcoxon dans le contexte de séries chronologiques stationnaires.

Dans le chapitre 4, nous développons un nouveau test pour tester un modèle CHARN contre un ensemble de modèle CHARN. Le but de cette étude est de détecter les ruptures dans le cas où l'amplitude du saut est quelconque et inconnue. Il s'agit pour nous de tester H_0 contre $H^{(n)} = \bigcup_{\beta \in \mathbb{R}^{k+1}} H_{\beta}^{(n)}$. Pour cette étude nous considérons un test de type Cramer-Von-Mises basée sur le processus de vraisemblance dont nous étudions la convergence faible dans un espace de Fréchet approprié. Pour plus des détails le lecteur peut consulter Henze et Wagner (1997), Csörgő (1983), Meintanis *et al.* (2015) et Ngatchou-Wandji (2009).

Après ces études théoriques, dans le chapitre 5, nous présentons les résultats d'une étude de simulation effectuée afin d'évaluer les performances de nos méthodes que nous appliquons par ailleurs à des données réelles.

De nombreuses questions sont soulevées au cours de notre travail. Des pistes à explorer dans des études futures sont finalement évoquées au chapitre 6.

1.2 Outils fondamentaux

Cette partie regroupe quelques notions et outils fondamentaux que nous utilisons dans la suite. Il s'agit des tests basés sur la vraisemblance, des séries chronologiques, des modèles non-linéaires, des modèles à coefficients variants avec le temps, de la normalité asymptotique locale (LAN), de la convergence faible et d'autres notions de probabilités

et statistiques.

1.2.1 Rappels sur les tests de vraisemblance

Nous présentons dans ce qui suit des rappels concernant les tests de vraisemblance. Il est à noter qu'il s'agit de tests paramétriques. Pour plus de détails, se référer par exemple à Monfort (1996) ou Lehmann et Romano (2006).

On commence par introduire la fonction de vraisemblance.

Soit (X_1, \dots, X_n) une suite d'observations dont la loi mère appartient à une famille paramétrique de densités (ou fonctions de probabilité) $\{f(x, \theta); \theta \in \Theta\}$, où $\Theta \subseteq \mathbb{R}^k$. On appelle fonction de vraisemblance de θ pour une réalisation donnée (x_1, x_2, \dots, x_n) de l'échantillon, la fonction suivante :

$$L(x_1, \dots, x_n, \theta) = \prod_{i=1}^n f(x_i, \theta).$$

Les tests de vraisemblance classiques sont le test de Wald, du score et du rapport de vraisemblance. Ce sont des tests asymptotiques et convergents. Ils permettent de tester $H_0 : g(\theta) = 0$ contre $H_1 : g(\theta) \neq 0$, où θ appartient à un ouvert de \mathbb{R}^k , $k \in \mathbb{N}$, et g une fonction vectorielle de classe C^1 . Par exemple, si on veut tester $\theta = \theta_0$ contre $\theta \neq \theta_0$, on prend $g(\theta) = \theta - \theta_0$.

Dans la suite, on note $L_n(\theta)$ la vraisemblance des observations et $\hat{\theta}_n$ l'estimateur du maximum de vraisemblance(EMV) de θ .

1.2.1.1 Test de Wald

L'idée du test de Wald est de comparer $g(\hat{\theta}_n)$ et $g(\theta)$. On va donc chercher la loi de la différence, ou, plus précisément, de la norme de la différence (on se ramène ainsi à quelque chose qui est toujours positif). Si on a la normalité asymptotique suivante :

$$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{loi} \mathcal{N}(0, I^{-1}(\theta)),$$

où $I(\theta)$ est l'information de Fisher, alors si g est différentiable, on a par la méthode dite delta,

$$\sqrt{n} \left(g(\hat{\theta}_n) - g(\theta) \right) \xrightarrow{loi} \mathcal{N}(0, \Sigma(\theta)),$$

où

$$\Sigma(\theta) := \frac{\partial g(\theta)}{\partial \theta^\top} I^{-1}(\theta) \frac{\partial g^\top(\theta)}{\partial \theta}.$$

Ainsi, en notant $\Sigma(\theta)^{1/2}$ la matrice telle que $\Sigma(\theta)^{1/2} \Sigma(\theta)^{1/2} = \Sigma(\theta)$, et $\Sigma(\theta)^{-1/2}$ son inverse, on a :

$$\sqrt{n} \Sigma(\theta)^{-1/2} \left(g(\hat{\theta}_n) - g(\theta) \right) \xrightarrow{loi} \mathcal{N}(0, Id),$$

d'où l'on peut en déduire que

$$n \left(g(\widehat{\theta}_n) - g(\theta) \right)^\top \Sigma(\theta)^{-1/2} \left(g(\widehat{\theta}_n) - g(\theta) \right) \xrightarrow{\text{loi}} \chi^2(\nu),$$

avec $\nu = \text{dimension du vecteur } g(\theta)$.

Comme θ est inconnu, on peut le remplacer par $\widehat{\theta}_n$. Ceci est possible car pour un modèle suffisamment régulier, la fonction $\theta \mapsto \Sigma(\theta)$ est continue, et l'EMV est convergent. Le théorème de Slutsky nous assure alors que l'on a toujours la même loi limite :

$$n \left(g(\widehat{\theta}_n) - g(\theta) \right)^\top \Sigma(\widehat{\theta}_n)^{-1} \left(g(\widehat{\theta}_n) - g(\theta) \right) \xrightarrow{\text{loi}} \chi^2(\nu).$$

Théorème 1.2.1. (*Test de Wald*)

La statistique du test est la suivante :

$$T_n = n \left(g(\widehat{\theta}_n) - g(\theta) \right)^\top \Sigma(\widehat{\theta}_n)^{-1} \left(g(\widehat{\theta}_n) - g(\theta) \right)$$

qui converge sous H_0 vers $\chi^2(\nu)$, où $\nu = \text{dimension du vecteur } g(\theta)$.

En particulier, quand $g(\theta) = \theta - \theta_0$, sous H_0 , on a

$$T_n = n \left(\widehat{\theta}_n - \theta \right)^\top I^{-1}(\widehat{\theta}_n) \left(\widehat{\theta}_n - \theta \right) \xrightarrow{\text{loi}} \chi^2(\nu),$$

où $\nu = \text{dimension de } \theta$.

On peut remarquer que la statistique de test diverge vers l'infini quand H_0 n'est pas satisfaite. La région de rejet est donc

$$W_n = \{x = (x_1, \dots, x_n) / T_n \geq q_{1-\alpha}(\chi^2(\nu))\},$$

où $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\chi^2(\nu)$.

Le théorème de Slutsky ci-dessus permet de remplacer $\Sigma(\theta)$ par $\Sigma(\widehat{\theta}_n)$. On aurait pu aussi le remplacer par tout autre estimateur $\widehat{\Sigma}(\widehat{\theta})$ de $\Sigma(\theta)$. Dans le cas où $g(\theta) = \theta - \theta_0$ par exemple, on peut remplacer $I(\theta)$ par son plug-in estimateur empirique

$$I(\widehat{\theta}_n) = -\frac{1}{n} \sum_{i=1}^n \frac{\partial^2}{\partial \theta \partial \theta^\top} \log[f(X_i, \widehat{\theta}_n)].$$

1.2.1.2 Test du multiplicateur de Lagrange ou test du score

L'idée est de comparer les EMV non contraint, $\widehat{\theta}_n$ et contraint par H_0 , $\widehat{\theta}_n^0$. En introduisant les multiplicateurs de Lagrange λ (vecteur de dimension égale au nombre de contraintes) pour gérer la contrainte $H_0 : g(\theta) = 0$, on trouve la relation suivante :

$$\frac{\partial L_n}{\partial \theta}(\widehat{\theta}_n^0) + \lambda^\top \frac{\partial g}{\partial \theta}(\widehat{\theta}_n^0) = 0$$

qui nous permet de transformer le test de Wald et d'avoir une statistique qui s'écrit en fonction des vecteurs de score, $\partial_\theta \log L_n(\hat{\theta}_n)$ et $\partial_\theta \log L_n(\hat{\theta}_n^0)$ et non plus des différences $g(\hat{\theta}_n) - g(\theta)$.

Théorème 1.2.2. (Test du score pour $g(\theta) = \theta - \theta_0$)

La statistique du test est la suivante :

$$T'_n = \frac{1}{n} \frac{\partial \log L_n(\hat{\theta}_n^0)}{\partial \theta^\top} I^{-1}(\hat{\theta}_n^0) \frac{\partial \log L_n(\hat{\theta}_n^0)}{\partial \theta}$$

qui converge sous H_0 vers $\chi^2(\nu)$.

La région de rejet est donc :

$$W_n = \{x = (x_1, \dots, x_n) / T'_n \geq q_{1-\alpha}(\chi^2(\nu))\},$$

où $q_{1-\alpha}$ est le quantile d'ordre $1 - \alpha$ de la loi $\chi^2(\nu)$.

On peut remarquer que le test de Wald nécessite uniquement le calcul de l'EMV non contraint, tandis que celui du score nécessite uniquement le calcul de l'EMV contraint. Selon le problème posé, on pourra donc judicieusement choisir l'un ou l'autre.

On montre (voir Monfort, 1996, chapitre 20) que T_n et T'_n ont asymptotiquement les mêmes lois. De même que précédemment, il est possible de remplacer l'information de Fisher par son équivalent empirique.

1.2.1.3 Test du rapport de vraisemblance

Une troisième idée est de comparer les niveaux des vraisemblances ou de façon équivalentes des log-vraisemblances $\log L_n(\theta_0)$ et $\log L_n(\hat{\theta})$. Si le rapport est proche de 1, alors les deux quantités sont équivalentes. Sinon, il est peu vraisemblable que θ_0 puisse correspondre au maximum de la vraisemblance théorique.

On appelle **test du rapport de vraisemblance** de l'hypothèse $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$ au niveau α , le test est défini par la région de rejet de la forme :

$$R = \frac{L_0}{L_1} = \frac{L(x_1, \dots, x_n, \theta_0)}{L(x_1, \dots, x_n, \theta_1)} < k,$$

où k est une valeur positive déterminée en fonction du risque α .

1.2.1.4 Optimalité

Définition 1.2.1. On dit qu'un test τ_1 est **plus puissant** qu'un test τ_2 au niveau α , si τ_2 est de niveau égal (ou inférieur) à α et s'il a une puissance inférieure à celle de τ_1 .

Définition 1.2.2. On dit que le test τ est le **plus puissant (ou optimal)** de niveau α s'il est plus puissant que tout autre test de niveau α .

Théorème 1.2.3. (Lemme de Neyman-Pearson)

Le test du rapport de vraisemblance est le plus puissant quel que soit le choix de $\alpha \in]0, 1[$.

1.2.2 Quelques rappels sur les séries chronologiques

1.2.2.1 Notions de séries chronologiques

Définition 1.2.3. On appelle *série chronologique (série temporelle, chronique)* une suite d'observations numériques d'une grandeur effectuées au cours du temps.

Les séries chronologiques sont généralement étudiées à travers des modèles statistiques tels que les modèles ARMA, ARCH, GARCH, CHARN etc.

• Stationnarité

La notion de stationnarité est très importante dans l'analyse et l'étude théorique des séries chronologiques. On distingue deux principaux types.

Définition 1.2.4. Soit la série temporelle (X_t)

• Elle est stationnaire au second ordre si :

i. $\forall t \in \mathbb{Z}, \mathbb{E}(|X_t|^2) < \infty$.

ii. $t \mapsto \mathbb{E}(X_t)$ est constante.

iii. $\forall s, t, h \in \mathbb{Z}, \text{Cov}(X_s, X_t) = \text{Cov}(X_{s+h}, X_{t+h})$.

• Elle est stationnaire au sens strict si les distributions de probabilité conjointes de X_{t_1}, \dots, X_{t_k} et de $X_{t_1+h}, \dots, X_{t_k+h}$ sont les mêmes pour tout $k \in \mathbb{N}^*$, pour tout $t_1, \dots, t_k \in \mathbb{Z}^k$ et pour tout $h \in \mathbb{Z}$.

Remarque 1. La stationnarité au sens strict implique la stationnarité au second ordre pour les processus du second ordre. C'est le cas des séries chronologiques du second ordre.

• Ergodicité

La notions d'ergodicité est très utile dans notre travail. Nous donnons ci-dessous une définition de cette notion.

Définition 1.2.5. Un processus stationnaire est dit *ergodique* si pour toute fonction borélienne f de \mathbb{R}^k dans \mathbb{R} , lorsque $n \rightarrow \infty$, on a

$$\frac{1}{n} \sum_{i=1}^n f(X_{t_1+i}, \dots, X_{t_k+i}) \xrightarrow{p.s.} \mathbb{E}[f(X_{t_1}, \dots, X_{t_k})].$$

Remarque 2. *On peut voir l'ergodicité comme une généralisation de la loi des grands nombres.*

- **Non-stationnarité**

La plupart des séries économiques sont non-stationnaires, c'est-à-dire que le processus qui les décrit ne vérifie pas au moins une des conditions de la définition d'un processus stationnaire du second ordre. Ceci a conduit à définir plusieurs types de non-stationnarité.

- **Non-stationnarité déterministe**

Définition 1.2.6. *On dit que le processus $(X_t)_t$ est non-stationnaire déterministe, s'il peut s'écrire sous la forme suivante :*

$$X_t = f(t) + Z_t,$$

où f est une fonction déterministe qui dépend du temps et $(Z_t)_t$ est un processus stationnaire, le bruit blanc étant un cas particulier.

Par exemple on peut prendre $f(t) = a + bt$, où $(a, b) \in \mathbb{R}^2$, c'est le cas couramment retenu, mais on pourrait aussi considérer entre autres une tendance quadratique $f(t) = a + bt + ct^2$, où $(a, b, c) \in \mathbb{R}^3$, ou une tendance segmentée comme le cas de Ciuperca (2011). Parmi les travaux qui ont étudié ce type de non-stationnarité, on a Lavielle et Lebarbier (2001) et Lebarbier (2005).

- **Stationnarité par morceaux**

Dans notre travail, nous nous intéressons à un autre type de non-stationnarité, dite par morceaux, quelque peu différente de celle introduite par Dahlhaus *et al.* (1997). Un processus $(X_t)_t$ est dit stationnaire par morceaux si sur son intervalle d'observation il est stationnaire sur des sous intervalles. Il est à noter qu'un tel processus est non stationnaire, s'il a des lois différentes sur au moins deux sous-intervalles de son intervalle d'observation.

- **Non-stationnarité au sens de Dahlhaus**

Dahlhaus *et al.* (1997) considère un processus asymptotiquement stationnaire. Il suppose que lorsque le nombre des observations tend vers l'infini, la série devient stationnaire. Cet article est une généralisation de Whittle (1953) pour les processus stationnaires. Dahlhaus *et al.* (1997) définit les processus non-stationnaires comme un processus ayant une représentation spectrale de la forme suivante :

$$X_{t,T} = \mu \left(\frac{t}{T} \right) + \int_{-\pi}^{\pi} e^{i\lambda t} A_{t,T}^0(\lambda) d\xi(\lambda),$$

où $X_{t,T}(t = 1, \dots, T)$ est la série d'observations, T est le nombre des observations, ξ est un processus stochastique dans $[-\pi, \pi]$ et $A_{t,T}^0(\lambda) = A(t/T, \lambda)$ telle que $A : [0, 1] \times \mathbb{R} \rightarrow \mathbb{C}$ est une fonction 2π -périodique avec $A(u, -\lambda) = \overline{A(u, \lambda)}$ et pour tout T , $A(u, \lambda)$ et $\mu(u)$ sont continues en u , \bar{x} est le conjugué de $x \in \mathbb{C}$.

Remarque 3. *Un autre type de non-stationnarité locale est considéré dans Ngatchou-Wandji et Harel (2013) et Ngatchou-Wandji et al. (2019). Elle a en commun avec la non-stationnarité au sens de Dahlhaus et al. (1997) la propriété que toute série les vérifiant sont stationnaires à partir d'un certain rang.*

1.2.2.2 Processus ARIMA et SARIMA

Parmi les processus non-stationnaires les plus connus, on trouve les modèles ARIMA et SARIMA.

Définition 1.2.7. *On dit qu'un processus $(X_t)_{t \in \mathbb{N}}$ est un processus ARIMA d'ordre (p, d, q) , noté $ARIMA(p, d, q)$, si :*

$$\Phi(B)\nabla^d X_t = \Theta(B)\varepsilon_t \quad t \in \mathbb{Z}.$$

On dit qu'un processus $(X_t)_{t \in \mathbb{N}}$ est un processus SARIMA d'ordre $(p, d, q)(P, D, Q)_s$, noté $SARIMA(p, d, q)(P, D, Q)_s$ si :

$$\Phi(B)\Phi'(B^s)\nabla^d\nabla_s^D X_t = \Theta(B)\Theta'(B^s)\varepsilon_t \quad t \in \mathbb{Z}.$$

Dans les deux définitions,

- $\nabla^d = (1 - B)^d$
- $\nabla_s^D = (1 - B^s)^D$
- $\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$, où $\varphi = (\varphi_1, \dots, \varphi_p) \in \mathbb{R}^p$ et $\varphi_p \neq 0$
- $\Phi'(B) = 1 - \varphi'_1 B - \dots - \varphi'_P B^P$, où $\varphi' = (\varphi'_1, \dots, \varphi'_P) \in \mathbb{R}^P$ et $\varphi'_P \neq 0$
- $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, où $\Theta = (\theta_1, \dots, \theta_q) \in \mathbb{R}^q$ et $\theta_q \neq 0$
- $\Theta'(B) = 1 - \theta'_1 B - \dots - \theta'_Q B^Q$, où $\Theta' = (\theta'_1, \dots, \theta'_Q) \in \mathbb{R}^Q$ et $\theta'_Q \neq 0$
- B l'opérateur de retard.

Un processus stochastique est dit linéaire s'il peut s'écrire comme somme pondérée bruit blanc, soit :

$$X_t = \varrho + \sum_{i=0}^{+\infty} \alpha_i \varepsilon_{t-i},$$

où ϱ est une constante, α_i des coefficients réels avec $\alpha_0 = 1$ et $(\varepsilon_t)_t$ des v.a iid, centrées et de variance $\sigma_\varepsilon^2 < +\infty$ et telles que $\mathbb{E}(\varepsilon_t | X_{t-1}, X_{t-2}, \dots) = 0$.

On cite comme exemples des modèles linéaires, les modèles ARMA, ARIMA, SARIMA

ci-dessous (voir par exemple Brockwell *et al.* (2002)).

L'écriture ci-dessus peut s'écrire aussi comme suit :

$$X_t = h(\varepsilon_t, \varepsilon_{t-1}, \varepsilon_{t-2}, \dots),$$

structure dans laquelle on impose la linéarité de la fonction h .

Il y a de nombreuses manières d'introduire les modèles non-linéaires de séries temporelles. La plus simple est de prendre la dernière écriture ci-dessus dans le cas où h est une fonction non-linéaire. Parmi les exemples les plus connus des modèles non-linéaires, on a les modèles bilinéaires proposés par Granger et Andersen (1978), les modèles TARARCH et EXPAR étudiés dans par exemple Guégan *et al.* (1994) et Tong (1990), les modèles ARCH(p) définis par exemple de la manière suivante :

$$\varepsilon_t = \sigma_t z_t,$$

où ε_t représente les innovations de la série, z_t est un bruit blanc centré réduit et $\sigma_t^2 = a_0 + \sum_{i=1}^p a_i \varepsilon_{t-i}^2$, avec les a_i , $i = 0, \dots, p$, des réelles tels que $a_0 > 0$, $a_1 \geq 0$ et $a_p \geq 0$ garantissant la stricte positivité de la variance conditionnelle. Pour plus des détails voir Engle (1982).

1.2.2.3 Les modèles à coefficients variant avec le temps

Beaucoup de phénomènes de la vie, comme par exemple les changements climatiques, la variation d'une machine dans une chaîne de fabrication, peuvent être modélisés par des modèles statistiques ayant des paramètres qui changent au cours du temps. Ces modèles sont appelés modèles à coefficients variant avec le temps. Le modèle (1.1) que nous étudions dans cette thèse et les modèles présentés dans le paragraphe précédent en sont des exemples. De tels modèles sont largement étudiés dans la littérature. Quelques références sur le sujet sont Badagián *et al.* (2015), Davis *et al.* (1995), Dehling *et al.* (2013) et Gardner (1969).

1.2.3 Normalité asymptotique locale et contiguïté

Dans cette partie, on présente une notion fondamentale sur laquelle s'appuie une grande partie de notre travail. On commence par présenter la normalité asymptotique locale (LAN) issue de la théorie de Le Cam (1986). Il existe plusieurs versions. Nous adoptons ici la formulation donnée par Dreesbeke et Fine (1996) avec laquelle nous avons travaillé. Puis, nous présentons les propriétés liées à la contiguïté entre des mesures de probabilités, une forme d'absolue continuité asymptotique.

1.2.3.1 Notations et définitions

Nous présentons ici les principales définitions et les principaux résultats associés à la notion de la normalité asymptotique locale.

On suppose qu'on a $\mathbf{X}^{(n)} = (X_1, \dots, X_n)$, $n \in \mathbb{N}$, suite d'observations décrites par la suite de modèles statistiques (ou expériences statistiques) $\xi_\theta^{(n)} = (\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)})$, où $\mathcal{P}^{(n)} = \{P_\theta^{(n)}, \theta \in \Theta\}$, Θ est un ouvert non vide de \mathbb{R}^k , une famille paramétrique des lois de probabilité définie sur $(\mathbb{R}^n, \mathcal{B}^n)$ indexée d'une façon injective c'est-à-dire pour tout $(\theta, \theta') \in \Theta \times \Theta$, on a :

$$\theta \neq \theta' \implies P_\theta^{(n)} \neq P_{\theta'}^{(n)}.$$

où \mathcal{B} est la tribu borélienne.

Autrement dit, observer une suite de variables aléatoires X_1, \dots, X_n sur l'espace mesurable $(\mathbb{R}, \mathcal{B})$ revient à observer le n -uplet $\mathbf{X}^{(n)} = (X_1, \dots, X_n)$ de loi de probabilité $\mathcal{P}^{(n)}$ sur l'espace $(\mathbb{R}^n, \mathcal{B}^n)$. Donc l'observation $\mathbf{X}^{(n)}$ est une v.a. réelle de loi $\mathcal{P}^{(n)}$.

Pour tout $\theta \in \Theta$, on définit des suites locales de voisinages $\theta^{(n)}$ de θ qui convergent vers θ lorsque $n \rightarrow \infty$. Ces suites locales sont de la forme

$$\theta^{(n)} = \theta + \nu(n)\tau, \quad \text{où } \nu(n) \text{ est une matrice } k \times k \text{ et } \tau \in \mathbb{R}^k. \quad (1.2)$$

On suppose que lorsque n est assez grand, $\theta^{(n)} \in \Theta$. Nous supposons aussi que la matrice $\nu(n)$ est inversible telle que $\|\nu(n)\| \rightarrow 0$ lorsque $n \rightarrow \infty$, où nous rappelons que $\|\cdot\|$ est la norme euclidienne. Ainsi, de façon plus générale,

$$\theta^{(n)} = \theta + \nu(n)\tau^{(n)}, \quad \text{où } \sup [(\tau^{(n)})'\tau^{(n)}] < +\infty. \quad (1.3)$$

La suite d'expériences $\xi^{(n)}(\theta) = (\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{(n)}(\theta))$, est appelée suite d'expériences locales (localisée en θ). Selon le problème à étudier, il y a quelques exemples sur le choix de $\nu(n)$ dans Dreesbeke et Fine (1996).

Prenons maintenant deux mesures de probabilité P et Q sur l'espace mesurable $(\mathcal{X}, \mathcal{A})$. D'après la décomposition de Lebesgue, il existe une unique fonction f non négative \mathcal{A} -mesurable et un unique événement N avec $P(N) = 0$, tels que

$$Q(A) = \int_A f dP + Q(A \cap N), \quad A \in \mathcal{A}. \quad (1.4)$$

L'égalité (1.4) est la décomposition de Lebesgue de Q par rapport à P et $Q(A \cap N)$ est la partie singulière dans cette décomposition.

Dans le cas où la mesure Q est absolument continue par rapport à P , la densité $f = dQ/dP$ est appelée rapport de vraisemblance de Q par rapport à P .

Revenons aux suites d'expériences statistiques $\xi_\theta^{(n)}$ décrites ci-dessus. Pour tester $P_\theta^{(n)}$

contre $P_{\theta+\nu(n)\tau(n)}^{(n)}$, d'après le Lemme de Neyman-Pearson, le rapport de vraisemblance est $dP_{\theta+\nu(n)\tau(n)}^{(n)}/dP_{\theta}^{(n)}$. Ainsi, le logarithme du rapport de vraisemblance est le suivant :

$$\Lambda_{\theta+\nu(n)\tau(n)} = \log \left(\frac{dP_{\theta+\nu(n)\tau(n)}^{(n)}}{dP_{\theta}^{(n)}} \right). \quad (1.5)$$

Le Cam (1986) a étudié la performance locale asymptotique des expériences statistiques décrites au point de départ. Il a trouvé que cette étude peut être caractérisée à partir du comportement de $\Lambda_{\theta+\nu(n)\tau(n)}$.

1.2.3.2 Normalité locale asymptotique

En utilisant les notations des paragraphes précédents, nous donnons la définition de la normalité locale asymptotique dont nous avons besoin.

Définition 1.2.8. (Définition 4.1 de Dreesbeke et Fine (1996))

La suite d'expériences $\mathcal{E}^{(n)}(\theta) = (\mathbb{R}^n, \mathcal{B}^n, P^{(n)}(\theta))$ est dite localement asymptotiquement normale LAN, si pour tout $\theta \in \Theta$, il existe une suite $\Delta^{(n)}(\theta)$ de vecteurs aléatoires \mathcal{B}^n -mesurables, et une matrice semi-définie positive $\Gamma(\theta)$ fonction continue de θ , telles que :

(i) : Pour toute suite de vecteurs $\tau^{(n)}$ telle que $\sup[(\tau^{(n)})'\tau^{(n)}] < +\infty$, sous $P^{(n)}(\theta)$, quand $n \rightarrow +\infty$, on a

$$\Lambda_{\theta+\nu(n)\tau^{(n)}}^{(n)} = \tau^{(n)'} \Delta^{(n)}(\theta) - \frac{1}{2} \tau^{(n)'} \Gamma(\theta) \tau^{(n)} + o_P(1).$$

(ii) : Sous $P^{(n)}(\theta)$ et quand $n \rightarrow +\infty$, on a

$$\Delta^{(n)}(\theta) \xrightarrow{Loi} \mathcal{N}(0, \Gamma(\theta)).$$

On appelle $\Delta^{(n)}(\theta)$ suite centrale.

Il y a plusieurs exemples de familles vérifiant la condition LAN présentée dans Dreesbeke et Fine (1996). On donne comme exemple le modèle de position à un échantillon, le modèle de régression simple, le modèle linéaire général, les familles régulières à la Cramer-Wald : par exemple le cas d'une structure d'échantillonnage $(\mathbb{R}^n, \mathcal{B}^n, \mathcal{P}^{\otimes n})$, où $\mathcal{P}^{\otimes n}$ est la famille des lois-produits caractérisées par les densités

$$f_{\theta}^{(n)}(X_1, \dots, X_n) = \prod_{i=1}^n f_{\theta}(X_i), \quad \theta \in \Theta \subset \mathbb{R},$$

lorsque $f_{\theta}(x) = \frac{1}{2} \exp\{-|x - \theta|\}$.

Les détails de ces exemples sont bien explicités dans Dreesbeke et Fine (1996).

1.2.3.3 Contiguïté

On désigne par $P^{(n)}$ et $Q^{(n)}$ deux mesures de probabilité sur l'espace mesurable $(\mathcal{X}^{(n)}, \mathcal{A}^{(n)})$. On a la définition de Dreesbeke et Fine (1996) suivante :

Définition 1.2.9. (Définition 4.2.1 de Dreesbeke et Fine (1996))

On dit que la suite $Q^{(n)}$ est contiguë à la suite $P^{(n)}$, si et seulement si pour toute suite $A^{(n)} \in \mathcal{A}^{(n)}$:

$$\lim_{n \rightarrow +\infty} P^{(n)}(A^{(n)}) = 0 \implies \lim_{n \rightarrow +\infty} Q^{(n)}(A^{(n)}) = 0.$$

Si deux suites $P^{(n)}$ et $Q^{(n)}$ sont mutuellement contiguës, on dit qu'elles sont contiguës. On peut remarquer que la contiguïté est une forme d'absolue continuité. En effet, Si $dQ^{(n)}/dP^{(n)}$ et $N^{(n)}$ fournissent une décomposition de Lebesgue de $Q^{(n)}$ par rapport à $P^{(n)}$. Puisque $P^{(n)}(N^{(n)}) = 0$ et la suite $Q^{(n)}$ est contiguë à la suite $P^{(n)}$, alors $Q^{(n)}(N^{(n)}) = 0$. Ainsi, la partie singulière de $Q^{(n)}$ par rapport à $P^{(n)}$ tend vers 0. L'un des intérêts est que, si on a une convergence en probabilité d'une suite d'événements par rapport à une mesure de probabilité $P^{(n)}$, on en déduit la convergence en probabilité de cette suite d'événements par rapport à toute mesure de probabilité $Q^{(n)}$ contiguë à $P^{(n)}$. On peut aussi remarquer que la contiguïté de $Q^{(n)}$ à $P^{(n)}$ conserve les ordres des convergences en probabilités : un $o_{P^{(n)}}(h(n))$ est également $o_{Q^{(n)}}(h(n))$.

En effet, si

$$\lim_{n \rightarrow \infty} P^{(n)} \left[\left| \frac{T^{(n)}}{h(n)} \right| > \varepsilon \right] = 0 \quad \forall \varepsilon > 0,$$

alors

$$\lim_{n \rightarrow \infty} Q^{(n)} \left[\left| \frac{T^{(n)}}{h(n)} \right| > \varepsilon \right] = 0 \quad \forall \varepsilon > 0.$$

Pour montrer la contiguïté d'une mesure à une autre, l'utilisation de la définition n'est pas toujours facile. Pour ceci on peut utiliser le corollaire ci-dessous qui est plus simple à appliquer.

Corollaire 1.2.1. (Corollaire 4.2 de Dreesbeke et Fine (1996))

La suite $\Lambda^{(n)} = \log \left(\frac{dQ^{(n)}}{dP^{(n)}} \right)$ est localement asymptotiquement normale sous $P^{(n)}$ c'est-à-dire

si il existe $\mu^{(n)}$ et $\sigma^{(n)}$ tels que, sous $P^{(n)}$, quand $n \rightarrow \infty$, $\frac{\Lambda^{(n)} - \mu^{(n)}}{\sigma^{(n)}} \xrightarrow{Loi} \mathcal{N}(0, 1)$, alors :

$P^{(n)}$ et $Q^{(n)}$ sont mutuellement contiguës si et seulement si :

- (i) : $\lim_{n \rightarrow +\infty} \sup \sigma^{(n)} < +\infty$.
- (ii) : $\lim_{n \rightarrow +\infty} \left(\mu^{(n)} + \frac{1}{2} \sigma^{(n)2} \right) = 0$.

Une conséquence du corollaire précédent en rapport avec la normalité locale asymptotique est donnée par le corollaire suivant :

Corollaire 1.2.2. (Corollaire 4.3 de Dreesbeke et Fine (1996))

Dans une famille LAN, les suites $P_{\theta+\nu(n)\tau(n)}^{(n)}$ et $P_{\theta}^{(n)}$ sont mutuellement contiguës pour tout $\theta \in \Theta$ et pour toute suite bornée $\tau^{(n)}$.

1.2.4 Puissance locale asymptotique des tests

Après la construction d'un test d'hypothèses, pour étudier un problème statistique quelconque, il est intéressant d'étudier sa puissance pour savoir la consistance et l'efficacité du test construit. Dans notre travail, nous nous intéressons à l'étude de la puissance locale asymptotique. Pour cela, nous présentons quelques notions utiles pour cette étude.

1.2.4.1 Le troisième lemme de Le Cam

Proposition 1.2.1. (Proposition 4.2 de Dreesbeke et Fine (1996))

Soit $S^{(n)}$ une suite de variables aléatoires et $\Lambda^{(n)} = \log \left(\frac{dQ^{(n)}}{dP^{(n)}} \right)$.

Si sous $P^{(n)}$,

$$\begin{pmatrix} S^{(n)} \\ \Lambda^{(n)} \end{pmatrix} \xrightarrow{Loi} \mathcal{N} \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right],$$

avec $\mu_2 = -\frac{\sigma_2^2}{2}$, alors $P^{(n)}$ et $Q^{(n)}$ sont mutuellement contiguës et sous $Q^{(n)}$,

$$S^{(n)} \xrightarrow{Loi} \mathcal{N}(\mu_1 + \sigma_{12}, \sigma_1^2).$$

1.2.4.2 Puissance locale

D'après Dreesbeke et Fine (1996), si on considère la suite des tests $T^{(n)}$ définie par :

$$T^{(n)} = \mathbb{1} \left\{ \frac{S^{(n)} - \mu^{(n)}}{\sigma_1^{(n)}} > Z_\alpha \right\},$$

où Z_α est le quantile d'ordre $1 - \alpha$ de la loi normale centrée réduite. Si sous $P^{(n)}$,

$$\begin{pmatrix} \frac{S^{(n)} - \mu^{(n)}}{\sigma_1^{(n)}} \\ \Lambda^{(n)} \end{pmatrix} \xrightarrow{Loi} \mathcal{N} \left[\begin{pmatrix} 0 \\ -\frac{1}{2}\sigma_2^2 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right].$$

Alors $T^{(n)}$ est de niveau asymptotique α (sous la suite d'hypothèses nulles $P^{(n)}$), et on a donc $\lim_{n \rightarrow +\infty} E_{P^{(n)}}(T^{(n)}) = \alpha$.

Le troisième lemme de Le Cam nous permet d'en déduire la puissance asymptotique de $T^{(n)}$ sous la suite des hypothèses alternatives $Q^{(n)}$, on obtient ainsi :

$$\lim_{n \rightarrow +\infty} E_{Q^{(n)}}(T^{(n)}) = 1 - \Psi(Z_\alpha - \sigma_{12}),$$

où Ψ est la fonction de répartition d'une loi normale centrée réduite.

1.2.5 Théorème limite centrale des martingales

L'un des outils fondamentaux que nous utilisons dans cette thèse pour étudier la loi asymptotique de la statistique du test considéré est le théorème limite centrale des martingales ci-dessous.

Théorème 1.2.4. (Théorème 3.2 de Hall et Heyde (1980))

Soit $\{S_{ni}, \mathcal{F}_{ni}, 1 \leq i \leq k_n, n \geq 1\}$ une martingale centrée et de carré intégrable et $X_{ni} = S_{ni} - S_{n,i-1}$ la différence des martingales correspondante. Soit η une variable aléatoire p.s. finie.

On suppose que :

$$\max_i |X_{ni}| \xrightarrow{P} 0, \quad (1.6)$$

$$\sum_{i=1}^n X_{ni}^2 \xrightarrow{P} \eta^2, \quad (1.7)$$

$$E(\max_i X_{ni}^2) \text{ est borné en } n, \quad (1.8)$$

et des σ -tribus sont croissantes, (i.e)

$$\mathcal{F}_{ni} \subseteq \mathcal{F}_{n+1,i} \text{ pour } 1 \leq i \leq k_n, n \geq 1. \quad (1.9)$$

On a alors,

$$S_{nk_n} = \sum_{i=1}^{k_n} X_{ni} \xrightarrow{Loi} Z,$$

où Z est une v.a. de fonction caractéristique $\mathbb{E}(\exp(\frac{1}{2}\eta^2 t^2))$.

Corollaire 1.2.3. (Corollaire 3.1 Hall et Heyde (1980))

On maintient dans l'énoncé du théorème précédent la condition (1.9).

Si les conditions (1.6) et (1.8) sont remplacées par la condition de Linderberg suivante :

$$\text{Pour tout } \varepsilon > 0, \lim_{n \rightarrow \infty} \sum_{i=1}^n E \left(X_{ni}^2 \mathbf{1}_{|X_{ni}| > \varepsilon} | \mathcal{F}_{n,i-1} \right) = 0 \quad \text{p.s.}$$

et si la condition (1.7) est remplacée par une condition du même type sur la variance conditionnelle :

$$\sum_{i=1}^{k_n} E(X_{ni}^2 | \mathcal{F}_{n,i-1}) \xrightarrow{P} \eta^2,$$

le théorème précédent reste vérifié.

Remarque 4. (Remarque de Hall et Heyde (1980))

Si on maintient dans l'énoncé du théorème 1.2.4 les conditions (1.6), (1.7) et (1.8) et on remplace la condition (1.9) par la condition de mesurabilité de la variable aléatoire η , le théorème reste aussi valable.

Dans notre travail nous étudions la convergence faible du processus log-vraisemblance dans l'espace $\mathbb{C}(\mathbb{R}^k, \mathbb{R})$ des fonctions définies et continues sur \mathbb{R}^k à valeurs dans \mathbb{R} , où $k \in \mathbb{N}$. Ces deux notions sont présentées dans Billingsley (1968) pour l'espace $\mathbb{C}([0, 1], \mathbb{R})$. Nous rappelons ici les principales définitions et les principaux résultats.

1.2.6 Convergence faible dans l'espace $\mathbb{C}([0, 1], \mathbb{R})$

Dans le chapitre 4 de cette thèse, nous étudions la convergence faible du processus de statistique centrale. Cette étude se fera dans l'espace de Fréchet des fonctions réelles définies sur \mathbb{R}^{k+1} . Dans ce paragraphe, nous définissons l'espace de Fréchet et nous donnons les principaux résultats utilisés.

Soit $\mathbf{C} = \mathbb{C}([0, 1], \mathbb{R})$ l'espace des fonctions continues sur $[0, 1]$ à valeurs réelles avec la métrique uniforme ρ définie pour $x = x(t)$ et $y = y(t)$ dans \mathbf{C} par

$$\rho(x, y) = \sup_t |x(t) - y(t)|.$$

L'espace \mathbf{C} est séparable, complet et par le théorème de Prohorov, la compacité relative d'une famille de mesures de probabilité dans $(\mathbf{C}, \mathcal{B}(\mathbf{C}))$, où $\mathcal{B}(\mathbf{C})$ est l'ensemble des boréliens de \mathbf{C} , est équivalente à la tension de cette famille. Pour plus des détails voir chapitre 2 de Billingsley (1968).

1.2.6.1 Critère de tension

Les critères de tension sont plus compliqués que dans \mathbb{R} , le plus général est le suivant : Soit le module de continuité d'une fonction x de \mathbf{C} suivant :

$$W_x(\delta) = w(x, \delta) = \sup_{|s-t| < \delta} |x(s) - x(t)|, \quad 0 < \delta \leq 1.$$

Soit $(P_n)_{n \in \mathbb{N}^*}$ une suite de mesures de probabilités sur $(\mathbf{C}, \mathcal{B}(\mathbf{C}))$.

Théorème 1.2.5. (Théorème 8.2 de Billingsley (1968))

La suite $(P_n)_{n \in \mathbb{N}^*}$ est tendue si et seulement si les deux conditions suivantes sont vérifiées :

(i) Pour tout $\eta > 0$, il existe $a \in \mathbb{R}_+$ tels que

$$P_n(\{|x; x(0) > a\}) \leq \eta, \quad n \geq 1.$$

(ii) Pour tout $\varepsilon > 0$ et $\eta > 0$, il existe $\delta \in]0, 1[$ et $n_0 \in \mathbb{N}$ telle que pour tout $n \geq n_0$,

$$P_n(\{x; W_x(\delta) \geq \varepsilon\}) \leq \eta.$$

Un critère plus accessible aux méthodes calculatoires est donné par le théorème ci-dessous.

Théorème 1.2.6. (Théorème 12.2 de Billingsley (1968))

Soit $(P_n)_{n \in \mathbb{N}^*}$ une suite de mesure de probabilités sur $(\mathbf{C}, \mathcal{B}(\mathbf{C}))$ et $X_n \sim P_n$.

S'il existe une constante $\gamma \geq 0$ et $\alpha > 1$ et une fonction F non décroissante, continue dans $[0, 1]$ telle que pour tout t_1, t_2 et n et pour tout réel λ positif

$$\mathbb{P}\{|X_n(t_2) - X_n(t_1)| \geq \lambda\} \leq \frac{1}{\lambda^\gamma} |F(t_2) - F(t_1)|^\alpha$$

et si de plus la suite $\{X_n(0)\}$, est tendue alors $(P_n)_{n \in \mathbb{N}^*}$ est tendue.

Remarque 5. $\{X_n(0)\}$ est tendue signifie pour tout $\varepsilon > 0$, il existe M telle que pour tout n ,

$$\mathbb{P}(|X_n(0)| \leq M) \geq 1 - \varepsilon.$$

1.2.6.2 Convergence faible

D'abord, les lois dans $\mathbb{C}([0, 1], \mathbb{R})$ sont caractérisées par ce qu'on appelle les lois finies-dimensionnelles : si on prend la loi μ d'un processus $X = (X_t)_{t \in [0, 1]}$, la loi de μ est caractérisée par la loi des vecteurs $(X_{t_1}, \dots, X_{t_k})$ pour $k \geq 1$, $0 \leq t_1 < \dots < t_k \leq 1$.

La convergence des lois finies-dimensionnelles est une condition nécessaire pour la convergence en loi, mais non suffisante. La question qui se pose maintenant est celle de savoir comment montrer la convergence en loi d'une mesure de probabilité dans $(\mathbf{C}, \mathcal{B}(\mathbf{C}))$. La réponse est donnée dans le Théorème 8.2 de Billingsley (1968) suivant :

Théorème 1.2.7. (Théorème 8.2 de Billingsley (1968))

Soit P_n et P deux mesures de probabilité dans $(\mathbf{C}, \mathcal{B}(\mathbf{C}))$. Si les lois finies-dimensionnelles de P_n convergent faiblement vers celles de P , et si $\{P_n\}$ est tendue, alors lorsque $n \rightarrow +\infty$, on a

$$P_n \xrightarrow{\text{loi}} P.$$

1.2.7 Espaces de Fréchet

1.2.7.1 Espaces semi-normés

Définition 1.2.10. (Définition 2.2 de Simon (2017))

Soit E un espace vectoriel. On appelle *semi-norme* sur E une application $\mathcal{P} : E \rightarrow [0, +\infty[$ vérifiant $\mathcal{P}(f + g) \leq \mathcal{P}(f) + \mathcal{P}(g)$ et $\mathcal{P}(\lambda f) = |\lambda| \mathcal{P}(f)$ pour $f, g \in E$ et λ scalaire.

Une norme est une semi-norme avec la condition supplémentaire que $\mathcal{P}(f) > 0$ si f est nulle.

Définition 1.2.11. (Définition 2.3 de Simon (2017))

- i. On appelle **espace normé** tout espace vectoriel muni d'une norme. La norme d'un espace normé E est notée $\|\cdot\|_E$.
- ii. On appelle **espace semi-normé** E tout espace vectoriel muni d'une famille non vide de semi-normes $\{\mathcal{P}_i : i \in \mathcal{N}_E\}$, où \mathcal{N}_E est un ensemble indexant les semi-normes est, a priori quelconque.
- iii. Un espace semi-normé E est dit **séparé** si, pour tout $f \in E$:

$$\forall \nu \in \mathcal{N}_E, \mathcal{P}_\nu(f) = 0 \implies f = 0.$$

Définition 1.2.12. (Définition 4.8 de Simon (2017))

Un espace semi-normé séparé est dit **séquentiellement complet** si toutes ses suites de Cauchy convergent.

1.2.7.2 Espaces de Fréchet

Définition 1.2.13. (Définition 4.10 de Simon (2017))

On appelle **espace de Fréchet** tout espace semi-normé métrisable séquentiellement complet.

Dans le chapitre 4 de cette thèse, nous considérons le cas où l'hypothèse alternative étudiée est une suite d'hypothèses. Nous voulons montrer que notre test détecte bien les points de ruptures lorsque le paramètre de saut du modèle n'est pas fixé. Nous allons faire cette étude dans l'espace $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$. Plus précisément, nous définissons la métrique ρ dans $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$ par

$$\rho(x, y) = \sum_{j=1}^{\infty} 2^{-j} \frac{\rho_j(x, y)}{1 + \rho_j(x, y)},$$

où

$$\rho_j(x, y) = \sup_{\|t\| < j} |x(t) - y(t)|$$

est la norme sup. Avec cette métrique, $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$ est un espace de Fréchet séparable et la convergence dans cette métrique correspond aux convergences uniformes dans tout ensemble compact. Nous travaillons dans cet espace de Fréchet parce que le paramètre de saut est quelconque. Pour cette raison, nous considérons cette mesure avec laquelle notre espace est de Fréchet. De plus, nous adoptons les résultats énoncés dans Billingsley (1968) ci-dessus pour les utiliser sur l'espace $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$.

Chapitre 2

État de l'art

2.1 Modèle CHARN

Dans ce paragraphe, nous donnons une définition générale de la classe des modèles CHARN et nous présentons quelques travaux qui se sont intéressés à ce type de modèle.

Définition 2.1.1. *On dit qu'un processus $(X_t)_{t \in \mathbb{Z}}$ est un processus CHARN(p) s'il est généré par le modèle suivant :*

$$X_t = M_{\theta_0}(X_{t-1}, X_{t-2}, \dots, X_{t-p})\varepsilon_t + f_{\theta_0}(X_{t-1}, X_{t-2}, \dots, X_{t-p}), \forall t \in \mathbb{Z}, \quad (2.1)$$

où $\theta_0 \in \Theta \subset \mathbb{R}^d$ c'est le paramètre d'intérêt, $M_{\theta_0}(X_{t-1}, X_{t-2}, \dots)$ c'est une matrice $(m \times p)$ de rang m , $f_{\theta_0}(X_{t-1}, X_{t-2}, \dots)$ est un vecteur de \mathbb{R}^m , les $\varepsilon_t = (\varepsilon_t^{(k)})_{1 \leq k \leq p}$ sont indépendantes identiquement distribuées telles que $\mathbb{E}(\varepsilon_0^{(k)} \varepsilon_0^{(k')}) = 0$ pour $k \neq k'$ et $\mathbb{E}(\varepsilon_0^{(k)2}) = \text{Var}(\varepsilon_0^{(k)}) = 1$. Le cas où $p = +\infty$ est considéré par exemple par Bardet et al. (2009).

Cette classe de modèles est considérée dans Härdle et Tsybakov (1997) sous la forme

$$Y_i = f(Y_{i-1}) + S(Y_{i-1})\varepsilon_i, \quad i = 1, 2, \dots$$

où $(\varepsilon_i)_i$ est une suite de variables aléatoires indépendantes et identiquement distribuées (iid) telle que $\mathbb{E}(\varepsilon_i) = 0$ et $\mathbb{E}(\varepsilon_i^2) = 1$, f et S sont des fonctions inconnues telle que $S(y) > 0, \forall y \in \mathbb{R}^I$ et Y_0 est une v.a. indépendante de $(\varepsilon_i)_i$. C'est le modèle (2.1) avec $M_{\theta_0} \equiv f$ et $f_{\theta_0} \equiv S$.

Dans ce travail, les auteurs donnent une estimation de la fonction de volatilité $V(x) = S^2(x)$ pour l'échantillon donné Y_1, \dots, Y_n , basé sur la méthode des polynômes locaux (LP) et ils prouvent la normalité asymptotique de cet estimateur.

Härdle *et al.* (1998) considèrent un modèle CHARN multivarié, décrit comme suit :

$$Y_i = f(X_i) + \Sigma^{1/2}(X_i)\varepsilon_i, \quad i = 1, 2, \dots$$

$$X_i = (Y_{i-1}^T, \dots, Y_{i-m}^T)^T \in \mathbb{R}^{md},$$

où $Y_i = (Y_{i1}, \dots, Y_{id})^T \in \mathbb{R}^d$, $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{id})^T \in \mathbb{R}^d$, $i = m, \dots, n$ et les ε sont indépendantes et identiquement distribuées avec $\mathbb{E}(\varepsilon_{1j}) = 0$ et $\mathbb{E}(\varepsilon_{1j}^2) = 1$, $\forall 1 \leq j \leq d$. Le vecteur moyenne $f : \mathbb{R}^{md} \rightarrow \mathbb{R}^d$ et la volatilité $\Sigma : \mathbb{R}^{md} \rightarrow \mathbb{R}^{2d}$ sont inconnus avec variable initiale $Y_m = (Y_{m-1}^T, \dots, Y_0^T)^T$ indépendante de $(\varepsilon_i)_i$.

Les auteurs s'intéressent à l'estimation des fonctions f et Σ , et ils généralisent les résultats de Härdle et Tsybakov (1997).

L'un des estimateurs fondamentaux pour les modèles de séries chronologiques financières est celui des moindres carrés conditionnels (CL). Cependant, Chandra et Taniguchi (2001) construisent un estimateur qu'ils notent G, et ils montrent qu'il est plus efficace que l'estimateur CL pour certains modèles de séries chronologiques. Dans Amano (2012), l'efficacité des estimateurs CL et G pour le modèle CHARN est étudié. Amano (2012) en déduit la condition selon laquelle l'estimateur G est asymptotiquement optimal. Il considère le modèle CHARN univarié. Ici, le paramètre de changement est $\theta = (\theta_1, \dots, \theta_k)^\top$ qui est supposé dans l'ouvert Θ de \mathbb{R}^k et le vrai paramètre du modèle est θ_0 . L'auteur définit deux estimateurs notés $\hat{\theta}_{EF}$ et $\hat{\theta}_G$ de θ_0 qui sont respectivement l'estimateur des moindres carrés et le G-estimateur de θ_0 . Amano (2012) montre la normalité asymptotique de ces deux estimateurs et conclut que l'estimateur $\hat{\theta}_G$ est moins efficace que $\hat{\theta}_{CL}$.

Ces articles mentionnés ci-dessus ne sont pas intéressés par la détection des ruptures mais on peut les utiliser dans notre travail pour estimer les paramètres du modèle considéré.

2.2 Estimation des points de ruptures

Beaucoup de travaux sont consacrés à l'estimation des points de ruptures. Nous en résumons dans cette sous-section quelques uns. Nous commençons par Wong *et al.* (2001) qui considèrent le modèle CHARN

$$X_t = T(X_{t-1}) + \sigma(X_{t-1})\varepsilon_t,$$

où $\{\varepsilon_t\}$ est un bruit blanc iid tel que $\mathbb{E}(\varepsilon_t) = 0$, $Var(\varepsilon_t) = 1$, $\mathbb{E}(|\varepsilon_t|^\nu) < \infty$, pour $\nu > 2$, et $\{\varepsilon_{t+1}\}$ est indépendant à $\mathcal{F}_t^0 = \mathcal{F}(X_t, X_{t-1}, \dots)$. $T(x)$ et $\sigma(x)$ sont bornés sur l'intervalle connu $[a; b]$ et ils ont des points de rupture aux instants t_l , $1 \leq l \leq p$, et s_m , $1 \leq m \leq q$, respectivement où $a < t_1 < t_2 < \dots < t_p < b$, $a < s_1 < s_2 < \dots < s_q < b$, avec p , q , t_l et s_m des constantes inconnues.

Sous des conditions d'ergodicité et de stationnarité, les auteurs estiment ces constantes et montrent leur consistance.

Bardet *et al.* (2010) considèrent une autre extension du modèle étudié dans Wong *et al.*

(2001), le modèle (2.1) ci-dessus. Ils étudient les propriétés asymptotiques de l'estimateur du maximum de vraisemblance de θ_0 . Ce travail est poursuivi dans Bardet *et al.* (2012) avec le modèle

$$X \in M_{T_j^*}(M_{\theta_j^*}, f_{\theta_j^*}), \quad \forall j = 1, \dots, K^*, \quad (2.2)$$

où

- $K^* \in \mathbb{N}^*$, $T_j^* = \{t_{j-1}^* + 1, t_{j-1}^* + 2, \dots, t_j^*\}$, $0 < t_1^* < \dots < t_{K^*-1}^* < n$, $t_j^* \in \mathbb{N}$ et par convention $t_0^* = -\infty$, $t_{K^*}^* = +\infty$ et $p = +\infty$.
- $\theta_j^* = (\theta_{j,1}^*, \dots, \theta_{j,d}^*) \in \Theta \subset \mathbb{R}^d$ pour $j = 1, \dots, K^*$.

Il est à noter que $X \in M_T(M, T)$ signifie que X satisfait la relation (2.1) avec $p = +\infty$. Le but est l'estimation des paramètres inconnus K^* , $(t_j^*)_{1 \leq j \leq K^*-1}$, et $(\theta_j^*)_{1 \leq j \leq K^*}$, qui sont respectivement les nombres de ruptures, les instants de ruptures et les paramètres du modèle. Les auteurs montrent que lorsque f_θ et M_θ sont lipschitziennes, les estimateurs considérés \widehat{K}_n , $(\widehat{t}_j/n)_{1 \leq j \leq \widehat{K}_n-1}$ et $(\widehat{\theta}_j)_{1 \leq j \leq \widehat{K}_n}$ de K^* , $(t_j^*)_{1 \leq j \leq K^*-1}$, et $(\theta_j^*)_{1 \leq j \leq K^*}$ respectivement, sont consistants lorsque ε_t est de moment d'ordre r , $r \geq 1$, et X admet des moments d'ordre > 2 . Ils montrent aussi que si de plus, les dérivées de f_θ et M_θ sont lipschitziennes et si $r \geq 4$, le taux de convergence de $(\widehat{t}_j/n)_{1 \leq j \leq \widehat{K}_n-1}$ est $O_P(n^{-1})$. Ils établissent le théorème central limite pour $(\widehat{\theta}_j)_{1 \leq j \leq \widehat{K}_n}$ (avec un taux de convergence \sqrt{n}). Le modèle utilisé dans Lavielle et Lebarbier (2001) est le suivant :

Soit $y = (Y_t)_{t \geq 1}$ un processus réel tel que pour tout $t \geq 1$,

$$Y_t = S(t) + \varepsilon_t,$$

où $(\varepsilon_t)_{t \geq 1}$ est une suite de variables aléatoires centrées et S une fonction constante par morceaux. La seconde hypothèse implique qu'il existe une suite de constantes $(\tau_k)_{k \geq 0}$ et une suite $(m_k)_{k \geq 1}$ telles que pour tout $k \geq 1$, $S(t) = m_k$, pour tout $\tau_{k-1} + 1 \leq t \leq \tau_k$ (c'est-à-dire que S est constante sur chaque segment $[\tau_{k-1}, \tau_k]$). Ce modèle est un cas particulier de celui de Bardet *et al.* (2012).

Lavielle et Lebarbier (2001) introduisent $r = (r_t)_{t \geq 1}$ comme processus des points de rupture tel que

$$r_t = \begin{cases} 1 & \text{s'il existe } k \text{ tel que } t = \tau_k \\ 0 & \text{sinon} \end{cases}$$

Autrement dit, $(r_t)_t$ est un indicateur de l'existence des points de rupture. Elle vaut 1 s'il existe et 0 sinon. Pour estimer les instants des points de rupture, il suffit de l'estimer, ce qui permet aussi d'estimer S . Les auteurs proposent une méthodologie intéressante dans un contexte Bayésien.

Le modèle probabiliste utilise une séquence non observée $(r_t)_{t \geq 1}$ dont un algorithme MCMC peut être utilisé pour estimer la distribution a posteriori. Des simulations numériques ont montré que cette procédure est beaucoup plus rapide que l'algorithme de

saut réversible.

Lebarbier (2005) considère le même modèle que Lavielle et Lebarbier (2001) lorsque

$$Y_t = S(X_t) + \varepsilon_t, \quad t = 1, \dots, n,$$

où $X_t = \frac{t}{n}$, $(\varepsilon_t)_t$ est une suite iid de variables aléatoires gaussiennes centrées et de variance $\sigma^2 < +\infty$, S est une fonction constante par morceaux. Donc il existe certains instants $\tau_0 = 0 < \tau_1 < \tau_2 < \dots < \tau_k = 1$ et des constantes S_k , $k = 1 \dots, K$ tels que

$$S = \sum_{k=1}^K S_k \mathbf{1}_{I_k}, \quad I_k =]\tau_{k-1}, \tau_k].$$

Donc on a $K - 1$ points de rupture de (Y_t) à des instants inconnus $(t_k, 1 \leq k \leq K - 1)$, avec $t_k = [n\tau_k]$. Le nombre $K - 1$ est supposé inconnu. Le problème est de détecter et localiser les instants des points de rupture de Y et estimer la moyenne des sauts. L'auteure propose un critère des moindres carrés pénalisés calibrés pour le problème des points de rupture par une approche non-asymptotique. Ceci conduit à une méthode automatique qui fonctionne bien et mieux que les critères Mallows C_p et BIC notamment pour les petits échantillons. Cependant, il est difficile d'obtenir un étalonnage qui fonctionne bien dans toutes les situations. L'un des objectifs de cette étude est d'évaluer la méthode heuristique proposée par Birgé et Massart (2001) pour l'estimation de la constante de pénalité et de la calibrer dans le cadre particulier de la régression gaussienne.

Des expériences numériques ont montré que la méthode étalonnée proposée fonctionne bien en terme de risque quadratique. Ceci permet d'espérer qu'il sera utile dans des situations plus complexes, en ce sens que la constante de pénalité n'est pas explicitement connue.

Horváth et Kokoszka (1997) considèrent aussi le cas d'une série d'observations présentant un changement dans la moyenne. Ils supposent que la série d'observations $(X_t)_{1 \leq t \leq n}$ présente une rupture à l'instant k^* et qu'elle est générée par le modèle suivant :

$$X_t = \begin{cases} \mu + \varepsilon_t & \text{pour } 1 \leq t \leq k^* \\ \mu + \delta + \varepsilon_t & \text{pour } k^* < t \leq n, \end{cases}$$

où μ, δ, k^* sont des paramètres inconnus, $(\varepsilon_t)_t$ est un bruit blanc centré et n est le nombre d'observations. Ici, la série d'observations est stationnaire par morceaux.

Horváth et Kokoszka (1997) étudient une famille d'estimateurs de point de rupture k^* lorsque l'erreur ε_t exhibe une dépendance de longue mémoire. C'est-à-dire lorsqu'il vérifie les conditions $C.1$ et $C.2$ de l'article. Ils montrent que les estimateurs de k^* proposés, après une normalisation, convergent en loi vers une fonction de mouvements browniens fractionnels.

Le problème de l'estimation d'un point de rupture inconnu dans le vecteur moyen ou la matrice de covariance d'une série des variables aléatoires gaussiennes multivariées indépendantes est considéré dans Fotopoulos *et al.* (2009). Ceux-ci supposent que pour tout $n \geq 2$, $Y_{n,1}, \dots, Y_{n,n}$ sont des observations gaussiennes indépendantes multivariées de densité de probabilité $f(y, \mu, \Sigma)$. Ils supposent aussi que le paramètre (μ, Σ) change de sa valeur initiale (μ_0, Σ_0) à (μ_1, Σ_1) . Pour un point d'index inconnu $\tau_n \in \{1, 2, \dots, n-1\}$, $(\mu_0, \Sigma_0) \neq (\mu_1, \Sigma_1)$. Ainsi,

$$f_i(Y_{n,i}, \mu, \Sigma) = \begin{cases} f(Y_{n,i}, \mu_0, \Sigma_0) & \text{pour } i = 1, 2, \dots, \tau_n \\ f(Y_{n,i}, \mu_1, \Sigma_1) & \text{pour } i = \tau_n + 1, \dots, n \end{cases}$$

Soit $\hat{\tau}_n$ l'estimateur du maximum de vraisemblance de τ_n (c'est le temps de changement). D'après Hinkley (1970),

$$\hat{\tau}_n = \arg \max_{1 \leq j \leq n-1} \sum_{i=1}^j a(Y_{n,i}),$$

où $a(Y_{n,i}) = \log \left(\frac{f(Y_{n,i}, \mu_0, \Sigma_0)}{f(Y_{n,i}, \mu_1, \Sigma_1)} \right)$, $i = 1, 2, \dots, n$ et $a : \mathbb{R}^d \rightarrow \mathbb{R}$.

Soit τ_n tel que $\tau_n = [n\rho]$, où $\rho \in]0, 1[$ constante inconnue qui représente le point de rupture inconnu et $\hat{\rho}_n = \hat{\tau}_n/n$ est un estimateur de ρ . Fotopoulos *et al.* (2009) expriment l'estimateur du maximum de vraisemblance centré comme suit :

$$\xi_n = \hat{\tau}_n - \tau_n = [n\hat{\rho}_n] - [n\rho] = \arg \max_{-[n\rho]+1 \leq j \leq n-[n\rho]-1} \sum_{i=1}^{[n\rho]+j} a(Y_{n,i})$$

et ils étudient ses propriétés.

Huh (2010) considère un modèle de régression linéaire. Il suppose que ce modèle présente des points de rupture. Le temps et le nombre de sauts sont estimés. L'auteur étudie la loi asymptotique des estimateurs proposés et il illustre les résultats obtenus par des simulations numériques. Plus précisément, il suppose qu'on observe au hasard (X_i, Y_i) copies de (X, Y) , $i = 1, \dots, n$, où la loi conditionnelle de Y sachant $X = x$ est donnée par :

$$f_{Y|X}(y|x) = \exp(y\theta(x) - b(\theta(x)) + c(y)),$$

où b, c sont des fonctions connues. Il s'intéresse à l'estimation de la fonction de régression $m(x) = \mathbb{E}(Y|X = x) = b(\theta(x))$.

Döring (2010, 2011) considère un problème de rupture multiple : une série finie des variables aléatoires indépendantes est constituée de segments donnés par un nombre connu de points de rupture de sorte que la distribution sous-jacente diffère d'un segment à l'autre. L'auteur estime ces points de ruptures dans un cadre complètement non-paramétrique.

L'estimateur proposé est défini comme le point de maximisation de U-statistiques multivariés pondérés.

Dans Ciuperca (2011), un critère général est proposé pour déterminer le nombre K des points de ruptures dans un modèle non-linéaire multidimensionnel paramétrique. L'auteure montre que l'estimateur proposé pour K est consistant. En particulier, elle a considéré le cas où

$$Y_i = f_{(\theta_{1,K}, \theta_{2,K})}(X_i) + \varepsilon_i, \quad i = 1, \dots, n,$$

où les Y_i , $i = 1, \dots, n$ sont les observations, $(X_i, \varepsilon_i)_{1 \leq i \leq n}$ est une suite de variables aléatoires iid, K est le nombre des points de rupture et

$$\begin{aligned} f_{(\theta_{1,k}, \theta_{2,k})} &: \mathbb{R} \longrightarrow \mathbb{R} \\ x &\mapsto h_{\beta_1}(x)\mathbb{1}_{x \leq \tau_1} + h_{\beta_2}(x)\mathbb{1}_{\tau_1 < x \leq \tau_2} + \dots + h_{\beta_{K+1}}(x)\mathbb{1}_{x > \tau_K}, \end{aligned}$$

où h_{β_k} est une fonction connue du paramètre β_k , avec $k = 1, \dots, K$.

Les paramètres de régression sont $\theta_{1,K} = (\beta_1, \dots, \beta_{K+1})$ et les points de changement sont $\theta_{2,K} = (\tau_1, \dots, \tau_K)$, avec $\tau_1 < \dots < \tau_K$. L'auteure propose un estimateur consistant de K . Les résultats de simulation pour les modèles non-linéaires ayant de nombreuses applications confirment la performance de l'estimateur construit. Il semble que cet estimateur est plus performant que celui de Nosek (2010).

Badagián *et al.* (2015) traitent le problème de la localisation et de l'estimation des points de rupture dans une série chronologique. Ils s'intéressent à l'étude des ruptures dans la moyenne et les coefficients autorégressifs dans les processus autorégressifs par morceaux, ainsi que les ruptures dans la variance des innovations. Ils proposent une approche basée sur le critère d'information bayésienne (BIC) et la segmentation binaire. Cette procédure est comparée à plusieurs autres méthodes. Ils calculent les propriétés empiriques de taille et de puissance des procédures disponibles dans plusieurs cas de Monte Carlo, et ils comparent leurs performances sur un jeu de données. Ils constatent qu'elle a d'excellentes performances.

Dans Kanai *et al.* (2010), les auteurs appliquent l'approche de la fonction d'estimation aux modèles CHARN qui comprennent des nombreux modèles de séries chronologiques non-linéaires bien connus comme cas spéciaux et que nous considérons dans cette thèse. Il s'agit du modèle de Wang *et al.* (2012) et Bardet *et al.* (2009) pour $M_\theta = F_\theta$, $f_\theta = H_\theta$ et $u_t = \varepsilon_t$, où $X_t = (x_{1,t}, \dots, x_{m,t})_{t \in \mathbb{Z}_+}^T$ est un processus stochastique multidimensionnel, $F_\theta : \mathbb{R}^{mp} \longrightarrow \mathbb{R}^m$ est une fonction mesurable, $H_\theta : \mathbb{R}^{mq} \longrightarrow \mathbb{R}^{2m}$ est une fonction mesurable et $u_t = (u_{1,t}, \dots, u_{m,t})^T$ est une suite de variables aléatoires iid telle que $\mathbb{E}(u_t) = 0$, $\mathbb{E}(|u_t|) < \infty$ et $(u_t)_t$ indépendante de $\{X_s, s < t\}$. Ici, le paramètre du modèle est $\theta = (\theta_1, \dots, \theta_r)^T \subset \Theta \subset \mathbb{R}^r$.

La vraie valeur et son estimateur sont respectivement notés par θ_0 et $\widehat{\theta}_{EF} = (\widehat{\theta}_1, \dots, \widehat{\theta}_r)^T$. Les propriétés asymptotiques de cet estimateur sont étudiées et il est prouvé que sa distribution asymptotique ne dépend pas de F_θ et H_θ . Ensuite, une région de confiance pour θ est construite, ce qui permet d'effectuer des tests d'hypothèses.

2.3 Tests d'existence de points de ruptures

2.3.1 Tests généraux

Nous présentons maintenant quelques travaux sur les tests d'existence de points de ruptures. Nous commençons par Kengne (2012) qui considère la classe des modèles $M_T(M_\theta, f_\theta)$ décrits plus haut. Il étudie le test de l'hypothèse H_0 contre H_1 où

H_0 : Il existe $\theta_0 \in \Theta$ tel que $(X_t)_{t \in \mathbb{Z}}$ appartient à la classe $M_{\{1, \dots, n\}}(M_{\theta_0}, f_{\theta_0})$

H_1 : Il existe $K \geq 2$, $(\theta_1^*, \dots, \theta_K^*) \in \Theta$ tel que $\theta_j^* \neq \theta_{j+1}^*$, tel que $(X_t)_{t \in \mathbb{Z}} \in \cap_{j=1}^k M_{T_j^*}(M_{\theta_j^*}, f_{j^*})$, où $T_j^* = \{t_{j-1}^* + 1, t_{j-1}^* + 2, \dots, t_j^*\}$, tel que $0 = t_0^* < t_1^* < \dots < t_{K-1}^* < t_K^* = n$.

H_1 implique que pour tout $j = 1, \dots, K$, $(X_t)_{t \in \mathbb{Z}}$ appartient à la classe $M_{T_j^*}(M_{\theta_j^*}, f_{j^*})$. Autrement dit, la série d'observations est stationnaire sur chaque intervalle $[t_{j-1}^*, t_j^*]$ et les paramètres du modèle qui changent sont θ_j , M_{θ_j} et f_{θ_j} . Donc la série est stationnaire par morceaux et le nombre de changements égal à K . Ainsi, il est facile de voir que sous H_1 , la propriété de stationnarité est perdue après le premier changement, ce qui est le cas dans notre travail, mais pas dans des nombreux travaux existants (par exemple Kouamo *et al.* (2010)), où la stationnarité du K -ème ordre après le changement est une hypothèse essentielle.

Kengne (2012) considère la statistique de test suivante :

$$\widehat{Q}_n = \max(\widehat{Q}_n^{(1)}, \widehat{Q}_n^{(2)}),$$

où les deux statistiques $\widehat{Q}_n^{(1)}$ et $\widehat{Q}_n^{(2)}$ sont définies comme suit :

$$\widehat{Q}_n^{(1)} := \max_{k \in \Pi_n} \widehat{Q}_{n,k}^{(1)}, \text{ avec } \widehat{Q}_{n,k}^{(1)} := \frac{k^2}{n} \left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n) \right)^\top \widehat{\Sigma}_{n,k} \left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n) \right),$$

$$\widehat{Q}_n^{(2)} := \max_{k \in \Pi_n} \widehat{Q}_{n,k}^{(2)}, \text{ avec } \widehat{Q}_{n,k}^{(2)} := \frac{(n-k)^2}{n} \left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n) \right)^\top \widehat{\Sigma}_{n,k} \left(\widehat{\theta}_n(T_k) - \widehat{\theta}_n(T_n) \right),$$

où pour tout $T_i \subset \{1, \dots, n\}$, $\widehat{\theta}_n(T_i) := \operatorname{argmax}_{\theta \in \Theta} \widehat{L}_n(T_i, \theta)$ est l'estimateur du maximum de log de vraisemblance de θ , avec $\widehat{L}_n(T_i, \theta)$ l'estimateur de la fonction log de vraisemblance, $\widehat{\Sigma}_{n,k}$ est une matrice semi-définie positive définie à partir des fonctions $\widehat{G}_n(T_i)$ et $\widehat{F}_n(T_i)$ données dans l'article, et $\Pi_n = [v_n, n - v_n] \cap \mathbb{N}$, où $(v_n)_n$ est une suite telle que $v_n \rightarrow \infty$ et $v_n/n \rightarrow 0$ lorsque $n \rightarrow \infty$.

La comparaison avec la procédure basée sur $\widehat{Q}_n^{(0)}$ proposée par Lee et Song (2008) est en faveur de ce test dans de nombreux cas. Selon les résultats de simulation, il est recommandé d'utiliser cette procédure lorsque $n \geq 1000$.

Bardet et Kengne (2014) considèrent aussi la classe des modèles $M_T(M_\theta, f_\theta)$. Ils supposent observées les données X_1, \dots, X_n , avec $(X_1, \dots, X_n) \in M_{\{1, \dots, n\}}(M_{\theta_0^*}, f_{\theta_0^*})$ et θ_0^* est inconnue. Ensuite, ils considèrent de nouvelles données $X_{n+1}, X_{n+2}, \dots, X_k, \dots$. Pour chaque nouvelle observation, ils veulent savoir si un changement se produit dans le paramètre θ_0^* . Plus précisément, ils considèrent le problème de test suivant :

H_0 : θ_0^* est constante sur les observations $X_1, \dots, X_n, X_{n+1}, \dots$, c'est-à-dire $(X_n)_{n \in \mathbb{N}} \in M_{\mathbb{N}}(M_{\theta_0^*}, f_{\theta_0^*})$

H_1 : Il existe $K^* > n$, $(\theta_0^*, \theta_1^*) \in \Theta^2$ tel que $\theta_0^* \neq \theta_1^*$

tel que $(X_1, \dots, X_{K^*}) \in M_{\{1, \dots, K^*\}}(M_{\theta_0^*}, f_{\theta_0^*})$ et $(X_{K^*+n})_{n \in \mathbb{N}} \in M_{\{K^*+1, \dots\}}(M_{\theta_1^*}, f_{\theta_1^*})$.

Ils proposent un test du rapport de vraisemblance. Ils trouvent que sous certaines hypothèses, sous H_1 , si $\theta_1^* \neq \theta_0^*$ et s'il existe $T^* \in]1, T[$ tel que $k^* = k^*(n) = [T^*n]$, alors pour $k_n = k^*(n) + n^\delta$, avec $\delta \in]\frac{1}{2}, 1[$,

$$\max_{l \in \Pi_{n, k_n}} \frac{\widehat{C}_{k_n, l}}{b(k_n - l)} \xrightarrow{p.s.} +\infty,$$

où $\widehat{C}_{k_n, l}$ est la statistique du test étudié.

Yang et Song (2014) généralisent le travail de Wong *et al.* (2001) en considérant le modèle de régression non-paramétrique suivant :

$$Y_t = m(X_t) + \sigma(X_t)\varepsilon_t, \quad t = 1, \dots, n,$$

où $(X_t, Y_t)_{t=1}^n$ est un processus stochastique strictement stationnaire bivarié, $\{\varepsilon_t\}_{t=1}^n$ est un bruit blanc conditionnel tel que $\mathbb{E}(\varepsilon_t | X_t = x) = 0$, $Var(\varepsilon_t | X_t = x) = 1$, $t = 1, \dots, n$, les fonctions m et σ sont définies sur l'intervalle $[a, b]$ et m est continue sur cet intervalle avec un nombre fini de points qu'on appelle sauts. Ils supposent inconnus, le nombre et les instants de ces points de rupture. Sous des conditions raisonnables, des statistiques de test sur l'existence de sauts sont fournies et leurs distributions limites sont étudiées sous l'hypothèse de non-changement. À l'aide des simulations numériques, ils montrent que la méthode fonctionne assez bien pour un ensemble de données de taille modérée.

Gombay et Serban (2009) s'intéressent au modèle AR(p) suivant :

$$Y_i - \mu = \phi_1(Y_{i-1} - \mu) + \dots + \phi_p(Y_{i-1} - \mu) + \varepsilon_i, \quad i \geq 1,$$

où $(Y_i)_i$ est une suite d'observations, μ , ϕ_i , $i = 1, \dots, p$ sont des constantes et $(\varepsilon_i)_i$ est une suite de variables aléatoires iid telles que $\mathbb{E}(\varepsilon_i) = 0$ et $Var(\varepsilon_i) = \sigma^2$.

Les auteurs étudient le problème de la détection de rupture. Ils supposent que le vrai paramètre du modèle est θ_0 et θ le paramètre de changement qui est l'une des composantes du vecteur $\xi = (\mu, \sigma^2, \phi_1, \dots, \phi_p)^\top$ inconnu.

Le test utilisé est basé sur une approximation du vecteur du score sous l'hypothèse nulle de non changement et sous l'alternative de changement. Ils proposent des estimateurs pour chaque paramètre. Ils montrent leur consistance et ils établissent des résultats asymptotiques pour chaque statistique du test considéré.

Gombay et Horváth (1994) considèrent X_1, \dots, X_n variables aléatoires indépendantes, chacune de fonction de répartition $F(x, \theta_1), F(x, \theta_2), \dots, F(x, \theta_n)$ respectivement, où $\theta_i \in \Theta \subset \mathbb{R}^d$, $1 \leq i \leq n$. Ils étudient le problème de test suivant :

$$H_0 : \theta_1 = \dots = \theta_n$$

contre

$$H_A : \text{Il existe } k^*, 1 \leq k^* < n \text{ telle que } \theta_1 = \dots = \theta_{k^*} \neq \theta_{k^*+1} = \dots = \theta_n.$$

Il est clair que sous H_A , après la k^* -ième observation, il y a un changement dans le paramètre de la suite d'observations. Ici, l'instant de rupture est k^* . La détection de ce point de rupture est faite par la considération d'une statistique de type maximum de vraisemblance. Ils trouvent la loi asymptotique du test sous l'hypothèse alternative de non-changement.

Prášková et Chochola (2014) étudient une procédure permettant de détecter un changement de paramètres de régression dans un modèle linéaire. La procédure de test proposée est basée sur les M-estimateurs et les M-résidus pondérés correspondants. Les distributions asymptotiques sont étudiées.

L'objectif de Jarušková (2010) est de détecter un changement brusque dans la moyenne d'une suite de vecteurs aléatoires indépendants normalement distribués. La formulation de ce problème est la suivante :

Soit $X(1) = (X_1(1), \dots, X_p(1))^\top, \dots, X(n) = (X_1(n), \dots, X_p(n))^\top$, les vecteurs observés.

Ils considèrent le problème de test d'hypothèses suivant :

$H_0 : X_i(j) = \mu_i + e_i(j), j = 1, \dots, n, i = 1, \dots, p$ contre $A : \text{Il existe } k^* \in \{1, \dots, n-1\}$ tel que

$$X_i(j) = \mu_i + e_i(j), \quad j = 1, \dots, k^*, \quad i = 1, \dots, p$$

$$X_i(j) = \mu_i + {}^n\sigma_i(j) + e_i(j), \quad j = k^* + 1, \dots, n, \quad i = 1, \dots, p$$

Les constantes $\mu_i, i = 1, \dots, p$ et ${}^n\sigma_i(j) \in \mathbb{R}, j = k^* + 1, \dots, n, i = 1, \dots, p$, sont inconnues. Pour résoudre ce problème de test, Jarušková (2010) considère la statistique

$$T_n^2 = \max_{1 \leq k < n} \frac{n}{k(n-k)} ({}^nS^X(k))^\top ({}^nS^X(k)),$$

où pour tout $k = 1, \dots, n$, ${}^n S^X(k) = ({}^n S_1^X(k), \dots, {}^n S_p^X(k))^\top$,

$${}^n S_i^X(k) = \sum_{j=1}^k \frac{X_i(j) - \bar{X}_i}{\sqrt{{}^n \sigma_{ii}^X}},$$

$$\bar{X}_i = \frac{1}{n} \sum_{j=1}^n X_i(j), \sqrt{{}^n \sigma_{ii}^X} = \frac{1}{n-1} \sum_{j=1}^n \frac{X_i(j) - \bar{X}_i}{X_l(j) - \bar{X}_l}, i, l = 1, \dots, p, p \in \mathbb{N}^*.$$

Le résultat important dans cet article, est le calcul de la distribution asymptotique de cette statistique de test.

Chen *et al.* (2011) montrent que la méthode de segmentation binaire (BSP) de Vostrikova (1981) et la méthode MRD de Cohen *et al.* (2009) détectent bien les points de ruptures dans le cadre multidimensionnel. Plus précisément, les auteurs considèrent le problème suivant : Soit X_{ij} , $i = 1, \dots, k$, $j = 1, \dots, n$ des v.a. indépendantes gaussiennes avec des moyennes μ_i et de variance connue σ^2 . Ils considèrent les hypothèses nulles $H_i : \mu_i - \mu_{i+1} = 0$ contre les alternatives $K_i : \mu_i - \mu_{i+1} \neq 0$ pour $i = 1, \dots, k-1$. Après avoir présenté les deux méthodes BSP et MRD, ils montrent que ces deux méthodes sont consistantes. C'est-à-dire que lorsque la taille des échantillons tend vers l'infini, la probabilité de faire une erreur (faux rejet ou fausse acceptation) tend vers zéro.

Le but de Zhou *et al.* (2014) est de développer des tests de spécification pour la régression non-paramétrique des séries chronologiques non-stationnaires. Ils considèrent le modèle suivant :

$$Y_i = X_i^T \beta(t_i) + \varepsilon_i, \quad i = 1, 2, \dots, n,$$

où $t_i = \frac{i}{n}$, $X_i = (x_{i1}, x_{i2}, \dots, x_{ip})^T$ est une matrice $p \times 1$, ε_i est le bruit satisfait $\mathbb{E}(\varepsilon_i | X_i) = 0$. Les processus $(X_i)_i$ et $(\varepsilon_i)_i$ sont non-stationnaires et ils peuvent être non-corrélés. Les paramètres de régression $\beta(\cdot) = (\beta_1(\cdot), \dots, \beta_p(\cdot))^T$ sont des fonctions à valeurs dans $[0, 1]$. L'auteur considère le problème de test suivant :

$$\begin{cases} H_0 : \beta(\cdot) = \beta_0(\cdot) \\ H_a : \beta(\cdot) \neq \beta_0(\cdot) \end{cases}$$

où $\beta_0(\cdot)$ est une fonction connue à valeur dans $[0, 1]$.

La statistique GLRT (Generalized Likelihood Radito Test) est la suivante :

$$\lambda_n = \frac{n}{2} \log \frac{RSS_0}{RSS_a} \approx -\frac{n}{2} \frac{RSS_a - RSS_0}{RSS_0},$$

où $RSS_0 = \sum_{i=1}^n (Y_i - X_i^T \beta_0(t_i))^2$ sous l'hypothèse nulle et $RSS_a = \sum_{i=1}^n (Y_i - X_i^T \hat{\beta}_0(t_i))^2$ sous l'hypothèse alternative, avec $\hat{\beta}(\cdot)$ est un estimateur de $\beta(\cdot)$.

Ici, le paramètre de changement est la fonction β qui est multipliée par X_i . Sous l'hypothèse que (X_i) est localement stationnaire, Zhou *et al.* (2014) étudient la convergence de λ_n sous H_0 .

On peut considérer aussi le problème de test du type :

$$\begin{cases} H_{01} : \beta^{(1)}(\cdot) = \beta_0^{(1)}(\cdot) \\ H_{a1} : \beta^{(1)}(\cdot) \neq \beta_0^{(1)}(\cdot) \end{cases}$$

où $\beta(t) = \begin{pmatrix} \beta^{(1)}(t) \\ \beta^{(2)}(t) \end{pmatrix}$, $\beta_0(t) = \begin{pmatrix} \beta_0^{(1)}(t) \\ \beta_0^{(2)}(t) \end{pmatrix}$ et $X_i = \begin{pmatrix} X_i^{(1)} \\ X_i^{(2)} \end{pmatrix}$.

$\beta^{(1)}(t)$, $\beta_0^{(1)}(t)$ et $X_i^{(1)}$ de dimension $p_1 < p$ et $\beta_0^{(1)}(t)$ une fonction connue.

Dans ce cas, sous H_{01} , les fonctions $\beta_j(\cdot)$; $j = p_1 + 1, \dots, p$ peuvent être estimées par $Y_i^* = Y_i - \left(\beta_0^{(1)}(t)\right)^T X_i^{(2)}$. Ici, Zhou *et al.* (2014) étudient le comportement asymptotique de λ_{1n} , la statistique du test de H_{01} contre H_{a1} . Ils étudient aussi la puissance locale sous les alternatives suivantes :

$$H_{a,n} : \beta(\cdot) = \beta_0(\cdot) + n^{-\frac{4}{9}} f_n(\cdot) \text{ où } f_n(\cdot) \in C^2[0, 1].$$

Ils proposent un autre type de test appelé test moyen (averaged test) basé sur λ_n^* et ils étudient sa loi asymptotique. Ils concluent que le test λ_n^* est plus puissant que le test basé sur le rapport de vraisemblance généralisé.

Dans Wang *et al.* (2012) le modèle non-linéaire étudié est le suivant :

$$y_{t-1} = f(x_t) + u_{t+1}, \quad t = 1, 2, \dots, n,$$

avec u_t est un processus stationnaire et (x_t) est une régression non-stationnaire.

Le but est de tester l'hypothèse nulle

$$H_0 : f(x) = f(x, \theta), \quad \theta \in \Omega_0,$$

où pour tout $x \in \mathbb{R}$, $f(x, \theta)$ est une fonction connue dépendant d'un paramètre inconnu $\theta \in \Omega_0$.

Le test est basé sur la statistique suivante :

$$S_n = \sum_{s,t=1, s \neq t}^n \hat{u}_{t+1} \hat{u}_{s+1} K\left(\frac{x_t - x_s}{h}\right),$$

où $\hat{u}_{t-1} = y_{t-1} - f(x_t, \hat{\theta})$, $K(x)$ est un noyau positif de paramètre de lissage $h \equiv h_n \rightarrow 0, n \rightarrow \infty$ et $\hat{\theta}$ est l'estimateur paramétrique de θ sous H_0 . La normalité asymptotique de S_n avec auto-normalisation est prouvée.

MacNeill (1974) considère une suite de variables aléatoires indépendantes $(X_j)_{j=1, \dots, n}$

d'une famille exponentielle à un paramètre inconnu θ_j . Donc cette distribution est de la forme

$$f(X_i|\theta_i) = h(X_i)\exp\{\psi_1(\theta_i)U(X_i) + \psi_2(\theta_i)\}, \quad i = 1, \dots, n,$$

avec $\psi_1(\cdot)$ et $\psi_2(\cdot)$ deux fonctions deux fois dérivables telles que

$$\psi_1'(0) \neq 0 \quad \text{et} \quad 0 < (\psi_1'(0))^{-1} (\psi_1''(\theta)\psi_2'(\theta) - \psi_2''(\theta)\psi_1'(\theta)) < \infty.$$

L'hypothèse de non-changement considéré est la suivante :

$$H_0 : \theta_i = \theta_0, \quad i = 1, \dots, n.$$

Pour cette étude, il considère le test du rapport de vraisemblance. Une approximation de la statistique du test est donnée sous l'hypothèse nulle de non-changement. Sa distribution asymptotique est donnée, en supposant que les θ_i , $i = 1, \dots, n$ prennent des valeurs proches de θ_0 . Cette distribution est celle de $\int \psi(t) [B(t) + H(t)]^2 dt$, où $B(t)$ est un mouvement brownien, $H(t)$ dépend des valeurs de θ_i , $i = 1, \dots, n$ et ψ est une fonction positive mesurable qui vérifie certaines conditions. La fonction caractéristique de la distribution de cette intégrale est exprimée en développement de Karhunen-Loève de $\psi^{\frac{1}{2}}B(t)$ et est présentée sous forme de tableau pour certaines ψ . Les résultats associés au pont brownien s'appliquent lorsque la valeur initiale θ_0 est inconnue et remplacée par un estimateur.

Sen et Srivastava (1975) considèrent X_1, X_2, \dots, X_N , N observations indépendantes. Le but est de savoir s'il y a un changement dans leur moyenne ou non. S'il y a un changement, alors les observations sont de la forme

$$X_i = \mu + \varepsilon_i, \quad i = 1, \dots, r$$

$$X_i = \mu^* + \varepsilon_i, \quad i = r + 1, \dots, N,$$

où $(\varepsilon_i)_i$ est une suite de v.a. indépendantes et r inconnu.

Dans cet article, les X_i sont supposées suivre la loi $\mathcal{N}(\mu_i, \sigma^2)$, pour $i = 1, \dots, N$.

Les auteurs considèrent les quatre problèmes suivants :

- Problème 1 : Tester

$$H : \mu_1 = \mu_2 = \dots = \mu_N = \mu$$

contre

$$A_1 : \mu_1 = \mu_2 = \dots = \mu_r < \mu_{r+1} = \dots = \mu_N,$$

où le point de rupture r est inconnu, μ est inconnue et σ est connue.

- Problème 2 : C'est le problème 1 mais dans le cas où μ est connue et égale à zéro.

- Problème 3 : Tester H contre

$$A : \mu_1 = \mu_2 = \dots = \mu_r \neq \mu_{r+1} = \dots = \mu_N,$$

où le point de rupture r est inconnu, μ est inconnue et $\sigma = 1$.

- Problème 4 : C'est le problème 3 mais dans le cas où μ est connue et égale à zéro.

Pour chaque type de problème, la statistique de test bayésien et la statistique de test du rapport de vraisemblance sont données. Les auteurs comparent les puissances de ces deux types de test pour chaque type de problème, et ils trouvent que dans le cas où $\Delta = \mu_N - \mu_1$ prend de petites valeurs, la puissance du test bayésien est au-dessus de la puissance du test du rapport de vraisemblance, et inversement lorsque Δ prend de grandes valeurs.

Dans le cas des problèmes 3 et 4, sous H , la distribution exacte du test bayésien est donnée et la distribution asymptotique du test bayésien pour le problème 3 est donnée par Gardner (1969).

Sen et Srivastava (1975) considèrent quelques tests non-paramétriques pour tester H contre A_1 et établissent que la puissance du test proposé est supérieure à celle du test de Bhattacharyya et Johnson (1968).

Vogelsang (1997) proposent des tests statistiques pour détecter une rupture en une date inconnue dans la tendance d'une série chronologique dynamique univariée. Le premier test considéré est basé sur la statistique de moyenne exponentielle de Andrews et Ploberger (1994) et le deuxième est basé sur la statistique du spectre de Andrews (1993). Ils montrent que la puissance du premier test est non-monotone par rapport à l'amplitude de la rupture et est dominée par celle de deuxième test.

Vogelsang (1997) applique ces deux tests sur plusieurs types des données macro-économiques et dans le plupart des cas, l'hypothèse nulle est rejetée. Donc les tests acceptent l'hypothèse d'une présence d'un point de rupture dans les données.

Vogelsang (1999) explique la raison de la non-monotonie de la puissance des tests considérés dans Vogelsang (1997). Il trouve que cette non-monotonie est due dans certain cas à l'estimation de la variance du bruit sous l'hypothèse alternative de changement dans la moyenne des observations. Et dans d'autres cas, cette non-monotonie est due à la présence d'une longue dépendance dans l'estimation de la régression. En particulier, Vogelsang (1999) introduit deux types des tests pour tester un changement dans la moyenne d'une série d'observations univariée. Le premier test est basé sur la statistique de régression (voir Vogelsang, 1999, Section 2.2) et le deuxième est basé sur la régression dynamique, c'est-à-dire la régression avec une large dépendance (voir Vogelsang, 1999, Section 2.3). Ensuite, il représente graphiquement la puissance de chaque type de tests et il trouve des conséquences pratiques énoncées dans la section 4 de l'article.

Horváth (2001) considère une série d'observations $(X_t)_{t \in \mathbb{Z}}$ longue mémoire. C'est-à-dire elle s'écrit de la forme $X_t = k_t \sum_{-\infty \leq j \leq +\infty} R(t-j, \lambda_t) \varepsilon_j$, où $(\varepsilon_t)_{t \in \mathbb{Z}}$ un bruit blanc tel que ε_0 est centré et de variance z^2 , k_t et λ_t sont des paramètres inconnus tels que $\lambda_t = (\lambda_{t,1}, \lambda_{t,2}, \dots, \lambda_{t,p})$ et R est une fonction telle que $\sum_{-\infty \leq j \leq +\infty} R^2(j, \lambda_t) < +\infty$. Dans cet article l'auteur considère le test ci-dessous pour détecter les ruptures dans les paramètres de la série d'observations :

$$H_0 : (k_1, \lambda_1) = (k_2, \lambda_2) = \dots = (k_n, \lambda_n)$$

contre

$$H_A : \exists k^*, 1 \leq k^* < n; (k_1, \lambda_1) = \dots = (k_{k^*}, \lambda_{k^*}) \neq (k_{k^*+1}, \lambda_{k^*+1}) = \dots = (k_n, \lambda_n).$$

Il calcule la loi asymptotique de la statistique du test considéré sous l'hypothèse nulle de non-changement et il étudie sa consistance.

Berkes *et al.* (2004) considèrent le modèle GARCH(p, q). Ils supposent qu'il présente un changement dans les paramètres. Pour détecter ce changement, ils considèrent un test basé sur une approximation de la vraisemblance. Ils déterminent la loi limite de la statistique du test sous des conditions faibles. Ce résultat est énoncé dans le Théorème 2.1 de l'article. L'un des tests les plus connus dans la littérature pour détecter les points de rupture est le CUSUM test. Dans la partie ci-dessous, on cite quelques exemples des travaux portant sur les CUSUM tests pour détecter un ou plusieurs points de rupture.

2.3.2 Le CUSUM test

Le CUSUM test standard est l'un des premiers tests de détection des points de rupture introduits par Brown *et al.* (1975). Celui basé sur les résidus des moindres carrés (noté $CUSUM^{ols}$) en est l'une des modifications. Dans Zeileis (2001, 2004), les deux tests sont introduits pour estimer les p-valeurs dans le cadre du modèle de régression linéaire standard du type

$$y_t = x_t^\top \beta + u_t, \quad t = 1, \dots, n, \quad (2.3)$$

où les y_t sont les observations de variables dépendantes, $x_t = (1, x_{t2}, x_{t3}, \dots, x_{tk})^\top$ est un vecteur $k \times 1$ des variables indépendantes, $(u_t)_t$ une suite de variables aléatoires iid centrées et de variance σ^2 et β est le vecteur $k \times 1$ des coefficients de régression.

Le CUSUM test standard est basé sur la somme cumulée des résidus, la statistique du test est la suivante :

$$w_n(t) = \frac{1}{\tilde{\sigma} \sqrt{n-k}} \sum_{i=k+1}^{[k+t(n-k)]} \tilde{u}_i, \quad 0 \leq t \leq 1,$$

où $\tilde{\sigma} = \left(\frac{1}{n-k} \sum_{i=k+1}^n (\tilde{u}_i - \bar{\tilde{u}})^2 \right)^{1/2}$, $\tilde{u}_i = \frac{y_i - x_i^\top \hat{\beta}^{(i-1)}}{\sqrt{1 + x_i^\top (\sum_{t=1}^{i-1} x_t x_t^\top)^{-1} x_i}}$, $i = k+1, \dots, n$,

$\hat{\beta}^{(i-1)}$ est l'estimateur des moindres carré de β , k est le nombre des composantes du vecteur x_t et $t \in [0, 1]$ est l'instant de rupture.

H_0 est rejetée à chaque fois que $w_n(t)$ croit vers $c(t)$ ou $-c(t)$, avec $c(t) = \lambda + 2\lambda t = \lambda(1 + 2t)$, où $\lambda \in]0, 1[$. Ce qui est équivalent à rejeter H_0 lorsque la statistique du test

$S = \sup_{0 \leq t \leq 1} \left| \frac{w_n(t)}{1 + 2t} \right|$ est supérieure à λ .

Krämer *et al.* (1988) montrent que lorsque $n \rightarrow +\infty$,

$$w_n(t) \xrightarrow{loi} B(t),$$

où $B(t)$ est un mouvement brownien standard.

Le $CUSUM^{ols}$ test est défini en utilisant l'estimateur des moindres carrés des résidus

$\hat{u}_t = y_t - x_t^\top \hat{\beta}$. Pour $t \in [0, 1]$, sa trajectoire est définie comme suit :

$$w_n^0(t) = \frac{1}{\hat{\sigma} \sqrt{n}} \sum_{i=1}^{[nt]} \hat{u}_i,$$

où $\hat{\sigma} = \sqrt{\frac{1}{n-k} \sum_{t=1}^n \hat{u}_t^2}$.

On rejette H_0 lorsque la statistique du test $S^0 = \sup_{0 \leq t \leq 1} \left| \frac{w_n^0(t)}{1 + 2t} \right|$ est supérieure à $\lambda \in]0, 1[$.

Krämer et Schotman (1992) montrent que lorsque $n \rightarrow +\infty$,

$$w_n^0(t) \xrightarrow{loi} B^0(t),$$

où $B(t)$ est un pont brownien standard.

La p -valeur est définie comme suit : $\mathbb{P}_{H_0}(S \geq s)$. Donc la p -valeur de CUSUM test standard est

$$\mathbb{P}_{H_0}(S \geq s) = \mathbb{P}_{H_0}(|w_n(t)| \geq s + 2ts, \quad 0 \leq t \leq 1).$$

La p -valeur asymptotique considérée par Zeileis (2001, 2004) est

$$P(s) = \mathbb{P}_{H_0}(|B(t)| \geq s + 2ts, \quad 0 \leq t \leq 1),$$

où s est le seuil de signification.

De même la p -valeur asymptotique pour le test $CUSUM^{ols}$ est

$$P'(s) = \mathbb{P}_{H_0}(|B^0(t)| \geq s + 2ts, \quad 0 \leq t \leq 1).$$

Zeileis (2001, 2004) donne deux bornes (borne supérieure et borne inférieure) approximatives de la p -valeur pour chaque test (le CUSUM test standard et $CUSUM^{ols}$ test). À l'aide des simulations numériques, il vérifie que ces bornes sont de bonnes approximations pour la p -valeur. Cependant, l'un des inconvénients majeurs est qu'ils ont une faible puissance lorsque les changements ont lieu dès les premières observations ou lorsqu'ils ont lieu à la fin (changements tardifs). Pour cela, Zeileis (2001, 2004) utilise une borne alternative pour chaque type de test. Pour le CUSUM test standard, il prend $b(t) = \lambda\sqrt{t}$, où $\lambda \in]0, 1[$, et pour le $CUSUM^{ols}$ test, $d(t) = \lambda\sqrt{t(1-t)}$.

Ces limites alternatives proportionnelles à l'écart-type des distributions limites n'améliorent pas les propriétés du CUSUM test standard. Cependant, ils peuvent mieux détecter les ruptures que le $CUSUM^{ols}$ test lorsque les changements ont lieu dès les premières observations ou lorsqu'ils ont lieu à la fin.

Dans ce contexte, Horváth *et al.* (2020) proposent une modification du CUSUM test standard. Cette modification est très efficace à la détection de rupture lorsqu'elle se produit dans les premières ou les dernières observations. Le test proposé est inspiré par Rényi (1953). Nous donnons plus de détails sur ce test dans le chapitre 5 pour faire la comparaison avec leur méthode.

Deng et Perron (2008) considèrent le même modèle que Zeileis (2001, 2004). Ils présentent les deux types de CUSUM tests et ils étudient une autre modification notée CUSQ. C'est une alternative introduite par Brown *et al.* (1975). La statistique de ce CUSQ test est la suivante :

$$CUSQ = \max_{k+1 < r \leq T} \sqrt{T} \left| S_T^{(r)} - \frac{r-k}{T-k} \right|,$$

où $S_T^{(r)} = (\sum_{t=k+1}^r \tilde{u}_t^2) / (\sum_{t=k+1}^T \tilde{u}_t^2)$ et T le nombre d'observations.

Ce test, construit avec les résidus des moindres carrés, a une puissance nulle lorsque le point de rupture se produit au milieu de l'échantillon.

Le modèle considéré dans Deng et Perron (2008) est le suivant :

$$y_t = x_t^\top \beta + u_t, \quad t = 1, \dots, T, \quad \beta_t = \beta + \delta \mathbb{1}_{(t > [T\lambda])}, \quad (2.4)$$

pour $0 < \lambda < 1$. Ce modèle est un cas particulier de celui de Zeileis (2001, 2004).

Ploberger et Krämer (1990) prouvent que la puissance du CUSUM test est supérieure à la puissance du CUSQ test qui, cependant, croît rapidement lorsque l'amplitude de la rupture augmente. Pour une rupture prématurée, la puissance avec ols (moindres carrés ordinaires) est plus élevée qu'avec les résidus récursifs et inversement pour une rupture tardive.

Sous certaines conditions portant sur $(x_t)_t$ et $(u_t)_t$, Deng et Perron (2008) établissent des résultats asymptotiques pour les $CUSQ^{ols}$, CUSQ, $CUSUM^{ols}$ et CUSUM tests. Les

résultats impliquent d'abord que les puissances des tests sont monotones, en ce sens que lorsque $\|\delta\|$ prend des grandes valeurs, la limite n'est pas décroissante (ceci est valable en général pour les CUSQ tests et pour les CUSUM tests, il est valable pour les cas considérés dans les simulations de l'article mais peut ne pas être vrai en général). Ils ont étudié aussi le cas d'un CUSQ test avec une erreur non-gaussienne.

D'après Ploberger *et al.* (1986), si les erreurs sont non corrélées et conditionnellement homoscédastiques, on a

$$CUSQ \xrightarrow{loi} \frac{\sqrt{\gamma}}{\sigma^2} \sup_{r \in [0,1]} |BB(r)|,$$

où $BB(r)$ est un pont brownien et $\gamma = \mathbb{E}(u_t^2 - \sigma^2)^2 = \mu_4 - \sigma^4$, avec $\mu_4 = \mathbb{E}(u_t^4)$.

Le résultat est vrai dans le cas où les résidus sont estimés par la méthode des moindres carrés ou par une méthode récursive car

$$CUSQ_M = \frac{\hat{\sigma}^2}{\sqrt{\hat{\gamma}}},$$

où $\hat{\gamma} = \hat{\mu}_4 - \hat{\sigma}^4$, avec $\hat{\mu}_4 = \frac{1}{T} \sum_{t=1}^T \bar{u}_t^4$, où \bar{u}_t^4 est un estimateur de résidu.

Deng et Perron (2008) donnent une représentation asymptotique équivalente au test CUSQ dans le cas où le bruit n'est pas gaussien qui est la suivante :

$$CUSQ_M = \frac{\max_{1 \leq r \leq T} \left| T^{-\frac{1}{2}} \sum_{s=1}^r \bar{u}_s^4 - \frac{[Tr]}{T} \sum_{s=1}^T \bar{u}_s^4 \right|}{\sqrt{\hat{\gamma}}}.$$

Le Théorème 2 de Deng et Perron (2008) montrent que la puissance de $CUSQ^{ols}$ test augmente de façon monotone, contrairement au cas où le bruit est gaussien. Aussi, les auteurs étudient le cas où les résidus sont corrélés et non-gaussiens. Dans ce cas, on a

$$CUSUM^{rec} \xrightarrow{loi} \left(\frac{h(0)}{\sigma^2} \right)^{\frac{1}{2}} \sup_{0 \leq r \leq 1} \left| \frac{w(r)}{1 + 2r} \right|,$$

où $h(0) = \lim_{T \rightarrow \infty} Var \left(T^{-\frac{1}{2}} \sum_{t=1}^T u_t \right)$.

Pour des résidus récurrents ou des moindres carrés, on a

$$CUSQ \xrightarrow{loi} \frac{\varphi^{\frac{1}{2}}}{\sigma^2} \sup_{0 \leq r \leq 1} |BB(r)|,$$

où $\varphi = \lim_{T \rightarrow \infty} Var \left(T^{-\frac{1}{2}} \sum_{t=1}^T \xi_t \right)$, $\xi_t = u_t^4 - \sigma^4$.

Le test modifié vérifie la relation suivante :

$$CUSUM_S = \left(\frac{\hat{\sigma}^2}{\hat{h}(0)} \right)^{\frac{1}{2}} CUSUM,$$

où $\widehat{h}(0)$ est un estimateur consistant de $h(0)$.

Deng et Perron (2008) montrent que si on considère l'estimateur basé sur

$$\widehat{w}_{ws}^2 = \widehat{\gamma}_0 + 2 \sum_{j=1}^{T-1} K(j, m) \widehat{\gamma}_j,$$

où $\widehat{\gamma}_j = \frac{1}{T} \sum_{t=j+1}^T \bar{u}_t \bar{u}_{t-j}$, on a $CUSUM_S = \left(\frac{\widehat{\sigma}^2}{\widehat{w}_{ws}^2} \right)^{\frac{1}{2}} CUSUM$. Ainsi,

$$CUSQ_S = \frac{\max_{1 \leq r \leq T} \left| T^{-\frac{1}{2}} \sum_{s=1}^r \bar{u}_s^4 - \frac{[Tr]}{T} \sum_{s=1}^T \bar{u}_s^4 \right|}{\sqrt{\widehat{\varphi}}},$$

où $\widehat{\varphi} = (1/T) \sum_{j=-(T-1)}^{T-1} K(j, m) \sum_{t=|j|+1}^T \widehat{\eta}_t \widehat{\eta}_{t-j}$ et $\widehat{\eta}_t = \bar{u}_t^2 - \widehat{\sigma}^2$.

Sous certaines conditions, ils établissent des résultats asymptotiques pour les tests $CUSUM_S^{ols}$, $CUSUM_S^{rec}$, $CUSQ_S^{ols}$ et $CUSQ_S^{rec}$. Ensuite, ils étudient le cas d'une régression dynamique, en considérant le modèle de régression suivant :

$$y_t = \alpha y_{t-1} + \beta_t x_t^\top + u_t,$$

où $\beta_t = \beta + \sigma \mathbb{1}_{(t > [T\lambda])}$, pour $0 < \lambda < 1$.

Le modèle de régression dynamique s'écrit comme suit :

$$y_t = z_t^\top (\phi + \gamma \mathbb{1}_{(t > [T\lambda])}) + u_t,$$

où $z_t^\top = (y_{t-1}, x_t^\top)$, $\phi = (\alpha, \beta)^\top$ et $\gamma = (0, \sigma)^\top$.

Des égalités asymptotiques sont prouvées dans le Théorème 4 de l'article, pour les tests $CUSUM^{rec}$, $CUSQ^{rec}$, $CUSUM^{ols}$ et $CUSQ^{ols}$ dans ce cas. Les auteurs traitent aussi le cas d'un changement dans la moyenne. Ils établissent des égalités asymptotiques pour les statistiques des $CUSUM^{rec}$, $CUSQ^{rec}$, $CUSUM^{ols}$ et $CUSQ^{ols}$ tests. Ces résultats sont énoncés dans la Proposition 1 de l'article. Les Propositions 2 et 3 impliquent que lorsque la rupture se produit au milieu de l'échantillon, la puissance du CUSQ test vaut 0.

Pour conclure cet article, les auteurs étudient les CUSUM tests et les CUSQ tests dans plusieurs scénarios dans un modèle de régression linéaire. Les implications pratiques de leurs résultats sont les suivantes :

s'il s'agit d'une régression sans corrélation dans les erreurs, le CUSUM test est le test préféré. Si les erreurs sont corrélées, les deux tests considérés ont de mauvaises propriétés et lorsqu'il s'agit d'un changement dans la moyenne, ils n'arrivent pas à le détecter. Ici, il est nécessaire d'ajouter des modifications sur le modèle de régression considéré. Le CUSQ test est préférable pour la plupart des types de modifications, à condition que les erreurs

soient normalement distribuées. Si les erreurs sont non-gaussiennes, le CUSQ test perd sa supériorité et les deux tests ont des propriétés médiocres sauf dans des cas spéciaux.

Dans Lee *et al.* (2003), les auteurs examinent le problème d'un changement de paramètre dans les modèles GARCH(1,1) basé sur le CUSUM test. Ils montrent que la statistique du test proposé a une distribution limite d'un sup d'un pont brownien standard. Plus précisément, ils considèrent le modèle GARCH(1,1) suivant :

$$\begin{aligned} y_t &= c + \varepsilon_t, \\ \varepsilon_t &= h_t \cdot \xi_t, \\ h_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}^2, \end{aligned}$$

où $(\xi_t)_t$ est une suite de variables aléatoires iid centrées et de variance 1, c est un nombre réel et ω , α et β sont des réels positifs tels que $\alpha + \beta < 1$. De plus, $\mathbb{E}(|\varepsilon_t|^{4+\delta}) < \infty$ et $\mathbb{E}(|\xi_t|^{4+\delta}) < \infty$, pour $\delta > 0$ et ξ_t indépendant de y_s pour $s < t$. L'objectif de cet article est de tester

$$H_0 : \theta = (\omega, \alpha, \beta) \text{ ne change pas} \quad \text{contre} \quad H_1 : \text{il y a un changement}$$

Le CUSUM test basé sur les résidus des moindres carrés est instable (d'après Inclan et Tiao (1994) et Kim *et al.* (2000)) en raison de la volatilité, et a de faibles puissances. Pour améliorer ça, il faut envisager de développer un nouveau test. Selon Lee *et al.* (2003), on peut utiliser le CUSUM test basé sur (ξ_t^2) et fondé sur la statistique

$$T_n = \frac{1}{\sqrt{n\tau}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \xi_t^2 - \frac{k}{n} \sum_{t=1}^n \xi_t^2 \right|,$$

où $\tau = \text{Var}(\xi_t^2)$.

Puisque ξ_t n'est pas observable, on remplace ξ_t^2 dans T_n par son estimateur $\hat{\xi}_t^2 = (y_t - \hat{c})^2 / \hat{h}_t^2$ obtenue par estimation des constantes ω, α, β , et c . On peut maintenant anticiper que le CUSUM test récursif sera plus stable et produira des meilleurs puissances. Le test CUSUM construit est basé sur la statistique suivante :

$$\hat{T}_n = \frac{1}{\sqrt{n\hat{\tau}}} \max_{1 \leq k \leq n} \left| \sum_{t=1}^k \hat{\xi}_t^2 - \frac{k}{n} \sum_{t=1}^n \hat{\xi}_t^2 \right|,$$

$$\text{où } \hat{\tau} = \frac{1}{n} \sum_{t=1}^n \hat{\xi}_t^4 - \left(\frac{1}{n} \sum_{t=1}^n \hat{\xi}_t^2 \right)^2.$$

Les auteurs montrent un résultat asymptotique pour \hat{T}_n énoncé dans le Théorème 1 de l'article. À l'aide des applications numériques, ils montrent que ce test est préférable au test initial car il a une puissance meilleure que celle du CUSUM test initial.

Andreou (2008) propose une estimation du CUSUM test pour résoudre le problème de la non-monotonie de la puissance du test pour détecter un seul point de rupture dans la moyenne d'un processus faiblement dépendant. Pour un échantillon fini, la puissance du CUSUM test peut être non-monotone et même nulle. Selon Andreou (2008), la non-monotonie de la puissance est due à l'estimation de la variance. Il propose d'utiliser l'estimateur HAC (Heteroskedastic and Autocorrelation Consistent estimator) pour résoudre ce problème sous l'alternative d'existence d'une rupture.

Andreou (2008) considère le modèle de régression linéaire suivant :

$$y_t = \mu + u_t, \quad t = 1, \dots, T,$$

où $(u_t)_t$ est un processus stationnaire centré et du second ordre.

La statistique du CUSUM test considéré pour détecter un point de rupture dans la moyenne de y_t est le suivant :

$$CUSUM = \frac{1}{\sigma\sqrt{T}} \sup_{1 \leq j \leq T} \left| \sum_{t=1}^j y_t - \sum_{t=1}^T y_t \right| \xrightarrow{loi} \sup |B(m)|,$$

où $m \in \mathbb{N}$, $\sigma = \lim_{T \rightarrow \infty} \mathbb{E} \left[\frac{1}{T} \left(\sum_{t=1}^T u_t \right)^2 \right]$, $B(m) = W(m) - mW(1)$, avec $W(m)$ est le processus de Wiener standard défini dans $[0, 1]$.

Le résidus des moindres est donné par

$$\hat{u}_t^{ols} = y_t - \frac{1}{T} \sum_{t=1}^T y_t.$$

La statistique du CUSUM test correspondante est

$$CUSUM = \frac{1}{\hat{\sigma}\sqrt{T}} \sup_{1 \leq j \leq T} \sum_{t=1}^j \hat{u}_t^{ols},$$

où $\hat{\sigma}$ est un estimateur consistant de σ sous l'hypothèse nulle.

On peut utiliser $\hat{\sigma}_{pw}^2$ comme estimateur de σ^2 proposé par Andrews et Monahan (1992). Dans l'expression de $\hat{\sigma}_{pw}^2$, il y a un paramètre noté $\hat{\rho}$ (voir l'article).

Andrews et Monahan (1992) proposent de remplacer chaque valeur de $\hat{\rho}$ supérieure à 0.97 par 0.97 et chaque valeur de $\hat{\rho}$ inférieure à -0.97 par -0.97. En utilisant cette règle, le CUSUM test présente une puissance non-monotone.

Sul *et al.* (2005) proposent une autre règle basée sur T en utilisant l'estimateur $\hat{\rho}^\top = \min \left[1 - 1/\sqrt{T}, \hat{\rho} \right]$.

Andreou (2008) généralise l'estimateur de Sul *et al.* (2005) par $\hat{\rho}_*^\top = \min \left[1 - c/\sqrt{T}, \hat{\rho} \right]$.

D'un point de vue théorique, pour tout $c > 0$, $\widehat{\rho}_*^\top$ est un estimateur consistant $\widehat{\rho}$ sous l'hypothèse nulle et l'hypothèse alternative. Les résultats de simulation montrent que l'utilisation de l'estimateur $\widehat{\rho}_*^\top$ peut corriger le problème de la non-monotonie de la puissance et, dans ce cas, l'estimateur HAC devient plus efficace à la fois sous H_0 et sous l'hypothèse alternative.

Le CUSUM test a joué un rôle important en théorie et dans les applications liées à la détection des points de rupture, mais son inconvénient est qu'il perd de la puissance lorsque la rupture a lieu dans la moyenne d'un modèle de régression.

Horváth et Steinebach (2000) considèrent le modèle suivant :

$$Z_t = \begin{cases} at + bY_t & \text{pour } 1 \leq t \leq T^* \\ Z_{T^*} + a^*(t - T^*) + b^*Y_{t-T^*}^* & \text{pour } T^* < t \leq T, \end{cases}$$

où $(Z_t)_{0 \leq t \leq \infty}$ est la série d'observations, a, b, a^*, b^* et T^* sont des paramètres inconnus, $(Y_t)_{0 \leq t \leq \infty}$ et $(Y_t)_{0 \leq t \leq \infty}^*$ sont des processus stochastiques inobservables satisfaisants chacun le principe d'invariance faible (voir Horváth et Steinebach, 2000, (1.2)-(1.3)) et T est le nombre des observations.

Horváth et Steinebach (2000) considèrent le CUSUM test pour détecter un changement dans la moyenne ou dans la variance du processus $(Z_t)_t$ ci-dessus. En particulier, pour étudier le changement dans la moyenne des observations, ils considèrent le problème de test suivant :

$$H_0 : T^* = T \text{ contre } H_A^m : 0 < T^* < T \text{ et } a \neq a^*.$$

D'une façon similaire, pour détecter le changement dans la variance des observations, ils considèrent le test de H_0 contre

$$H_A^\nu : 0 < T^* < T \text{ et } b \neq b^*.$$

Pour chaque cas étudié, ils trouvent la loi asymptotique sous H_0 et sous l'hypothèse alternative, de la statistique du test considéré.

Dans Jiang et Kurozumi (2019) deux CUSUM tests modifiés sont considérés pour détecter un changement et pour examiner les propriétés de la puissance asymptotique de ces tests sous des alternatives fixes. Les auteurs montrent que la puissance de ces tests est asymptotiquement supérieure à la puissance des tests classiques. Leur étude est basée sur le modèle

$$y_t = x_t^\top \beta_t + u_t, \quad t = 1, \dots, T,$$

où $x_t = (1, x_{2t}, x_{3t}, \dots, x_{kt})^\top$, $(u_t)_t$ est un processus non observable et β_k est un vecteur de dimension k , $k \in \mathbb{R}$.

Les hypothèses à tester sont

$$H_0 : \beta_t = \beta, \quad \forall t \quad \text{contre} \quad H_1 : \beta_t = \beta + \delta \mathbf{1}_{(t > [T\lambda])},$$

où $\lambda \in]0, 1[$.

Tout d'abord, ils présentent le CUSUM test initial dans le cas où il est basé sur les résidus des moindres carrés, et dans le cas où il est basé sur les résidus récurrents. Ensuite, ils considèrent les CUSUM tests modifiés basés sur $\hat{\omega}_{jt} = x_{jt}\hat{u}_t$ et $\tilde{\omega}_{jt} = x_{jt}\tilde{u}_t$ et respectivement notés $CUSUM_m^{ols}$ et $CUSUM_m^{rec}$. Ils établissent des résultats asymptotiques pour chaque type de test. Ces résultats sont énoncés dans la Proposition 1 de l'article. Cette proposition montre que la modification pourrait bien fonctionner.

De plus, ils proposent quatre autres modifications sur les CUSUM tests standards notés dans l'article Q^{ols} , Q^{rec} , M^{ols} et M^{rec} . Ils étudient les comportements asymptotiques pour chaque test. Les résultats sont énoncés dans les Théorèmes 1 et 2 de l'article.

Dans le cas où les erreurs sont corrélées, des tests modifiés sont proposés et le Théorème 3 nous donne leur comportement asymptotique. Cette modification entraîne une réduction de la puissance du test. Des comparaisons de ces tests sont faites à l'aide des simulations numériques.

Pour conclure, lorsqu'un changement a lieu dans la moyenne d'une régression, les tests CUSUM standards perdent de leur puissance. En conséquence, plusieurs CUSUM tests modifiés ont été proposés explicitement ou implicitement dans la littérature pour détecter les points de ruptures. Jiang et Kurozumi (2019) étudient les propriétés asymptotiques de tels tests modifiés et montrent qu'ils peuvent rejeter avec succès l'hypothèse nulle. En ce sens, les tests modifiés pourraient compléter les CUSUM tests standards.

Horváth *et al.* (2004) considèrent le modèle (2.3), le même que dans Jiang et Kurozumi (2019) et Zeileis (2001, 2004). Ils proposent deux classes de systèmes de surveillance pour détecter (séquentiellement) un changement dans ce modèle linéaire après une période de taille m . La première catégorie de procédures est basée sur les CUSUM tests, dans lesquels le paramètre inconnu de contrôle a été remplacé par son estimation des moindres carrés, alors que la seconde classe utilise les CUSUM tests basés sur les résidus récurrents. La fonction de pondération peut être choisie de manière extensible en fonction d'un changement anticipé ou tardif après le temps m . Les procédures sont conçues pour que les tests aient une faible probabilité de fausse alarme (lorsque le nombre des observations est assez grand i.e $m \rightarrow \infty$) et de puissance asymptotique. Une petite étude de simulation illustre les performances de la méthode considérée en utilisant plusieurs choix de fonctions de pondération.

Hušková *et al.* (2008) étudient le même modèle que Deng et Perron (2008). Ils obtiennent par diverses méthodes de bootstrap, des approximations des valeurs critiques pour les tests de points de rupture. Les résultats théoriques montrent que les procédures de bootstrap ont les mêmes propriétés asymptotiques que ceux de Hušková *et al.* (2007). Une étude de simulation montre que les tests bootstrap se comportent mieux que les tests asymptotiques

initiaux si les performances sont mesurées par les erreurs de première et de deuxième espèces.

Hušková et Koubková (2005) considèrent le modèle (2.3) dans un cadre multidimensionnel. Ils étudient les deux types de CUSUM test, le test basé sur les erreurs des moindres carrés et le test basé sur les erreurs récurrents, pour détecter un point de rupture. Leurs propriétés sont étudiées lorsqu'il n'y a pas de rupture et ces résultats asymptotiques sont vérifiés par une étude de simulation.

Horváth *et al.* (2006) proposent des méthodes pour détecter les changements dans la variance inconditionnelle dans une série d'observations en ligne. Selon ces auteurs, cette procédure est applicable à la classe des modèles GARCH sous certaines conditions. Ils présentent deux méthodes. La première est basée sur les test CUSUM, et la deuxième sur la somme partielle des résidus et des résidus récurrents. Donc au total il y a quatre procédures nommées dans l'article CS, FL, PS et RR respectivement. Horváth *et al.* (2006) comparent ces méthodes par des simulations numériques. Ils montrent l'efficacité et la performance de chaque méthode.

Dans Davis *et al.* (1995), le problème de la détection de rupture est étudié. Les auteurs établissent que si le bruit de ce modèle est faiblement stationnaire avec un moment d'ordre 4 fini, sous l'hypothèse nulle, la statistique de test du rapport de vraisemblance gaussien normalisée converge en loi vers la distribution des valeurs extrêmes de Gumbel. Le modèle considéré est le suivant :

$$X_t = \begin{cases} \phi_0 + \phi_1 X_{t-1} + \dots + \phi_p X_{t-p} + \varepsilon_t, & -\infty < t \leq \tau, \\ \alpha_0 + \alpha_1 X_{t-1} + \dots + \alpha_p X_{t-p} + \varepsilon_t, & t \geq \tau + 1, \end{cases}$$

où $\tau \in]p, n]$, $(\varepsilon_t)_t$ est un bruit de moment d'ordre 4, n est le nombre d'observations, p l'ordre de modèle et τ l'instant de rupture.

Ils proposent un test du rapport de vraisemblance pour détecter ce point de rupture. Ils montrent un résultat asymptotique sous l'hypothèse nulle, pour la statistique de test considérée. Ils étudient le cas où la série d'observations $(X_t)_{t \geq 0}$ est issue du modèle

$$X_t = \begin{cases} \phi_0 + \phi_1 X_{t-1} + \dots + \phi_{p_0} X_{t-p_0} + \varepsilon_t, & -\infty < t \leq \tau, \\ \alpha_0 + \alpha_1 X_{t-1} + \dots + \alpha_{p_1} X_{t-p_1} + \varepsilon_t, & t \geq \tau + 1, \end{cases}$$

où $\tau \in]p_1, n]$, $p_0 \leq p_1$. Sous la condition que $\mathbb{E}(\varepsilon_t^3) = 0$ et $\mathbb{E}(\varepsilon_t^4) = 3\sigma^4$, où $\sigma^2 = Var(\varepsilon_t)$, ils proposent un test statistique noté Λ'_n et sous l'hypothèse de non changement, la loi asymptotique de Λ'_n est obtenue. Ensuite, ils refont la même chose dans le cas où $\mathbb{E}(\varepsilon_t^3) = 0$ et $\mathbb{E}(\varepsilon_t^4) \neq 3\sigma^4$.

2.3.3 La segmentation binaire

La segmentation binaire est une autre méthode de détection de rupture. Nous résumons ici quelques articles qui étudient cette méthode.

Vostrikova (1981) montre la consistance de cette méthode lorsque le nombre des ruptures est connu. Le travail est fait dans un cadre multivarié. Le test utilisé à chaque étape de la méthode est un CUSUM test, mais le critère d'arrêt proposé n'est pas facile à calculer en pratique. Pour cela, plusieurs améliorations ont été proposées. Olshen *et al.* (2004) proposent une modification de la segmentation binaire (BS) de Vostrikova (1981). Leur méthode est nommée segmentation binaire circulaire (CBS). Ils testent l'hypothèse de non-changement dans la moyenne des données contre l'hypothèse d'une présence de ruptures dans les extrémités. C'est-à-dire lorsque le segment est épissé aux deux extrémités pour former un cercle. Dans ce cas, la statistique de test du rapport de vraisemblance pour tester l'hypothèse que l'arc de $i + 1$ à j et son complément ont des moyennes différentes est nommée $Z_{i,j}$. La méthode de Olshen *et al.* (2004) est basée sur la statistique $Z = \max_{1 \leq i, j \leq n} |Z_{i,j}|$, avec n est le nombre de données.

La méthode est proposée pour identifier les altérations génomiques dans les expériences de nombre de copies de tableaux d'ADN. Olshen *et al.* (2004) appliquent leur méthode pour copier les données numériques des expériences aCGH sur 15 lignes cellulaires de fibroblastes et ils l'appliquent également aux données ROMA de 23 lignées cellulaires de cancer du sein. Ils observent des fausses détections dues à la présence des tendances locales. Cette méthode est évaluée par des simulations.

Fryzlewicz *et al.* (2014) proposent une nouvelle technique pour estimer le nombre des ruptures et leurs localisations dans une série stochastique univariée et non-stationnaire déterministe. Cette technique est appelée Wild Binary Segmentation (WBS). C'est une modification de la méthode de segmentation binaire standard (BS). Fryzlewicz *et al.* (2014) prouvent la consistance de leur méthode et montrent qu'elle fonctionne même lorsque les ruptures sont très proches et/ou de petites amplitudes contrairement à la méthode BS. Les deux types de méthodes (la méthode BS et la méthode WBS) sont basées sur la statistique CUSUM. Fryzlewicz *et al.* (2014) proposent deux critères d'arrêt pour la WBS. Le premier est basé sur le seuil critique et le deuxième sur le critère "strengthened Schwarz Information Criterion" (sSIC). Ils appliquent aussi leur méthode à la détection des ruptures dans la tendance des 500 indices de S&P qui est un indice boursier basé sur 500 grandes sociétés cotées sur les bourses aux États-Unis. Ils considèrent en particulier la série log-returns des données de S&P 500 quotidiens durant 8 ans à partir du 16 octobre 2012. Ils montrent qu'ils ont de bonnes performances pratiques. De plus, cette méthode est facile à coder sous R. Elle est implémentée avec le package *wbs* dans CRAN.

Une autre modification de la WBS existe. C'est la méthode dite NOT proposée par Baranowski *et al.* (2018). Elle consiste à estimer le nombre des ruptures et leurs localisations dans une série stochastique stationnaire par morceaux. Les auteurs étudient pour cela le modèle statistique univarié, celui considéré par Fryzlewicz *et al.* (2014), suivant :

$$Y_t = f_t + \varepsilon_t, \quad t = 1, \dots, T,$$

où $(Y_t)_t$ est la série observée, f_t est une fonction inconnue vérifiant certaines conditions, $(\varepsilon_t)_t$ est un bruit blanc centré et T est le nombre des observations.

Baranowski *et al.* (2018) considèrent le cas où les changements sont en f_t lorsqu'elle est constante par morceaux, en f'_t lorsque f_t est linéaire et continue par morceaux et en la discontinuité de f_t lorsqu'elle est discontinue ou en sa première dérivée lorsque f_t est linéaire par morceaux mais sans la contrainte de continuité. Ils traitent le cas où $(\varepsilon_t)_t$ est un processus gaussien et le cas où il n'est pas gaussien. Leur objectif est de détecter les ruptures dans la moyenne et/ou la variance de la série étudiée.

La méthode NOT a comme avantages sa faible complexité de calcul, sa facilité de mise en œuvre et sa précision. Cette méthode et la méthode WBS ont la même idée mais elles sont différentes et ont des objectifs différents. En effet, la méthode NOT se concentre sur les intervalles les plus étroits des données sur lesquelles il est possible de localiser la rupture. C'est cette localisation qui permet de ne pas aller bien au-delà de la simple détection lorsque f_t est constante par morceaux. Cette dernière est le seul objectif de la méthode WBS.

L'absence de mise au point de l'intervalle le plus étroit dans les méthodes BS et WBS signifie qu'elles ne sont pas applicables à la détection des ruptures dans un cadre général. Les simulations numériques faites dans Baranowski *et al.* (2018) montrent que la méthode NOT fait partie des méthodes les plus performantes en terme d'estimation du nombre de ruptures et de leurs localisations, ce d'autant qu'elle est plus rapide à calculer. Ce qui lui donne un avantage particulier par rapport aux autres méthodes existantes dans la littérature comme par exemple la méthode WBS de Fryzlewicz *et al.* (2014), la méthode NMCD de Zou *et al.* (2014), la méthode SMVCE de Frick *et al.* (2014) et la méthode FDRSeg de Li *et al.* (2016) dans certains scénarios considérés dans Baranowski *et al.* (2018). Plus particulièrement, les scénarios (S2), (S3) et (S5) de ce dernier article.

Baranowski *et al.* (2018) appliquent aussi leurs méthodes aux données réelles de la température d'anomalie globale et aux données des logements à Londres. L'implémentation de la méthode NOT est fournie dans le package R *not* (voir Baranowski *et al.* (2016)).

Anastasiou et Fryzlewicz (2019) introduisent une autre approche appelé ID (Isolate-Detect) pour estimer le nombre et les localisations des ruptures dans des suites de données. Cette méthode est basée sur la technique d'isolation, ce qui empêche la prise en compte

d'intervalles contenant plus d'une rupture. La méthode ID peut détecter les ruptures de petites amplitudes. Elle est implémentée sous R avec le package *IDetect* valable dans CRAN. Anastasiou et Fryzlewicz (2019) considèrent plus particulièrement le modèle suivant :

$$X_t = f_t + \sigma\varepsilon_t, \quad t = 1, \dots, T,$$

où $(X_t)_t$ est la série des observations, $(\varepsilon_t)_t$ est un bruit blanc et f_t est une fonction supposée dans un premier temps constante par morceaux et dans un second temps linéaire par morceaux avec ou sans la contrainte de continuité.

L'avantage de la méthode ID par rapport à la méthode NOT et la méthode WBS c'est qu'elle est plus rapide. Aussi, dans NOT et WBS, il n'est pas certain que les intervalles qu'ils prennent couvrent tout le domaine de données sans ignorer les zones dans lesquelles il y a des points de ruptures. En particulier, dans le cas où la série présente un très grand nombre des ruptures, les méthodes NOT et WBS doivent augmenter le nombre M des intervalles dessinés. Cela augmente le temps de calcul. En revanche, la méthode ID proposée couvre tous les emplacements des ruptures et ne nécessite aucun choix de M . Ce qui conduit aux meilleures performances pratiques avec des temps d'exécution plus prévisibles. Contrairement à la méthode WBS, les méthodes ID et NOT peuvent être appliquées lorsque les données sont constantes par morceaux.

Cho et Fryzlewicz (2020) proposent une autre méthodologie pour détecter plusieurs ruptures dans la moyenne d'une série chronologique stationnaire par morceaux linéaire autocorrélée. Ils démontrent la consistance de leur méthode nommée WEM.gsc. La bonne performance de cette méthode est aussi démontrée par des études de simulations et ils l'appliquent aux données sur la qualité de l'air de Londres.

La méthode WEM.gsc combine la solution générale basée sur le principe "wild energy maximisation principle" (WEM) utilisée dans Fryzlewicz *et al.* (2014) et Fryzlewicz (2020) pour les noyaux iid basée sur l'utilisation de la statistique CUSUM, et la stratégie de sélection de modèle basée sur des critères d'informations appelée critère gappy Schwartz (voir Schwarz *et al.* (1978)). Le logiciel R implémentant WEM.gSC est disponible sur <https://github.com/haeran-cho/wem.gsc>.

Chapitre 3

Détection des ruptures et estimation de leurs localisations : utilisation d'un test du rapport de vraisemblance

3.1 Introduction

Les travaux présentés dans le chapitre précédent s'intéressent essentiellement à l'étude de ruptures abruptes. Leurs puissances locales ne sont en général pas étudiées, donc leur capacité à détecter les petites ruptures n'est pas évaluée. Or la détection de ces ruptures peut être très importante dans la modélisation du phénomène étudié.

L'objectif de ce chapitre est de détecter les petites ruptures dans la moyenne des modèles CHARN(p) suivants :

$$X_t = T(Z_{t-1}) + \gamma^\top \omega(t) + V(Z_{t-1})\varepsilon_t, \quad t \in \mathbb{Z}, \quad (3.1)$$

où $\gamma = (\gamma_1, \dots, \gamma_k, \gamma_{k+1})^\top \in \mathbb{R}^{k+1}$ et pour t_1, \dots, t_k , $1 < t_1 < \dots < t_k < n$, $\omega(t) = (\mathbb{1}_{[1, t_1[}(t), \mathbb{1}_{[t_1, t_2[}(t), \dots, \mathbb{1}_{[t_k, n[}(t))^\top \in \{0, 1\}^{k+1}$, $(X_t)_{t \in \mathbb{Z}}$ est un processus stationnaire par morceaux et ergodique, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc centré réduit de densité f , pour tout $t \in \mathbb{Z}$, $Z_t = (X_t, \dots, X_{t-p+1})^\top$, $p \in \mathbb{N}$, et T et V sont des fonctions réelles telles que $V(x) > 0, \forall x \in \mathbb{R}^p$. Cette classe de modèles est très vaste. Elle contient entre autres, les modèles AR, EXPAR, TARCH, ARCH, dont les propriétés statistiques et probabilistes sont abondamment étudiées dans la littérature (voir par exemple Guégan *et al.* (1994) et Tong (1990)).

Soit n le nombre des observations. Notre travail consiste donc à tester

$$H_0 : \gamma = \gamma_0 \text{ contre } H_\beta^{(n)} : \gamma = \gamma_0 + \frac{\beta}{\sqrt{n}} = \gamma_n, \quad n > 1,$$

pour $\gamma_0 \in \mathbb{R}^{k+1}$ et $\beta \in \mathbb{R}^{k+1}$ dépend des instants de rupture $t^k = (t_1, \dots, t_k)$.

Nous utilisons pour cela, le test du rapport de vraisemblance. Dans un premier temps, nous écrivons le logarithme du rapport de vraisemblance de H_0 contre $H_\beta^{(n)}$. Puis, nous divisons notre travail en deux parties. La première consiste à supposer que les fonctions T et V sont connues. Ici, nous montrons que notre statistique de test vérifie la propriété LAN énoncée dans Dreesbeke et Fine (1996). En utilisant le troisième lemme de Le Cam (Proposition 4.2 de Dreesbeke et Fine (1996)), nous calculons la puissance locale du test et nous montrons qu'il est optimal. L'optimalité ici signifie que notre test est le plus puissant que n'importe quel autre test de même niveau de signification.

Dans la deuxième partie, nous supposons que les fonctions T et V sont de formes connues mais dépendent de paramètres inconnus : $T(x) = m(x; \rho)$, $V(x) = \sigma(x; \theta)$, $\rho \in \Theta \subset \mathbb{R}^l$ et $\theta \in \tilde{\Theta} \subset \mathbb{R}^q$. Nous considérons alors $\psi_0 = (\rho_0^\top, \theta_0^\top)^\top \in \Theta \times \tilde{\Theta} \subset \mathbb{R}^l \times \mathbb{R}^q$ le vrai paramètre du modèle (3.1). Nous définissons un estimateur consistant ψ_n de ψ_0 , et nous montrons que sous certaines conditions, sous H_0 , la suite de vecteurs $\sqrt{n}(\psi_n - \psi_0)$ converge en loi vers un vecteur gaussien centré et sous $H_\beta^{(n)}$, elle converge en loi vers un vecteur gaussien non-centré. Le test construit dans ce cas reste optimal. Enfin, nous considérons la situation où γ_0 est inconnue et nous montrons que le test reste optimal.

3.1.1 Hypothèses générales

Soit pour tout $t = 1, \dots, n$, $G_t = \sigma(Z_1, \dots, Z_t)$, la σ -algèbre engendrée par Z_1, \dots, Z_t , telle que ε_t soit indépendant de G_{t-1} . Pour une fonction différentiable h de dérivé h' , on note

$$\phi_h = -\frac{h'}{h} \text{ et } I(h) = \int_{\mathbb{R}} \phi_h^2(x) h(x) dx.$$

On fait les hypothèses suivantes :

- (A') On suppose que pour tout $j = 1, \dots, k+1$, $n_j(n) = t_j - t_{j-1} \rightarrow \infty$ lorsque $n \rightarrow \infty$ et que la loi stationnaire sur chaque intervalle est de fonction de répartition F_j , où $j = 1, \dots, k+1$. On suppose aussi que pour tout $j = 1, \dots, k+1$, $n_j(n)/n \rightarrow \alpha_j$ lorsque $n \rightarrow \infty$.
- (A₀) : $\int_{\mathbb{R}} x f(x) dx = 0$ et $\int_{\mathbb{R}} x^2 f(x) dx = 1$.
- (A₁) : f est dérivable avec dérivée f' .
- (A₂) : $\mu_{j,\ell}(\gamma_0) = I(f) \int_{\mathbb{R}^p} \left(\frac{1}{V(x)} \right)^\ell dF_j(x) < \infty$, où $j = 1, \dots, k+1$, $\ell \leq 3$.
- (A₃) : ϕ_f est dérivable de dérivée ϕ'_f , c_ϕ -lipschitzienne, où $0 < c_\phi < \infty$.
- (A₄) : $\lim_{x \rightarrow +\infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0$ et $\lim_{x \rightarrow +\infty} f'(x) = \lim_{x \rightarrow -\infty} f'(x) = 0$.
- (A₅) : $\int_{\mathbb{R}} |\phi_f(x)|^3 f(x) dx < \infty$.

Remarques 1. *i) Sur chaque intervalle $[t_{j-1}, t_j], j = 1, \dots, k+1$, il y a au plus p vecteurs aléatoires Z_t dont les composantes sont associées à des indices à la fois dans $[t_{j-1}, t_j]$ et $[t_{j-2}, t_{j-1})$. Ils peuvent ne pas avoir la même loi stationnaire. Mais comme leur nombre est négligeable devant $n_j(n)$, leurs lois n'affectent pas les résultats asymptotiques.*

ii) Pour tout $\ell \leq 3$, $\mu_{j,\ell}$ dépend de γ_0 car les F_j dépendent des composantes de γ_0 .

iii) L'hypothèse (A_4) implique $\int_{\mathbb{R}} \phi_f(x) f(x) dx = 0$ et $\int_{\mathbb{R}} \phi_f^2(x) f(x) dx = \int_{\mathbb{R}} \phi_f'(x) f(x) dx$.

Démonstration. iii) On a

$$\begin{aligned} \int_{\mathbb{R}} \phi_f(x) f(x) dx &= - \int_{\mathbb{R}} \frac{f'(x)}{f(x)} f(x) dx \\ &= - \int_{\mathbb{R}} f'(x) dx \\ &= 0 \end{aligned}$$

et

$$\begin{aligned} \int_{\mathbb{R}} \phi_f'(x) f(x) dx &= \int_{\mathbb{R}} - \left(\frac{f'(x)}{f(x)} \right)' f(x) dx \\ &= - \int_{\mathbb{R}} \frac{f''(x) f(x) - f'^2(x)}{f^2(x)} f(x) dx \\ &= - \int_{\mathbb{R}} \left[\frac{f''(x)}{f(x)} - \left(\frac{f'(x)}{f(x)} \right)^2 \right] f(x) dx \\ &= \int_{\mathbb{R}} f''(x) dx + \int_{\mathbb{R}} \phi_f^2(x) f(x) dx \\ &= \int_{\mathbb{R}} \phi_f^2(x) f(x) dx. \end{aligned}$$

□

3.1.2 Détection des ruptures

De nombreuses stratégies de détection des ruptures et d'estimation de leurs localisations dans les données hors-ligne ont été proposées dans la littérature. Les stratégies que nous proposons ici reposent sur la puissance du test étudié. Avant de les présenter, nous commençons par décrire les formes générales des paramètres γ_0 et β . Nous rappelons que nous supposons qu'on a un changement dans la moyenne du modèle CHARN donné par (3.1). C'est-à-dire nous supposons que nous avons une série d'observations générées à partir de (3.1) et certains indices de temps, la moyenne des observations change (voir les exemples de la Figure 5.1). Nous voulons détecter ces instants de changement.

Nous pouvons commencer, d'après (Brockwell *et al.*, 2002, Page 14), la modélisation de la

série des données, X_1, X_2, \dots, X_n par l'analyse de son graphe ou son chronogramme. Cela permet d'avoir quelques informations sur ces données, comme par exemple, les données historiques X_1, X_2, \dots, X_m , $m \ll n$ qui sont les premières m observations à peu près stationnaires, le nombre maximum K de points de ruptures dans la série, leurs localisations approximatives et la distance minimale $h \ll n$ entre elles.

Pour détecter les faibles ruptures, nous considérons les stratégies ci-dessous.

Pour $k \geq 0$, notons \mathcal{P}_{k,t^k} la puissance théorique du test H_0 contre $H_\beta^{(n)}$ en $t^k = (t_1, \dots, t_k)$. Par convention on a $\mathcal{P}_{0,t^0} = \alpha$, où α est le seuil de signification du test. On précise que dans le cas où aucune rupture dans les données, γ_0 contient la même valeur sur chaque composante. Soit $\widehat{\mathcal{P}}_{k,t^k}$ un estimateur de \mathcal{P}_{k,t^k} obtenu en remplaçant les paramètres par leurs estimateurs dans l'expression de \mathcal{P}_{k,t^k} . En particulier pour $j = 2, \dots, k+1$, nous pouvons estimer la j -ème composante de β par $\sqrt{n}(\overline{X}_{j,n} - \widehat{\gamma}_{j,n})$ où $\overline{X}_{j,n}$ est la moyenne des observations dans l'intervalle $[t_j, t_{j+1})$ et $\widehat{\gamma}_{j,n}$ la j -ème composante de l'estimateur de γ_0 . Par exemple, le test d'hypothèse de non-changement c'est-à-dire lorsque $\gamma_0 = (\mu_1, \mu_1, \dots, \mu_1)$ contre l'hypothèse qu'il existe au moins un point de rupture, la j -ème composante de $\widehat{\beta}$ peut être estimée par $\sqrt{n}(\overline{X}_{j,n} - \overline{X}_{1,n})$. Dans ce cas $\widehat{\gamma}_{j,n} = \overline{X}_{1,n}$.

Soit $\zeta \in (0, 1)$ un nombre petit. Nos procédures de détection des changements et d'estimation de leur localisations dans une série temporelle X_1, X_2, \dots, X_n commencent par trouver les informations a priori ci-dessus et se poursuivent comme suit.

- **S_1 : Détection des ruptures**

Nous prenons $k = 1$ et nous appliquons notre test pour tous les t_1 , où $m \leq t_1 \leq n$.

- i. Si $|\widehat{\mathcal{P}}_{1,t^1} - \mathcal{P}_{0,t^0}| \leq \zeta$ pour tout t_1 , alors aucun point de rupture détecté dans la série.
- ii. Si $|\widehat{\mathcal{P}}_{1,t^1} - \mathcal{P}_{0,t^0}| > \zeta$ pour un t_1 , alors il existe au moins une rupture dans la série.

- **S_2 : Estimation des localisations des ruptures**

Soit $m < \tau_1^0 < \dots < \tau_k^0 \leq n - h$, $\tau_j^0 - \tau_{j-1}^0 \geq h$, $j = 2, \dots, K$, les localisations des potentiels points de rupture obtenus à partir du chronogramme. Soit C_j un ensemble arbitraire des indices du temps autour des τ_j^0 , et $\Omega_j \cap \Omega_\ell = \emptyset$, $j \neq \ell = 1, \dots, K$. Nous considérons $S_\ell = C_1 \times C_2 \times \dots \times C_\ell$. Pour tout ℓ -uplet $\tau^\ell = (\tau_1, \dots, \tau_\ell) \in S_\ell$, nous appliquons le problème de test ci-dessus pour $t_j = \tau_j$, $j = 1, \dots, \ell$ et nous calculons $\widehat{\mathcal{P}}_{\ell,t^\ell}$. Nous prenons $\widehat{\mathcal{P}}_{0,t^0} = \mathcal{P}_{0,t^0} = \alpha$ et nous supposons que C_1 contient la localisation de la première rupture.

- Pour $\ell = 2, \dots, K$,

- i- Si $|\max_{t^\ell \in S_\ell} \widehat{\mathcal{P}}_{\ell, t^\ell} - \max_{t^{\ell-1} \in S_{\ell-1}} \widehat{\mathcal{P}}_{\ell-1, t^{\ell-1}}| \leq \zeta$, il n'y a probablement pas de localisation d'une rupture dans C_ℓ . Nous prenons $\widetilde{S}_\ell = S_{\ell-1}$.
- ii- Si $|\max_{t^\ell \in S_\ell} \widehat{\mathcal{P}}_{\ell, t^\ell} - \max_{t^{\ell-1} \in S_{\ell-1}} \widehat{\mathcal{P}}_{\ell-1, t^{\ell-1}}| > \zeta$, il y a probablement de localisation d'une rupture dans C_ℓ . Nous prenons $\widetilde{S}_\ell = S_\ell$.

Notons ι la longueur des vecteurs dans \widetilde{S}_K . Alors l'estimateur $(\widehat{k}, \widehat{t}^k)$ du couple du nombre de ruptures et du vecteur des localisations peut être obtenu comme

$$(\widehat{k}, \widehat{t}^k) = \arg \max_{t^\iota \in \widetilde{S}_K} \widehat{\mathcal{P}}_{\iota, t^\iota}.$$

– Clairement, si k , le nombre de localisations est connu, t^k peut être estimé par

$$\widehat{t}^k = \arg \max_{t^k \in S_k} \widehat{\mathcal{P}}_{k, t^k}. \quad (\star\star)$$

• **S3 : Une alternative à S2**

À la place de S_2 , nous pouvons utiliser la stratégie séquentielle suivante. Nous considérons le segment $\mathcal{U}_1 = \{1, \dots, m, \dots, \tau_1^0 + h\}$, $\mathcal{U}_\ell = \{\tau_\ell^0 + h, \dots, \tau_{\ell+1}^0 + h\}$, $\ell = 2, \dots, K$.

- i. Étape 1 : Nous appliquons notre test pour $t_1 \in \mathcal{U}_1$, $t > m$ avec n remplacé par $\tau_\ell^0 + h$.
Si $|\widehat{\mathcal{P}}_{1, t_1} - \mathcal{P}_{0, t_0}| \leq \zeta$ pour tous les t_1 considérés, alors, il n'y a pas de localisation de changement dans \mathcal{U}_1 . Nous mettons à jour en \mathcal{U}_2 en fusionnant le \mathcal{U}_2 courant avec \mathcal{U}_1 .
Si $|\widehat{\mathcal{P}}_{1, t_1} - \mathcal{P}_{0, t_0}| > \zeta$ pour un t_1 considéré, alors, il y a une localisation de changement dans \mathcal{U}_1 estimée par

$$\widehat{t}_1^1 = \arg \max_{t^1 \in \mathcal{U}_1} \widehat{\mathcal{P}}_{1, t^1}.$$

Ensuite, nous mettons à jour \mathcal{U}_2 en remplaçant \widehat{t}_1^1 par τ_1^0 dans le \mathcal{U}_2 courant.

- ii. Étape ℓ : Nous appliquons notre test pour tous $t_1 \in \mathcal{U}_\ell$, $t_1 > \tau_\ell^0 + h$ ou $t_1 > \widehat{t}_\ell^1 + h$, avec n remplacé par la longueur de \mathcal{U}_k .
Si $|\widehat{\mathcal{P}}_{k, t^k} - \mathcal{P}_{0, t^0}| \leq \zeta$ pour tout les t_1 considérés, alors, il n'y a pas de localisation de changement dans \mathcal{U}_ℓ . Nous mettons à jour $\mathcal{U}_{\ell+1}$ en fusionnant le $\mathcal{U}_{\ell+1}$ courant avec \mathcal{U}_ℓ .
Si $|\widehat{\mathcal{P}}_{\ell, t^\ell} - \mathcal{P}_{0, t^0}| > \zeta$ pour un t_1 considéré, alors, il y a une localisation de changement dans \mathcal{U}_ℓ estimée par

$$\widehat{t}_\ell^1 = \arg \max_{t^1 \in \mathcal{U}_\ell} \widehat{\mathcal{P}}_{\ell, t^\ell}.$$

Ensuite, nous mettons à jour $\mathcal{U}_{\ell+1}$ en remplaçant τ_ℓ^0 par \hat{t}_ℓ^ℓ dans le $\mathcal{U}_{\ell+1}$ courant. Il est clair que le nombre total des changements dans la série est le nombre des \mathcal{U}_ℓ où des changements ont été détectés, et leurs localisations sont estimées par les \hat{t}_ℓ^1 .

La partie i de $S1$ vient de l'idée qu'en cas d'absence de rupture, toutes les estimations possibles seraient proches de 0, et la puissance estimée serait proche du niveau du test. La partie ii est motivée par le fait que s'il y a au moins un changement, alors on peut en trouver un pour lequel les moyennes des observations avant et après cet indice de temps sont différentes. Le test rejeterait alors l'hypothèse nulle d'absence de rupture. $S2$ et $S3$ sont construits à partir des mêmes idées.

Remarques 2. *i. En pratique, les ensembles C_j dans la stratégie $S2$ ne doivent pas être trop grands pour des K grands.*

ii. Dans la stratégie $S1$, pour les t_1 se rapprochant de n , les données peuvent être épissées comme dans la méthode dite de segmentation binaire circulaire décrite dans Olshen et al. (2004). Ensuite, dans le cas d'un seul changement, pour estimer sa localisation, on peut appliquer directement $(\star\star)$ avec $C_1 = \{m + 1, \dots, n\}$ ou $S3$ avec $\mathcal{U}_1 = \{m + 1, \dots, n\}$, épissant les données lorsque l'indice de temps courant se rapproche de n .

iii. L'utilisation du graphique ne signifie pas que le nombre de ruptures et leurs localisations sont connus. Il a pour but d'éviter une recherche exhaustive trop coûteuse en temps, et de concentrer la recherche dans les "zones" où il est plus probable de trouver les ruptures. Malheureusement lorsque les ruptures sont de trop petites amplitudes, le graphique ne renseigne pas toujours sur ces zones. Par contre, comme on le verra dans les simulations, nos méthodes permettent de mieux détecter les petites ruptures qu'on peut voir sur les graphiques par rapport aux méthodes auxquelles elles sont comparées.

iv. L'utilisation de nos méthodes peut être précédée par l'utilisation d'autres méthodes. Nos méthodes permettront ainsi de valider les ruptures obtenues et/ou d'estimer celles qui n'auraient pas été détectées, comme par exemple, les petites ruptures.

Nous commençons dans la section suivante par étudier le cas où les fonctions T , V et la densité du bruit sont connues.

3.2 Cas où les fonctions T , V et la densité du bruit sont connues

Pour $\gamma = (\gamma_1, \dots, \gamma_{k+1})^\top \in \mathbb{R}^{k+1}$ et pour $t \in \mathbb{Z}$, on définit

$$\varepsilon_t(\gamma) = \frac{X_t - T(Z_{t-1}) - \gamma^\top \omega(t)}{V(Z_{t-1})}.$$

On veut tester

$$H_0 : \gamma = \gamma_0 \text{ contre } H_\beta^{(n)} : \gamma = \gamma_0 + \frac{\beta}{\sqrt{n}} = \gamma_n, \quad n > 1.$$

Pour cela, on utilise un test du rapport de vraisemblance. Le log du rapport de vraisemblance de H_0 contre $H_\beta^{(n)}$ est le suivant :

$$\Lambda_n = \sum_{t=1}^n \{\log[f(\varepsilon_t(\gamma_n))] - \log[f(\varepsilon_t(\gamma_0))]\}.$$

En effet, soit $\mathbf{X}^{(n)} = (X_n, \dots, X_1)$. On commence par calculer la loi de $\mathbf{X}^{(n)}$ sachant $Z_0 = (X_0, \dots, X_{1-p})^\top$, $p \in \mathbb{N}$. Par une application successive de la formule de Bayes, on a

$$\begin{aligned} & \mathfrak{L}_{X_n, X_{n-1}, \dots, X_1 | Z_0 = z_0}(x_n, \dots, x_1) \\ &= \mathfrak{L}_{X_n | (X_{n-1}, \dots, X_1, Z_0) = (x_{n-1}, \dots, x_1, z_0)}(x_n) \mathfrak{L}_{X_{n-1} | (X_{n-2}, \dots, X_1, Z_0) = (x_{n-2}, \dots, x_1, z_0)}(x_{n-1}) \\ &\times \mathfrak{L}_{X_{n-2} | (X_{n-3}, \dots, X_1, Z_0) = (x_{n-3}, \dots, x_1, z_0)}(x_{n-2}) \dots \mathfrak{L}_{X_1 | Z_0 = z_0}(x_1) \\ &= \mathfrak{L}_{X_n | X_{n-1} = x_{n-1}}(x_n) \mathfrak{L}_{X_{n-1} | X_{n-2} = x_{n-2}}(x_{n-1}) \dots \mathfrak{L}_{X_1 | Z_0 = z_0}(x_1), \end{aligned}$$

où $\mathfrak{L}_{X|Y}$ désigne la loi de X sachant Y .

Pour calculer par exemple $\mathfrak{L}_{X_1 | Z_0}$, on calcule d'abord la fonction de répartition de X_1 sachant Z_0 . On a pour tout $x \in \mathbb{R}$,

$$\begin{aligned} & \mathbb{P}(X_1 < x | Z_0 = z_0) \\ &= \mathbb{P}(T(Z_0) + V(Z_0)\varepsilon_0 < x | Z_0 = z_0) \\ &= \mathbb{P}\left(\varepsilon < \frac{x - T(z_0)}{V(z_0)}\right). \end{aligned}$$

En dérivant le membre de droite par rapport à x , on obtient

$$\mathfrak{L}_{X_1 | Z_0 = z_0}(x) = \frac{1}{V(z_0)} f\left(\frac{x - T(z_0)}{V(z_0)}\right), \quad x \in \mathbb{R}.$$

Ainsi, la vraisemblance de $\mathbf{X}^{(n)}$ sachant Z_0 a l'expression

$$\mathcal{L}(\mathbf{X}^{(n)} | Z_0) = \prod_{t=1}^n \frac{1}{V(Z_{t-1})} f[\varepsilon_t(\gamma)], \quad \gamma \in \mathbb{R}^{k+1}.$$

Le log-vraisemblance conditionnelle s'écrit donc :

$$L(\mathbf{X}^{(n)}) = \sum_{t=1}^n \{-\log[V(Z_{t-1})] + \log[f(\varepsilon_t(\gamma))]\}.$$

D'après ce qui précède, le logarithme du rapport de vraisemblance de H_0 contre $H_\beta^{(n)}$ s'écrit :

$$\begin{aligned} \Lambda_n &= \sum_{t=1}^n \log \left\{ \frac{f[\varepsilon_t(\gamma_n)]}{f[\varepsilon_t(\gamma_0)]} \right\} \\ &= \sum_{t=1}^n \{\log[f(\varepsilon_t(\gamma_n))] - \log[f(\varepsilon_t(\gamma_0))]\}. \end{aligned}$$

On définit pour tout $\beta \in \mathbb{R}^k$,

$$\Delta_n(\gamma_0, \beta) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)].$$

Théorème 3.2.1. *Sous les hypothèses (A'), (A₀)-(A₅), sous H_0 , on a*

$$\Lambda_n = \Delta_n(\gamma_0, \beta) - \frac{\mu(\gamma_0, \beta)}{2} + o_P(1),$$

où sous H_0 ,

$$\Delta_n(\gamma_0, \beta) \xrightarrow{loi} \mathcal{N}(0, \mu(\gamma_0, \beta)), n \rightarrow \infty,$$

avec $\mu(\gamma_0, \beta) = \sum_{j=1}^{k+1} \alpha_j \beta_j^2 \mu_{j,2}(\gamma_0) < \infty$.

Démonstration. On commence par montrer que

$$\Lambda_n = \Lambda_{1n} - \Lambda_{2n} + o_P(1),$$

où

$$\Lambda_{1n} := \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)]$$

et

$$\Lambda_{2n} := \frac{1}{2n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi'_f[\varepsilon_t(\gamma_0)].$$

En effet, on a

$$\Lambda_n = \sum_{t=1}^n \{\log[f(\varepsilon_t(\gamma_n))] - \log[f(\varepsilon_t(\gamma_0))]\}.$$

On applique la formule de Taylor d'ordre 2 au point γ_0 . Il existe $\tilde{\gamma}$ entre γ_n et γ_0 tel que

$$\Lambda_n = \sum_{t=1}^n \left. \partial_\gamma [\log f(\varepsilon_t(\gamma))] \right|_{\gamma=\gamma_0} (\gamma_n - \gamma_0) + \frac{1}{2} (\gamma_n - \gamma_0)^\top \left. \partial_\gamma^2 [\log f(\varepsilon_t(\gamma))] \right|_{\gamma=\tilde{\gamma}} (\gamma_n - \gamma_0).$$

Or, pour tout $\gamma \in \mathbb{R}^{k+1}$, on a

$$\begin{aligned} \partial_\gamma \{\log f[\varepsilon_t(\gamma)]\} &= \frac{1}{f(\varepsilon_t(\gamma))} \partial_\gamma [f(\varepsilon_t(\gamma))] \\ &= \frac{f'[\varepsilon_t(\gamma)]}{f[\varepsilon_t(\gamma)]} \partial_\gamma [\varepsilon_t(\gamma)] \\ &= \frac{f'[\varepsilon_t(\gamma)]}{f[\varepsilon_t(\gamma)]} \partial_\gamma \left[\frac{X_t - T(Z_{t-1}) - \gamma^\top \omega(t)}{V(Z_{t-1})} \right] \\ &= \frac{\omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma)]. \end{aligned}$$

Ainsi,

$$\begin{aligned} \partial_\gamma^2 \{\log f[\varepsilon_t(\gamma)]\} &= \frac{\omega(t)}{V(Z_{t-1})} \partial_\gamma \phi_f[\varepsilon_t(\gamma)] \\ &= \frac{\omega(t)}{V(Z_{t-1})} \phi_f'[\varepsilon_t(\gamma)] \partial_\gamma [\varepsilon_t(\gamma)] \\ &= -\frac{\omega(t)\omega(t)^\top}{V^2(Z_{t-1})} \phi_f'[\varepsilon_t(\gamma)]. \end{aligned}$$

En remplaçant $\partial_\gamma [\log f(\varepsilon_t(\gamma_0))] \Big|_{\gamma=\gamma_0}$ et $\partial_\gamma^2 [\log f(\varepsilon_t(\gamma))] \Big|_{\gamma=\tilde{\gamma}}$ par leurs expressions respectives, on obtient

$$\begin{aligned} \Lambda_n &= (\gamma_n - \gamma_0)^\top \sum_{t=1}^n \frac{\omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)] - \frac{1}{2} (\gamma_n - \gamma_0)^\top \sum_{t=1}^n \frac{\omega(t)\omega(t)^\top}{V^2(Z_{t-1})} \phi_f'[\varepsilon_t(\tilde{\gamma})] (\gamma_n - \gamma_0) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)] - \frac{1}{2n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi_f'[\varepsilon_t(\tilde{\gamma})] \\ &= \Lambda_{1n} - \Lambda_{1n}^{(1)}. \end{aligned}$$

Or, pour tout $t = 1, \dots, n$, et pour tout $\gamma \in \mathbb{R}^{k+1}$,

$$\begin{aligned} \varepsilon_t(\gamma) &= \frac{X_t - T(Z_{t-1}) - \gamma^\top \omega(t)}{V(Z_{t-1})} \\ &= \varepsilon_t(\gamma_0) + \frac{\gamma_0^\top \omega(t)}{V(Z_{t-1})} - \frac{\gamma^\top \omega(t)}{V(Z_{t-1})} \\ &= \varepsilon_t(\gamma_0) + (\gamma_0 - \gamma)^\top \frac{\omega(t)}{V(Z_{t-1})}. \end{aligned}$$

Par suite ,

$$\Lambda_{1n}^{(1)} := \frac{1}{2n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi_f' \left[\varepsilon_t(\gamma_0) + (\gamma_0 - \tilde{\gamma})^\top \frac{\omega(t)}{V(Z_{t-1})} \right].$$

D'après (A₃), on a

$$\left| \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi_f' \left[\varepsilon_t(\gamma_0) + (\gamma_0 - \tilde{\gamma})^\top \frac{\omega(t)}{V(Z_{t-1})} \right] - \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi_f' [\varepsilon_t(\gamma_0)] \right|$$

$$\leq \frac{1}{n} C_\phi \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \left| (\gamma_0 - \tilde{\gamma})^\top \frac{\omega(t)}{V(Z_{t-1})} \right|.$$

Il en découle que,

$$\begin{aligned} & \left| \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi'_f \left[\varepsilon_t(\gamma_0) + (\gamma_0 - \tilde{\gamma})^\top \frac{\omega(t)}{V(Z_{t-1})} \right] - \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi'_f [\varepsilon_t(\gamma_0)] \right| \\ & \leq \frac{1}{n} C_\phi \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2 \|\omega(t)\|}{V^3(Z_{t-1})} \|\gamma_0 - \tilde{\gamma}\| \\ & \leq (k+1) C_\phi \|\gamma_0 - \tilde{\gamma}\| \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^3(Z_{t-1})}. \end{aligned}$$

On a,

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^3(Z_{t-1})} &= \frac{1}{n} \sum_{j=1}^{k+1} \sum_{t=t_{j-1}}^{t_j} \frac{\beta_j^2}{V^3(Z_{t-1})} \\ &= \sum_{j=1}^{k+1} \frac{1}{n} \sum_{t=t_{j-1}}^{t_j} \frac{\beta_j^2}{V^3(Z_{t-1})} \\ &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{\beta_j^2}{n_j(n)} \sum_{i=t_{j-1}}^{t_j} \frac{1}{V^3(Z_{t-1})}. \end{aligned}$$

D'après le théorème ergodique, pour tout $j = 1, \dots, k+1$, presque sûrement,

$$\lim_{n \rightarrow \infty} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^3(Z_{t-1})} = \frac{\mu_{j,3}(\gamma_0)}{I(f)} < \infty.$$

Ainsi, presque sûrement,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^3(Z_{t-1})} = \frac{1}{I(f)} \sum_{j=1}^{k+1} \lim_{n \rightarrow \infty} \frac{n_j(n)}{n} \beta_j^2 \mu_{j,3}(\gamma_0) = \frac{1}{I(f)} \sum_{j=1}^{k+1} \alpha_j \beta_j^2 \mu_{j,3}(\gamma_0) < \infty.$$

Puisque $\tilde{\gamma}$ est entre γ_0 et γ_n et que $\gamma_n - \gamma_0 = \beta/\sqrt{n}$, on a alors que $\|\gamma_0 - \tilde{\gamma}\| \leq \|\beta\|/\sqrt{n}$ qui tend vers zéro lorsque n tend vers l'infini. Il s'ensuit que, presque sûrement,

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^3(Z_{t-1})} \|\gamma_0 - \tilde{\gamma}\| = 0.$$

D'après ce qui précède, on a

$$\Lambda_n = \Lambda_{1n} - \Lambda_{2n} + o_P(1),$$

où

$$\Lambda_{1n} := \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)]$$

et

$$\Lambda_{2n} := \frac{1}{2n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi'_f[\varepsilon_t(\gamma_0)].$$

Montrons à présent que sous H_0 ,

$$\Lambda_{1n} \xrightarrow{loi} \mathcal{N}(0, \mu(\gamma_0, \beta)).$$

Pour cela, on procède comme dans Ngatchou-Wandji (2005). On considère la suite :

$$\Lambda_{n,j} = \frac{1}{\sqrt{n}} \sum_{t=1}^j \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)], \quad j = 1, \dots, n,$$

et on définit pour tout $t = 1, \dots, n$,

$$\xi_{n,t} = \frac{1}{\sqrt{n}} \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)].$$

On observe alors que

$$\Lambda_{n,j} = \sum_{t=1}^j \xi_{n,t}, \quad j = 1, \dots, n,$$

et il est facile de voir que, pour tout $j = 1, \dots, n$, $\Lambda_{n,j}$ est centré.

Nous montrons d'abord que $\{(\Lambda_{n,j}, G_j), j = 1, \dots, n\}$ est une suite de martingales où nous rappelons que $G_{t-1} = \sigma(Z_1, \dots, Z_t)$ est la σ -algèbre engendrée par Z_1, \dots, Z_t , $t \in \mathbb{Z}$.

En effet, soient $j_1, j_2 \in \mathbb{Z}$ tels que $j_1 < j_2$ on a :

$$\begin{aligned} \mathbb{E} \left(\Lambda_{n,j_2} \middle| G_{j_1} \right) &= \mathbb{E} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^{j_2} \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)] \middle| G_{j_1} \right\} \\ &= \mathbb{E} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^{j_1} \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)] \middle| G_{j_1} \right\} \\ &\quad + \mathbb{E} \left\{ \frac{1}{\sqrt{n}} \sum_{t=j_1+1}^{j_2} \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)] \middle| G_{j_1} \right\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^{j_1} \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \phi_f[\varepsilon_t(\gamma_0)] + \frac{1}{\sqrt{n}} \sum_{t=j_1+1}^{j_2} \mathbb{E} \left\{ \frac{\beta^\top \omega(t)}{V(Z_{t-1})} \middle| G_{j_1} \right\} \mathbb{E} \{ \phi_f[\varepsilon_t(\gamma_0)] \}. \end{aligned}$$

D'après la remarque 1, on a $\mathbb{E}[\phi_f(\varepsilon_t(\gamma_0))] = 0$. Donc

$$\mathbb{E} \left(\Lambda_{n,j_2} \middle| G_{j_1} \right) = \Lambda_{n,j_1}.$$

On a aussi

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(\xi_{n,t}^2 | G_{t-1}) &= \sum_{t=1}^n \mathbb{E} \left\{ \frac{1}{n} \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi_f^2[\varepsilon_t(\gamma_0)] \middle| G_{t-1} \right\} \\ &= \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\gamma_0)] \middle| G_{t-1} \right\}. \end{aligned}$$

Puisque pour tout $t = 1, \dots, n$, ε_t est indépendant à G_{t-1} , on a

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(\xi_{n,t}^2 | G_{t-1}) &= \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \mathbb{E}[\phi_f^2(\varepsilon_t(\gamma_0))] \\ &= \frac{1}{n} \sum_{j=1}^{k+1} \sum_{t=t_{j-1}}^{t_j} \frac{\beta_j^2}{V^2(Z_{t-1})} \mathbb{E}[\phi_f^2(\varepsilon_t(\gamma_0))] \\ &= \sum_{j=1}^{k+1} \frac{1}{n} \sum_{t=t_{j-1}}^{t_j} \frac{\beta_j^2}{V^2(Z_{t-1})} \mathbb{E}[\phi_f^2(\varepsilon_t(\gamma_0))] \\ &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{\beta_j^2}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^2(Z_{t-1})} \mathbb{E}[\phi_f^2(\varepsilon_t(\gamma_0))]. \end{aligned}$$

D'après nos hypothèses, du fait que les ε_t , $t \in \mathbb{Z}$, sont iid, on a

$$\mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\gamma_0)] \right\} = \int_{\mathbb{R}} \phi_f^2(x) f(x) dx = I(f).$$

D'après le théorème ergodique, pour tout $j = 1, \dots, k+1$, presque sûrement,

$$\lim_{n \rightarrow \infty} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^2(Z_{t-1})} \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\gamma_0)] \right\} = \mu_{j,2}(\gamma_0) < \infty.$$

Ainsi,

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n \mathbb{E}(\xi_{n,t}^2 | G_{t-1}) = \mu(\gamma_0, \beta) = \sum_{j=1}^{k+1} \alpha_j \beta_j^2 \mu_{j,2}(\gamma_0) < \infty.$$

Maintenant on montre la condition de Lindeberg :

Soit $\epsilon > 0$, à l'aide de l'inégalité de Hölder conditionnelle, on a

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(\xi_{n,t}^2 \mathbf{1}_{|\xi_{n,t}| > \epsilon} | G_{t-1}) &\leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}}(|\xi_{n,t}|^3 | G_{t-1}) \mathbb{E}^{\frac{1}{3}}(\mathbf{1}_{|\xi_{n,t}| > \epsilon} | G_{t-1}) \\ &= \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ n^{-\frac{3}{2}} \frac{|\beta^\top \omega(t)|^3}{V^3(Z_{t-1})} |\phi_f^3[\varepsilon_t(\gamma_0)]| \middle| G_{t-1} \right\} \mathbb{P}^{\frac{1}{3}}(|\xi_{n,t}| > \epsilon | G_{t-1}). \end{aligned}$$

D'après l'inégalité de Markov,

$$\begin{aligned}
& \sum_{t=1}^n \mathbb{E} (\xi_{n,t}^2 \mathbf{1}_{|\xi_{n,t}| > \epsilon} | G_{t-1}) \\
& \leq n^{-1} \sum_{t=1}^n \left\{ [\beta^\top \omega(t)]^2 \frac{1}{V^2(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 | G_{t-1} \right\} \frac{\mathbb{E}^{\frac{1}{3}} [|\xi_{n,t}|^3 | G_{t-1}]}{\epsilon^{\frac{1}{3}}} \right\} \\
& \leq Cst \frac{1}{n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \mathbb{E}^{\frac{2}{3}} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 | G_{t-1} \right\} \times \mathbb{E}^{\frac{1}{3}} \left\{ n^{-\frac{3}{2}} \frac{|\beta^\top \omega(t)|^3}{V^3(Z_{t-1})} \|\phi_f[\varepsilon_t(\gamma_0)]\|^3 | G_{t-1} \right\} \\
& = Cst \sum_{t=1}^n n^{-\frac{3}{2}} \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \mathbb{E}^{\frac{2}{3}} \left\{ \phi_f^3[\varepsilon_t(\gamma_0)] | G_{t-1} \right\} \frac{|\beta^\top \omega(t)|}{V(Z_{t-1})} \times \mathbb{E}^{\frac{1}{3}} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 | G_{t-1} \right\} \\
& \leq Cst \sum_{t=1}^n n^{-\frac{3}{2}} \frac{|\beta^\top \omega(t)|^3}{V^3(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 | G_{t-1} \right\} \\
& = Cst \sum_{t=1}^n n^{-\frac{3}{2}} \frac{|\beta^\top \omega(t)|^3}{V^3(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 \right\} \\
& = Cst \sum_{j=1}^{k+1} n^{-\frac{3}{2}} \sum_{t=t_{j-1}}^{t_j} \frac{|\beta_j|^3}{V^3(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 \right\} \\
& = Cst \sum_{j=1}^{k+1} \left(\frac{n_j(n)}{n} \right)^{\frac{3}{2}} \frac{|\beta_j|^3}{[n_j(n)]^{\frac{3}{2}}} \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^3(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 \right\}
\end{aligned}$$

D'après le théorème ergodique, pour tout $\epsilon > 0$, pour tout $j = 1, \dots, k+1$, on a presque sûrement,

$$\begin{aligned}
& \lim_{n \rightarrow \infty} n_j^{-\frac{3}{2}}(n) \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^3(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 \right\} \\
& = \lim_{n \rightarrow \infty} n_j^{-\frac{1}{2}}(n) n_j^{-1}(n) \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^3(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 \right\} \\
& = \lim_{n \rightarrow \infty} n_j^{-\frac{1}{2}}(n) \times \lim_{n \rightarrow \infty} n_j^{-1}(n) \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^3(Z_{t-1})} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\gamma_0)]|^3 \right\} \\
& = 0 \times \int_{\mathbb{R}^p} \left(\frac{1}{V(x)} \right)^3 dF_j(x) \times \int_{\mathbb{R}} |\phi_f^3(x)| f(x) dx = 0.
\end{aligned}$$

Ainsi,

$$\lim_{n \rightarrow \infty} \sum_{t=1}^n \mathbb{E} (\xi_{n,t}^2 \mathbf{1}_{|\xi_{n,t}| > \epsilon} | G_{t-1}) = 0 \quad p.s.$$

Les conditions du Corollaire 3.1 de Hall et Heyde (1980) sont ainsi vérifiées. Donc d'après ce corollaire, sous H_0 ,

$$\Lambda_{1n} \xrightarrow{loi} \mathcal{N}(0, \mu(\gamma_0, \beta)).$$

Étudions maintenant la convergence de Λ_{2n} . On a

$$\begin{aligned}\Lambda_{2n} &= \frac{1}{2n} \sum_{t=1}^n \frac{[\beta^\top \omega(t)]^2}{V^2(Z_{t-1})} \phi'_f[\varepsilon_t(\gamma_0)] \\ &= \frac{1}{2} \sum_{j=1}^{k+1} \frac{1}{n} \sum_{t=t_{j-1}}^{t_j} \frac{\beta_j^2}{V^2(Z_{t-1})} \phi'_f[\varepsilon_t(\gamma_0)] \\ &= \frac{1}{2} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{\beta_j^2}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^2(Z_{t-1})} \phi'_f[\varepsilon_t(\gamma_0)].\end{aligned}$$

Sous nos hypothèses, $\mathbb{E}\{\phi'_f[\varepsilon_1(\gamma_0)]\} = \int_{\mathbb{R}} \phi'_f(x) f(x) dx = I(f)$. D'après le théorème ergodique, pour tout $j = 1, \dots, n$, presque sûrement, on a

$$\lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^2(Z_{t-1})} \phi'_f[\varepsilon_t(\gamma_0)] = \mathbb{E}\{\phi'_f[\varepsilon_1(\gamma_0)]\} \int_{\mathbb{R}^p} \left(\frac{1}{V(x)}\right)^2 dF_j(x) = \mu_{j,2}(\gamma_0).$$

D'où, presque sûrement,

$$\lim_{n \rightarrow +\infty} \Lambda_{2n} = \frac{1}{2} \sum_{j=1}^{k+1} \alpha_j \beta_j^2 \mu_{j,2}(\gamma_0) = \frac{\mu(\gamma_0, \beta)}{2}.$$

On en conclut que

$$\Lambda_n = \Lambda_{1n} - \frac{\mu(\gamma_0, \beta)}{2} + o_P(1).$$

Puisque sous H_0 , Λ_{1n} converge en loi vers une $\mathcal{N}(0, \mu(\gamma_0, \beta))$, on a que sous H_0 , lorsque $n \rightarrow \infty$,

$$\Lambda_n \xrightarrow{\text{loi}} \mathcal{N}\left(-\frac{\mu(\gamma_0, \beta)}{2}, \mu(\gamma_0, \beta)\right).$$

Ainsi, la propriété LAN énoncée dans Dreesbeke et Fine (1996) est vérifiée pour suite centrale $\Delta_n(\gamma_0, \beta) := \Lambda_{1n}$. \square

Corollaire 3.2.1. *Supposons que les hypothèses (A') et (A₀)-(A₅) sont vérifiées. Alors les suites d'hypothèses $\{H_\beta^{(n)} : n \geq 1\}$ et $\{H_0^{(n)} = H_0 : n \geq 1\}$ sont contigües.*

Démonstration. Dans la preuve du théorème précédent, on a montré que sous H_0 , lorsque $n \rightarrow \infty$,

$$\Delta_n(\gamma_0, \beta) \xrightarrow{\text{loi}} \mathcal{N}(0, \mu(\gamma_0, \beta)).$$

Et d'après le théorème, sous H_0 , lorsque $n \rightarrow \infty$,

$$\Lambda_n \xrightarrow{\text{loi}} \mathcal{N}\left(-\frac{\mu(\gamma_0, \beta)}{2}, \mu(\gamma_0, \beta)\right).$$

Ainsi, on a facilement que sous H_0 , lorsque $n \rightarrow \infty$,

$$\begin{pmatrix} \Delta_n(\gamma_0, \beta) \\ \Lambda_n \end{pmatrix} \xrightarrow{\text{loi}} \mathcal{N} \left(\begin{pmatrix} 0 \\ -\frac{\mu(\gamma_0, \beta)}{2} \end{pmatrix}, \begin{pmatrix} \mu(\gamma_0, \beta) & \mu(\gamma_0, \beta) \\ \mu(\gamma_0, \beta) & \mu(\gamma_0, \beta) \end{pmatrix} \right).$$

D'après la Proposition 4.2 de Dreesbeke et Fine (1996) (troisième lemme de Le Cam), les suites d'hypothèses $\{H_\beta^{(n)} : n \geq 1\}$ et $\{H_0^{(n)} = H_0 : n \geq 1\}$ sont contigües et sous $H_\beta^{(n)}$,

$$\Delta_n(\gamma_0, \beta) \xrightarrow{\text{loi}} \mathcal{N}(\mu(\gamma_0, \beta), \mu(\gamma_0, \beta)).$$

□

Pour tester H_0 contre $H_\beta^{(n)}$, on considère la statistique du test suivante :

$$\mathcal{T}_n(\gamma_0, \beta) = \frac{\Delta_n(\gamma_0, \beta)}{\widehat{\varpi}_n(\gamma_0, \beta)},$$

où $\widehat{\varpi}_n(\gamma_0, \beta) = \widehat{\mu}_n^{\frac{1}{2}}(\gamma_0, \beta)$, avec

$$\widehat{\mu}_n(\gamma_0, \beta) = \sum_{j=1}^{k+1} \widehat{\alpha}_j \beta_j^2 \widehat{\mu}_{j,2}(\gamma_0), \quad \widehat{\mu}_{j,2}(\gamma_0) = I(f) \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{V^2(Z_{t-1})},$$

et $\widehat{\alpha}_j$ est un estimateur de $\alpha_j = \lim_{n \rightarrow \infty} n_j(n)/n$, $j = 1, \dots, k+1$. En pratique, on prend $\widehat{\alpha}_j = n_j(n)/n$. Ici, $\widehat{\mu}_{j,2}(\gamma_0)$ est un estimateur de $\mu_{j,2}(\gamma_0)$. On obtient le résultat suivant :

Théorème 3.2.2. *On suppose que les conditions (A') et (A₀)-(A₅) sont satisfaites. Alors,*

- i. *Sous $H_\beta^{(n)}$, $\mathcal{T}_n(\gamma_0, \beta) \xrightarrow{\text{loi}} \mathcal{N}(\varpi(\gamma_0, \beta), 1)$, lorsque $n \rightarrow \infty$.*
- ii. *Sous $H_\beta^{(n)}$, la puissance asymptotique locale du test basé sur $\mathcal{T}_n(\gamma_0, \beta)$ est égale à $1 - \Phi(u_\alpha - \varpi(\gamma_0, \beta))$ où Φ est la fonction de répartition de la loi normale centrée réduite et u_α le quantile d'ordre $1 - \alpha$, $\alpha \in (0, 1)$.*
- iii. *Le test basé sur $\mathcal{T}_n(\gamma_0, \beta)$ est localement asymptotiquement optimal.*

Démonstration. i) D'après le Théorème 3.2.1, sous H_0 , lorsque $n \rightarrow \infty$,

$$\begin{pmatrix} \Delta_n(\gamma_0, \beta) \\ \Lambda_n(\gamma_0, \beta) \end{pmatrix} \xrightarrow{\text{loi}} \mathcal{N} \left(\begin{pmatrix} 0 \\ -\frac{\mu(\gamma_0, \beta)}{2} \end{pmatrix}, \begin{pmatrix} \mu(\gamma_0, \beta) & \mu(\gamma_0, \beta) \\ \mu(\gamma_0, \beta) & \mu(\gamma_0, \beta) \end{pmatrix} \right).$$

D'après la Proposition 4.2 de Dreesbeke et Fine (1996), sous $H_\beta^{(n)}$, lorsque $n \rightarrow \infty$,

$$\Delta_n(\gamma_0, \beta) \xrightarrow{\text{loi}} \mathcal{N}(\mu(\gamma_0, \beta), \mu(\gamma_0, \beta)).$$

On observe que pour tout $n \geq 1$,

$$\frac{\Delta_n(\gamma_0, \beta)}{\widehat{\varpi}_n(\gamma_0, \beta)} = \frac{\Delta_n(\gamma_0, \beta)}{\varpi(\gamma_0, \beta)} \times \frac{\varpi(\gamma_0, \beta)}{\widehat{\varpi}_n(\gamma_0, \beta)}.$$

Puisque, sous H_0 , $\widehat{\varpi}_n(\gamma_0, \beta) \rightarrow \varpi(\gamma_0, \beta)$, lorsque $n \rightarrow \infty$, par contiguïté, le résultat est encore valable sous $H_\beta^{(n)}$. Et d'après le troisième lemme de Le Cam, sous $H_\beta^{(n)}$, lorsque $n \rightarrow \infty$,

$$\frac{\Delta_n(\gamma_0, \beta)}{\varpi(\gamma_0, \beta)} \xrightarrow{\text{loi}} \mathcal{N}(\varpi(\gamma_0, \beta), 1).$$

Donc, sous $H_\beta^{(n)}$, lorsque $n \rightarrow \infty$,

$$\frac{\Delta_n(\gamma_0, \beta)}{\widehat{\varpi}_n(\gamma_0, \beta)} \xrightarrow{\text{loi}} \mathcal{N}(\varpi(\gamma_0, \beta), 1).$$

ii) On obtient donc la fonction de répartition asymptotique de $\frac{\Delta_n(\gamma_0, \beta)}{\widehat{\varpi}_n(\gamma_0, \beta)}$:

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\Delta_n(\gamma_0, \beta)}{\widehat{\varpi}_n(\gamma_0, \beta)} > u_\alpha | H_\beta^{(n)} \right) = \lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{\Delta_n(\gamma_0, \beta)}{\varpi(\gamma_0, \beta)} > u_\alpha | H_\beta^{(n)} \right) = 1 - \Phi(u_\alpha - \varpi(\gamma_0, \beta)),$$

où Φ est la fonction de répartition de la loi Gaussienne centrée réduite et u_α sont quantile d'ordre $1 - \alpha$ pour $\alpha \in (0, 1)$. Cela définit la puissance du test sous $H_\beta^{(n)}$.

La partie iii) est une conséquence directe de la Remarque 4.1 du paragraphe 4.4.3 de Dreesbeke et Fine (1996). \square

3.3 Cas où les fonctions T et V sont paramétriques et la densité du bruit est connue

Maintenant on se place dans le cadre du modèle (3.1) avec les fonctions T et V de formes connues mais dépendant de paramètres inconnus. Plus précisément, pour tout $x \in \mathbb{R}^p$ $T(x) = m(x; \rho)$, $V(x) = \sigma(x; \theta)$, $\rho \in \Theta \subset \mathbb{R}^l$ et $\theta \in \widetilde{\Theta} \subset \mathbb{R}^q$. Soit $\psi_0 = (\rho_0^\top, \theta_0^\top)^\top \in \Theta \times \widetilde{\Theta} \subset \mathbb{R}^l \times \mathbb{R}^q$ le vrai paramètre du modèle (3.1). Soit pour tout $t = 1, \dots, n$, $\gamma \in \mathbb{R}^{k+1}$ et $\psi = (\rho^\top, \theta^\top)^\top \in \Theta \times \widetilde{\Theta}$,

$$\varepsilon_t(\psi, \gamma) = \frac{X_t - m(Z_{t-1}, \rho) - \gamma^\top \omega(t)}{\sigma(Z_{t-1}, \theta)}. \quad (3.2)$$

On considère les hypothèses supplémentaires suivantes :

- (B_0) : Pour tout $\theta \in \widetilde{\Theta}$ et $z \in \mathbb{R}^p$, $\sigma(z, \theta) > \tau$, où τ est un réel positif.
- (B_1) : Il existe $\delta \geq 0$ tel que $\int_{\mathbb{R}^p} \|x\|^{2+\delta} dF_j(x) < \infty$, $j = 1, \dots, k+1$.

- (B_2) : $\int_{\mathbb{R}} |x\phi'_f(x)|f(x)dx < \infty$.
- (B_3) : Soient $\text{Int}(\Theta)$ et $\text{Int}(\tilde{\Theta})$, les intérieurs respectifs de Θ et $\tilde{\Theta}$. Les fonctions $m(z, \rho)$ et $\sigma(z, \theta)$ sont continues et différentiables par rapport à $\rho \in \text{Int}(\Theta)$ et $\theta \in \text{Int}(\tilde{\Theta})$ respectivement, et il existe deux nombres finis positifs s_1 et s_2 tels que $\bar{B}(\rho_0, s_1) \subset \text{Int}(\Theta)$, $\bar{B}(\theta_0, s_2) \subset \text{Int}(\tilde{\Theta})$ et une fonction positive ϑ définie sur \mathbb{R}^p , telle que les fonctions $\sup_{\rho \in \bar{B}(\rho_0, s_1)} \|\partial m(z, \rho)\|$, $\sup_{\theta \in \bar{B}(\theta_0, s_2)} |\sigma(z, \theta)|$, $\sup_{\theta \in \bar{B}(\theta_0, s_2)} \|\partial \sigma(z, \theta)\|$ et $\sup_{\theta \in \bar{B}(\theta_0, s_2)} \|\partial^2 \sigma(z, \theta)\|_M$ soient majorées par $\vartheta(z)$.
- (B_4) : Le vrai paramètre $\psi_0 = (\rho_0^\top, \theta_0^\top)^\top$ admet un estimateur consistant $\psi_n = (\rho_n^\top, \theta_n^\top)^\top$ qui satisfait l'égalité de Bahadur suivante :

$$n^{\frac{1}{2}}(\psi_n - \psi_0) = n^{-\frac{1}{2}} \sum_{t=1}^n \Psi(Z_{t-1}, \psi_0) \Omega(\varepsilon_t(\psi_0, \gamma_0)) + o_P(1),$$

où pour tout $z \in \mathbb{R}^p$,

$$\Psi(z, \psi_0) = (\Psi_1^\top(z, \psi_0), \Psi_2^\top(z, \psi_0))^\top,$$

$$\Psi_1(z, \psi_0) = (\Psi_{11}(z, \psi_0), \dots, \Psi_{1l}(z, \psi_0))^\top \in \mathbb{R}^l,$$

$$\Psi_2(z, \psi_0) = (\Psi_{21}(z, \psi_0), \dots, \Psi_{2q}(z, \psi_0))^\top \in \mathbb{R}^q,$$

tel que pour tout $j = 1, \dots, k+1$, $\mathbb{E}(\|\Psi(Z_j, \psi_0)\|^{2+\delta}) < \infty$, et pour tout $x \in \mathbb{R}$, $\Omega(x) = (\Omega_1(x), \Omega_2(x))^\top \in \mathbb{R}^2$ tel que $\int_{\mathbb{R}} \|\Omega(x)\|^{2+\delta} f(x) dx < \infty$ et $\int_{\mathbb{R}} \Omega(x) f(x) dx = 0 \in \mathbb{R}^2$.

Remarque 6. Les fonctions Ψ et Ω dans l'égalité de Bahadur ne sont pas explicitées car elles dépendent des fonctions T et V . Cependant, (B_4) est vérifiée par les estimateurs classiques comme l'estimateur du maximum de vraisemblance et celui des moindres carrés, pour les modèles usuels.

Soit pour tout $j = 1, \dots, k+1$, pour tout $i = 1, 2, 3$

$$\chi_{i,j} = \int_{\mathbb{R}^p} \vartheta^i(x) dF_j(x).$$

Soit pour tout $\psi = (\rho^\top, \theta^\top)^\top \in \Theta \times \tilde{\Theta}$ et $(\gamma, \beta) \in \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$,

$$\Lambda_n(\psi, \gamma, \beta) = \sum_{t=1}^n \log \left\{ f \left[\varepsilon_t \left(\psi, \gamma + \frac{\beta}{\sqrt{n}} \right) \right] \right\} - \log \{ f[\varepsilon_t(\psi, \gamma)] \}.$$

Avec cette définition, le logarithme du rapport de vraisemblance de H_0 contre $H_\beta^{(n)}$ c'est $\Lambda_n(\psi_0, \gamma_0, \beta)$.

On a les propriétés suivantes pour ψ_n :

Proposition 3.3.1. *Sous (A') , (A_0) - (A_5) et sous les conditions données dans (B_4) , on a,*

i) *Sous H_0 ,*

$$\sqrt{n}(\psi_n - \psi_0) \xrightarrow{Loi} \mathcal{N}(0, \Sigma),$$

ii) *Sous $H_\beta^{(n)}$,*

$$\sqrt{n}(\psi_n - \psi_0) \xrightarrow{Loi} \mathcal{N}(\nu, \Sigma),$$

où

$$\nu := \int_{\mathbb{R}} \Omega(x) \phi_f(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \beta_j \int_{\mathbb{R}^p} \frac{\Psi(x, \psi_0)}{\sigma(x, \theta_0)} dF_j(x),$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

avec

$$\Sigma_{11} := \int_{\mathbb{R}} \Omega_1^2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^p} \Psi_1(x, \psi_0) \Psi_1^\top(x, \psi_0) dF_j(x),$$

$$\Sigma_{12} := \int_{\mathbb{R}} \Omega_1(x) \Omega_2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^p} \Psi_1(x, \psi_0) \Psi_2^\top(x, \psi_0) dF_j(x),$$

$$\Sigma_{21} := \int_{\mathbb{R}} \Omega_1(x) \Omega_2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^p} \Psi_1^\top(x, \psi_0) \Psi_2(x, \psi_0) dF_j(x),$$

$$\Sigma_{22} := \int_{\mathbb{R}} \Omega_2^2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^p} \Psi_2(x, \psi_0) \Psi_2^\top(x, \psi_0) dF_j(x).$$

Démonstration. i) Soit pour tout $1 \leq j \leq n$,

$$L_{n,j} = n^{-\frac{1}{2}} \sum_{t=1}^j (\Psi_1(Z_{t-1}, \psi_0) \Omega_1(\varepsilon_t(\psi_0, \gamma_0)), \Psi_2(Z_{t-1}, \psi_0) \Omega_2(\varepsilon_t(\psi_0, \gamma_0))) + o_P(1),$$

et $\eta = (\eta_1^\top, \eta_2^\top)^\top \in \mathbb{R}^{l+q}$ tel que $\eta_1 = (\eta_1^1, \dots, \eta_1^l)^\top$ et $\eta_2 = (\eta_2^1, \dots, \eta_2^q)^\top$.

On constate que pour tout $n \geq 1$ et $1 \leq j \leq n$,

$$\eta^\top L_{n,j} = n^{-\frac{1}{2}} \sum_{t=1}^j a_t(\eta),$$

où

$$a_t(\eta) = \sum_{\ell=1}^2 \eta_\ell^\top \Psi_\ell(Z_{t-1}, \psi_0) \Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)].$$

Il est facile de voir que, pour tout $j = 1, \dots, n$, $\eta^\top L_{n,j}$ est centré. Nous montrons d'abord que $\{(\eta^\top L_{n,j}, G_j), j = 1, \dots, n\}$ est une suite de martingales.

En effet, soient $j_1, j_2 \in \mathbb{Z}$ tels que $j_1 < j_2$ on a :

$$\begin{aligned} \mathbb{E}(\eta^\top L_{n,j_2} | G_{j_1}) &= \mathbb{E}(\eta^\top L_{n,j_1} | G_{j_1}) + n^{-\frac{1}{2}} \sum_{t=j_1+1}^{j_2} \mathbb{E}[a_t(\eta) | G_{j_1}] \\ &= \eta^\top L_{n,j_1} + n^{-\frac{1}{2}} \sum_{t=j_1+1}^{j_2} \sum_{\ell=1}^2 \mathbb{E}[\eta^\top \Psi_\ell(Z_{t-1}, \psi_0) \Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)] | G_{j_1}] \\ &= \eta^\top L_{n,j_1} + n^{-\frac{1}{2}} \sum_{t=j_1+1}^{j_2} \sum_{\ell=1}^2 \mathbb{E}[\eta^\top \Psi_\ell(Z_{t-1}, \psi_0)] \mathbb{E}\{\Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)]\}. \end{aligned}$$

Or, pour tout $\ell = 1, 2$ et $t \in \mathbb{Z}$, $\mathbb{E}\{\Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)]\} = 0$. Par suite, $\mathbb{E}(\eta^\top L_{n,j_2} | G_{j_1}) = \eta^\top L_{n,j_1}$. On a aussi

$$\sum_{t=1}^n \mathbb{E}\{[n^{-\frac{1}{2}} a_t(\eta)]^2 | G_{t-1}\} = \frac{1}{n} \sum_{t=1}^n \mathbb{E}[a_t^2(\eta) | G_{t-1}]. \quad (3.3)$$

En remplaçant $a_t^2(\eta)$ par son expression, on obtient

$$\begin{aligned} &\sum_{t=1}^n \mathbb{E}\{[n^{-\frac{1}{2}} a_t(\eta)]^2 | G_{t-1}\} \\ &= \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left[\left(\sum_{\ell=1}^2 \eta_\ell^\top \Psi_\ell(Z_{t-1}, \psi_0) \Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)] \right)^2 \middle| G_{t-1} \right] \\ &= \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left[(\eta_1^\top \Psi_1(Z_{t-1}, \psi_0) \Omega_1[\varepsilon_t(\psi_0, \gamma_0)] + \eta_2^\top \Psi_2(Z_{t-1}, \psi_0) \Omega_2[\varepsilon_t(\psi_0, \gamma_0)])^2 \middle| G_{t-1} \right]. \end{aligned}$$

En développant cette expression et en utilisant le fait que pour tout $t \in \mathbb{Z}$, ε_t est indépendante à G_{t-1} , on a

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}\{[n^{-\frac{1}{2}} a_t(\eta)]^2 | G_{t-1}\} &= \frac{1}{n} \sum_{t=1}^n [\eta_1^\top \Psi_1(Z_{t-1}, \psi_0)]^2 \mathbb{E}\{\Omega_1^2[\varepsilon_t(\psi_0, \gamma_0)]\} \\ &\quad + \frac{1}{n} \sum_{t=1}^n [\eta_2^\top \Psi_2(Z_{t-1}, \psi_0)]^2 \mathbb{E}\{\Omega_2^2[\varepsilon_t(\psi_0, \gamma_0)]\} \\ &\quad + \frac{1}{n} \sum_{t=1}^n \eta_1^\top \Psi_1(Z_{t-1}, \psi_0) \eta_2^\top \Psi_2(Z_{t-1}, \psi_0) \\ &\quad \times \mathbb{E}\{\Omega_1[\varepsilon_t(\psi_0, \gamma_0)] \Omega_2[\varepsilon_t(\psi_0, \gamma_0)]\} \end{aligned}$$

$$\begin{aligned}
&= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} (\eta_1^\top \Psi_1(Z_{t-1}, \psi_0))^2 \mathbb{E} \left\{ \Omega_1^2[\varepsilon_t(\psi_0, \gamma_0)] \right\} \\
&+ \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} [\eta_2^\top \Psi_2(Z_{t-1}, \psi_0)]^2 \mathbb{E} \left\{ \Omega_2^2[\varepsilon_t(\psi_0, \gamma_0)] \right\} \\
&+ 2 \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \eta_1^\top \Psi_1(Z_{t-1}, \psi_0) \eta_2^\top \Psi_2(Z_{t-1}, \psi_0) \\
&\times \mathbb{E} \left\{ \Omega_1[\varepsilon_t(\psi_0, \gamma_0)] \Omega_2[\varepsilon_t(\psi_0, \gamma_0)] \right\} \\
&= S_n^{(1)} + S_n^{(2)} + S_n^{(3)}.
\end{aligned}$$

D'après le théorème ergodique,

$$S_n^{(1)} \xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^p} [\eta_1^\top \Psi_1(x, \psi_0)]^2 dF_j(x) \int_{\mathbb{R}} \Omega_1^2(x) f(x) dx < \infty,$$

$$S_n^{(2)} \xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^p} t [\eta_2^\top \Psi_2(x, \psi_0)]^2 dF_j(x) \int_{\mathbb{R}} \Omega_2^2(x) f(x) dx < \infty,$$

et

$$S_n^{(3)} \xrightarrow{p.s.} 2 \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^p} \eta_1^\top \Psi_1(x, \psi_0) \eta_2^\top \Psi_2(x, \psi_0) dF_j(x) \int_{\mathbb{R}} \Omega_1(x) \Omega_2(x) f(x) dx < \infty.$$

Ainsi,

$$\sum_{t=1}^n \mathbb{E}[(n^{-\frac{1}{2}} a_t(\eta))^2 | G_{t-1}] \xrightarrow{p.s.} \xi,$$

avec

$$\begin{aligned}
\xi &:= \sum_{j=1}^{k+1} \alpha_j \left\{ \int_{\mathbb{R}} \Omega_1^2(x) f(x) dx \int_{\mathbb{R}^p} [\eta_2^\top \Psi_1(x, \psi_0)]^2 dF_j(x) \right. \\
&+ \int_{\mathbb{R}} \Omega_2^2(x) f(x) dx \int_{\mathbb{R}^p} [\eta_1^\top \Psi_2(x, \psi_0)]^2 dF_j(x) \\
&\left. + \int_{\mathbb{R}} \Omega_1(x) \Omega_2(x) f(x) dx \int_{\mathbb{R}^p} \eta_1^\top \Psi_1(x, \psi_0) \eta_2^\top \Psi_2(x, \psi_0) dF_j(x) \right\}.
\end{aligned}$$

Maintenant on montre la condition de Lindeberg :

Soit $\epsilon > 0$, par l'inégalité de Hölder conditionnelle, on a

$$\begin{aligned}
&\sum_{t=1}^n \mathbb{E} \left\{ [n^{-\frac{1}{2}} a_t(\eta)]^2 \mathbb{1}_{|n^{-\frac{1}{2}} a_t(\eta)| > \epsilon} | G_{t-1} \right\} \\
&\leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ |n^{-\frac{1}{2}} a_t(\eta)|^3 | G_{t-1} \right\} \mathbb{E}^{\frac{1}{3}} \left[\mathbb{1}_{|n^{-\frac{1}{2}} a_t(\eta)| > \epsilon} | G_{t-1} \right] \\
&= \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ |n^{-\frac{1}{2}} a_t(\eta)|^3 | G_{t-1} \right\} P^{\frac{1}{3}} \left[|n^{-\frac{1}{2}} a_t(\eta)| > \epsilon | G_{t-1} \right].
\end{aligned}$$

D'après l'inégalité de Markov conditionnelle, on a

$$\begin{aligned} & \sum_{t=1}^n \mathbb{E} \left\{ [n^{-\frac{1}{2}} a_t(\eta)]^2 \mathbf{1}_{|n^{-\frac{1}{2}} a_t(\eta)| > \epsilon} | G_{t-1} \right\} \\ & \leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ |n^{-\frac{1}{2}} a_t(\eta)|^3 | G_{t-1} \right\} \frac{\mathbb{E}^{\frac{1}{3}} \left[|n^{-\frac{1}{2}} a_t(\eta)|^3 > \epsilon | G_{t-1} \right]}{\epsilon^{\frac{1}{3}}} \\ & \leq Cst \times n^{-\frac{3}{2}} \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ |a_t(\eta)|^3 | G_{t-1} \right\} \mathbb{E}^{\frac{1}{3}} \left[|a_t(\eta)|^3 > \epsilon | G_{t-1} \right]. \end{aligned}$$

Puisque

$$[a_t(\eta)]^3 \leq Cst \sum_{\ell=1}^2 \|\eta_\ell^\top \Psi_\ell(Z_{t-1}, \psi_0)\|^3 |\Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)]|^3,$$

on a

$$\begin{aligned} & \sum_{t=1}^n \mathbb{E} \left\{ [n^{-\frac{1}{2}} a_t(\eta)]^2 \mathbf{1}_{|n^{-\frac{1}{2}} a_t(\eta)| > \epsilon} | G_{t-1} \right\} \\ & \leq Cst.n^{-\frac{3}{2}} \sum_{t=1}^n \sum_{\ell=1}^2 \|\eta_\ell^\top \Psi_\ell(Z_{t-1}, \psi_0)\|^3 \mathbb{E} \{ |\Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)]|^3 \} \\ & = Cst \frac{1}{\sqrt{n}} \sum_{\ell=1}^2 \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \|\eta_\ell^\top \Psi_\ell(Z_{t-1}, \psi_0)\|^3 \\ & \times \mathbb{E} \{ |\Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)]|^3 \}. \end{aligned}$$

D'après le théorème ergodique, lorsque $n \rightarrow \infty$,

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{\ell=1}^2 \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \|\eta_\ell^\top \Psi_\ell(Z_{t-1}, \psi_0)\|^3 \mathbb{E} \{ |\Omega_\ell[\varepsilon_t(\psi_0, \gamma_0)]|^3 \} \\ & \xrightarrow{p.s.} 0 \times \sum_{\ell=1}^2 \sum_{j=1}^{k+1} \alpha_j \mathbb{E} \left[\|\eta_\ell^\top \Psi_\ell(Z_{t-1}, \psi_0)\|^3 \right] \int_{\mathbb{R}} |\Omega_\ell(x)|^3 f(x) dx = 0. \end{aligned}$$

D'après corollaire 3.1 de Hall et Heyde (1980), sous H_0 , lorsque $n \rightarrow \infty$,

$$\eta^\top \sqrt{n}(\psi_n - \psi_0) \xrightarrow{Loi} \mathcal{N}(0, \xi).$$

Ce qui implique que, sous H_0 , lorsque $n \rightarrow \infty$,

$$\sqrt{n}(\psi_n - \psi_0) \xrightarrow{Loi} \mathcal{N}(0, \Sigma),$$

où

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix},$$

avec Σ_{11} , Σ_{12} , Σ_{21} et Σ_{22} définies dans l'énoncé de la proposition.

ii) On sait que, sous H_0 ,

$$\begin{aligned} \sqrt{n}(\psi_n - \psi_0) &\xrightarrow{Loi} \mathcal{N}(0, \Sigma), \\ \Lambda_n(\psi_0, \gamma_0, \beta) &\xrightarrow{Loi} \mathcal{N}\left(-\frac{1}{2}\mu(\psi_0, \gamma_0, \beta), \mu(\psi_0, \gamma_0, \beta)\right), \end{aligned}$$

avec

$$\begin{aligned} \mu(\psi_0, \gamma_0, \beta) &= \sum_{j=1}^{k+1} \alpha_j \beta_j^2 \mu_{j,2}(\gamma_0, \psi_0), \\ \mu_{j,2}(\gamma_0, \psi_0) &= I(f) \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_j(x), \quad j = 1, \dots, k+1. \end{aligned}$$

Soit $S_n = \sqrt{n}(\psi_n - \psi_0)$. Alors, sous H_0 , $(S_n^\top, \Lambda_n(\psi_0, \gamma_0, \beta))^\top$ suit asymptotiquement une loi normale d'espérance $(0^\top, -\mu(\psi_0, \beta)/2)^\top$ et de vecteur de covariance :

$$\begin{aligned} \lim_{n \rightarrow \infty} Cov(S_n, \Lambda_n(\psi_0, \gamma_0, \beta)) &= \lim_{n \rightarrow \infty} Cov\left(S_n, \Delta_n(\psi_0, \gamma_0, \beta) - \frac{1}{2}\mu(\psi_0, \beta)\right) \\ &= \lim_{n \rightarrow \infty} Cov(S_n, \Delta_n(\psi_0, \gamma_0, \beta)). \end{aligned}$$

En remplaçant $\Delta_n(\psi_0, \gamma_0, \beta)$ par son expression, on a

$$\begin{aligned} &Cov(S_n, \Lambda_n(\psi_0, \gamma_0, \beta)) \\ &= Cov\left\{S_n, \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)]\right\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) Cov\left\{S_n, \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \left\{ \mathbb{E}\left[S_n \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right] - \mathbb{E}(S_n) \mathbb{E}\left[\frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right] \right\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \mathbb{E}\left\{S_n \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right\} - \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \mathbb{E}(S_n) \mathbb{E}\left\{\frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right\} \\ &= C_{1n} + C_{2n}. \end{aligned}$$

Puisque $\mathbb{E}\{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]\} = 0$, $t \in \mathbb{Z}$, alors $C_{2n} = 0$, $\forall n$.

$$\begin{aligned} C_{1n} &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \mathbb{E}\left\{S_n \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right\} \\ &= \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \sum_{j=1}^n \mathbb{E}\left\{\Psi(Z_{j-1}, \psi_0) \Omega[\varepsilon_j(\psi_0, \gamma_0)] \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right\} \\ &= \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \mathbb{E}\left\{\Psi(Z_{t-1}, \psi_0) \Omega[\varepsilon_t(\psi_0, \gamma_0)] \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right\} \\ &\quad + \frac{1}{n} \sum_{t,s=1, t \neq s}^n \beta^\top \omega(t) \mathbb{E}\left\{\Psi(Z_{s-1}, \psi_0) \Omega[\varepsilon_s(\psi_0, \gamma_0)] \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)}\right\} \end{aligned}$$

Le deuxième terme de la dernière égalité est nul. En effet, puisque pour tout $t \in \mathbb{Z}$, les ε_t sont indépendantes et ε_t indépendante de $G_{t-1} = \sigma(Z_1, \dots, Z_t)$, alors,

$$\begin{aligned} & \frac{1}{n} \sum_{t,s=1, t \neq s}^n \beta^\top \omega(t) \mathbb{E} \left\{ \Psi(Z_{s-1}, \psi_0) \Omega[\varepsilon_s(\psi_0, \gamma_0)] \frac{\phi_f[\varepsilon_t(\psi_0, \gamma_0)]}{\sigma(Z_{t-1}, \theta_0)} \right\} \\ &= \frac{1}{n} \sum_{t,s=1, t \neq s}^n \beta^\top \omega(t) \mathbb{E} \left[\frac{\Psi(Z_{s-1}, \psi_0)}{\sigma(Z_{t-1}, \theta_0)} \right] \mathbb{E} \left\{ \Omega[\varepsilon_s(\psi_0, \gamma_0)] \right\} \mathbb{E} \left\{ \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right\} \\ &= 0 \end{aligned}$$

car pour tout $t \in \mathbb{Z}$, $\mathbb{E} \left\{ \Omega[\varepsilon_t(\psi_0, \gamma_0)] \right\} = 0$ et pour tout $t \in \mathbb{Z}$, $\mathbb{E} \left\{ \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right\} = 0$.

Donc

$$\begin{aligned} \text{Cov}(S_n, \Lambda_n(\psi_0, \gamma_0, \beta)) &= \int_{\mathbb{R}} \Omega(x) \phi_f(x) f(x) dx \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \mathbb{E} \left(\frac{\Psi(Z_{t-1}, \psi_0)}{\sigma(Z_{t-1}, \theta_0)} \right) \\ &= \int_{\mathbb{R}} \Omega(x) \phi_f(x) f(x) dx \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{\beta_j}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \mathbb{E} \left(\frac{\Psi(Z_{t-1}, \psi_0)}{\sigma(Z_{t-1}, \theta_0)} \right). \end{aligned}$$

Or, le membre de droite de l'égalité ci-dessus tend presque sûrement vers

$$\int_{\mathbb{R}} \Omega(x) \phi_f(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \beta_j \int_{\mathbb{R}^p} \frac{\Psi(x, \psi_0)}{\sigma(x, \theta_0)} dF_j(x) < \infty.$$

D'où, sous H_0 , lorsque $n \rightarrow \infty$,

$$\text{Cov}(S_n, \Lambda_n(\psi_0, \gamma_0, \beta)) \xrightarrow{p.s.} \nu,$$

$$\text{où } \nu = \int_{\mathbb{R}} \Omega(x) \phi_f(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \beta_j \int_{\mathbb{R}^p} \frac{\Psi(x, \psi_0)}{\sigma(x, \theta_0)} dF_j(x) < \infty.$$

Ainsi, sous H_0 ,

$$\begin{pmatrix} S_n \\ \Lambda_n(\psi_0, \gamma_0, \beta) \end{pmatrix} \xrightarrow{loi} \mathcal{N} \left(\begin{pmatrix} 0 \\ -\frac{1}{2} \mu(\psi_0, \gamma_0, \beta) \end{pmatrix}, \begin{pmatrix} \Sigma & \nu^\top \\ \nu & \mu(\psi_0, \gamma_0, \beta) \end{pmatrix} \right).$$

D'après troisième lemme de Le Cam, sous $H_\beta^{(n)}$,

$$S_n = \sqrt{n}(\psi_n - \psi_0) \xrightarrow{Loi} \mathcal{N}(\nu, \Sigma).$$

□

3.3.1 Cas où γ_0 est connue

Puisque pour tout $\psi \in \mathbb{R}^{l+q}$ et $j = 1, \dots, k+1$, $\mu_{j,2}(\psi, \gamma_0)$ dépend de F_j et ψ qui sont inconnues, on estime $\mu_{j,2}(\psi_0, \gamma_0)$ par $\widehat{\mu}_{j,2}(\widehat{\psi}_n, \gamma_0)$ défini pour tout $j = 1, \dots, k+1$ par

$$\widehat{\mu}_{j,2}(\widehat{\psi}_n, \gamma_0) = I(f) \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \widehat{\theta}_n)}.$$

Donc, on prend

$$\widehat{\mu}_n(\widehat{\psi}_n, \gamma_0, \beta) = \sum_{j=1}^{k+1} \widehat{\alpha}_j \beta_j^2 \widehat{\mu}_{j,2}(\widehat{\psi}_n, \gamma_0), \text{ et } \widehat{\varpi}_n(\widehat{\psi}_n, \gamma_0, \beta) = \widehat{\mu}_n^{\frac{1}{2}}(\widehat{\psi}_n, \gamma_0, \beta),$$

où $\widehat{\alpha}_j$ est un estimateur de $\alpha_j = \lim_{n \rightarrow \infty} n_j(n)/n$, $j = 1, \dots, k+1$. En pratique on peut prendre $\widehat{\alpha}_j = n_j(n)/n$.

Soit

$$\mathcal{T}_n(\psi_{N(n)}, \gamma_0, \beta) = \frac{\Delta_n(\psi_{N(n)}, \gamma_0, \beta)}{\widehat{\varpi}_n(\psi_{N(n)}, \gamma_0, \beta)},$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$.

Proposition 3.3.2. *Sous les hypothèses (A'), (A₀)-(A₅) et (B₀)-(B₄), pour tout estimateur consistant et asymptotiquement normal ψ_n de ψ_0 , on a pour tout $n \geq 1$,*

i)

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_{N(n)}, \gamma_0, \beta) + o_P(1),$$

ii)

$$\widehat{\varpi}_n(\widehat{\psi}_n, \gamma_0, \beta) \rightarrow \varpi(\psi_0, \gamma_0, \beta),$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$.

Démonstration. i) Par un développement de Taylor à l'ordre 2 au point ψ_n on a,

$$\begin{aligned} \Delta_n(\psi_0, \gamma_0, \beta) &= \Delta_n(\psi_n, \gamma_0, \beta) + (\psi_0 - \psi_n)^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) \\ &\quad + \frac{1}{2} (\psi_0 - \psi_n)^\top \partial_\psi^2 \Delta_n(\widetilde{\psi}_n, \gamma_0, \beta) (\psi_0 - \psi_n), \end{aligned}$$

où $\widetilde{\psi}_n$ comprise entre ψ_n et ψ_0 .

Montrons d'abord que pour tout $\widetilde{\psi}_n$ entre ψ_n et ψ_0 ,

$$(\psi_0 - \psi_n)^\top \partial_\psi^2 \Delta_n(\widetilde{\psi}_n, \gamma_0, \beta) (\psi_0 - \psi_n) = o_P(1).$$

On a,

$$\begin{aligned} \left| (\psi_0 - \psi_n)^\top \partial_\psi^2 \Delta_n(\widetilde{\psi}_n, \gamma_0, \beta) (\psi_0 - \psi_n) \right| &\leq \|\sqrt{n}(\psi_0 - \psi_n)\| \times \frac{1}{\sqrt{n}} \left\| \partial_\psi^2 \Delta_n(\widetilde{\psi}_n, \gamma_0, \beta) \right\|_M \\ &\quad \times \|\psi_0 - \psi_n\|. \end{aligned}$$

Or, d'après la Proposition 3.3.1, $\sqrt{n}(\psi_0 - \psi_n)$ converge en loi vers une loi normale et $\psi_0 - \psi_n$ tend vers zéro en probabilité. Donc pour montrer que $(\psi_0 - \psi_n)^\top \partial_\psi^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta)(\psi_0 - \psi_n) = o_P(1)$, il suffit de montrer que $\left\| \partial_\psi^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta) \right\|_M / \sqrt{n}$ converge en probabilité vers une variable aléatoire positive.

Pour tout $(\psi, \gamma, \beta) \in \mathbb{R}^{l+q} \times \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$, on définit

$$D_n(\psi, \gamma, \beta) = \frac{1}{\sqrt{n}} \partial_\psi^2 \Delta_n(\psi, \gamma, \beta) = \begin{pmatrix} D_{n,1,1}(\psi, \gamma, \beta) & D_{n,1,2}(\psi, \gamma, \beta) \\ D_{n,2,1}(\psi, \gamma, \beta) & D_{n,2,2}(\psi, \gamma, \beta) \end{pmatrix},$$

où

$$\begin{aligned} D_{n,1,1}(\psi, \gamma, \beta) &:= \frac{1}{\sqrt{n}} \partial_\rho^2 \Delta_n(\psi, \gamma, \beta), \\ D_{n,1,2}(\psi, \gamma, \beta) &:= \frac{1}{\sqrt{n}} \partial_{\rho\theta}^2 \Delta_n(\psi, \gamma, \beta), \\ D_{n,2,1}(\psi, \gamma, \beta) &:= \frac{1}{\sqrt{n}} \partial_{\theta\rho}^2 \Delta_n(\psi, \gamma, \beta), \\ D_{n,2,2}(\psi, \gamma, \beta) &:= \frac{1}{\sqrt{n}} \partial_\theta^2 \Delta_n(\psi, \gamma, \beta), \end{aligned}$$

et pour tout $\psi = (\rho^\top, \theta^\top)^\top \in \Theta \times \tilde{\Theta}$, on a

$$\Delta_n(\psi, \gamma_0, \beta) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma(Z_{t-1}, \theta)} \phi_f[\varepsilon_t(\psi, \gamma_0)].$$

On rappelle que pour tout $t = 1, \dots, n$, $\gamma \in \mathbb{R}^{k+1}$ et $\psi = (\rho^\top, \theta^\top)^\top \in \Theta \times \tilde{\Theta}$,

$$\varepsilon_t(\psi, \gamma) = \frac{X_t - m(Z_{t-1}, \rho) - \gamma^\top \omega(t)}{\sigma(Z_{t-1}, \theta)}.$$

D'où, pour tout $t = 1, \dots, n$, $\gamma \in \mathbb{R}^{k+1}$ et $\psi = (\rho^\top, \theta^\top)^\top \in \Theta \times \tilde{\Theta}$,

$$\partial_\rho \varepsilon_t(\psi, \gamma) = -\frac{1}{\sigma(Z_{t-1}, \theta)} \partial m(Z_{t-1}, \rho)$$

et

$$\partial_\theta \varepsilon_t(\psi, \gamma) = -\frac{\partial \sigma(Z_{t-1}, \theta)}{\sigma(Z_{t-1}, \theta)} \varepsilon_t(\psi, \gamma).$$

Ainsi,

$$\begin{aligned} \partial_\rho \Delta_n(\psi, \gamma_0, \beta) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)] \frac{\partial \varepsilon_t(\psi, \gamma_0)}{\partial \rho} \\ &= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \partial m(Z_{t-1}, \rho) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \end{aligned} \quad (3.4)$$

et

$$\begin{aligned}
 \partial_\theta \Delta_n(\psi, \gamma_0, \beta) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \\
 &\times \left\{ \sigma(Z_{t-1}, \theta) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \frac{\partial \varepsilon_t(\psi, \gamma_0)}{\partial \theta} - \partial \sigma(Z_{t-1}, \theta) \phi_f[\varepsilon_t(\psi, \gamma_0)] \right\} \\
 &= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)] \partial \sigma(Z_{t-1}, \theta) \varepsilon_t(\psi, \gamma_0) \\
 &- \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \partial \sigma(Z_{t-1}, \theta) \phi_f[\varepsilon_t(\psi, \gamma_0)].
 \end{aligned} \tag{3.5}$$

Étude du comportement asymptotique de $D_{n,1,1}(\tilde{\psi}_n, \gamma_0, \beta)$

Pour tout $\psi \in \Theta \times \tilde{\Theta}$, on a

$$\begin{aligned}
 \partial_\rho^2 \Delta_n(\psi, \gamma_0, \beta) &= \frac{\partial}{\partial \rho} \left(\frac{\partial \Delta_n(\psi, \gamma_0, \beta)}{\partial \rho} \right) \\
 &= \frac{\partial}{\partial \rho} \left\{ -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \partial_\rho m(Z_{t-1}, \rho) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \right\} \\
 &= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \left\{ \partial_\rho^2 m(Z_{t-1}, \rho) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \right. \\
 &\quad \left. + \partial_\rho \varepsilon_t(\psi, \tilde{\gamma}_n) \phi''_f[\varepsilon_t(\psi, \gamma_0)] \partial_\rho m(Z_{t-1}, \rho) \right\} \\
 &= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \partial_\rho^2 m(Z_{t-1}, \rho) \phi'_f(\varepsilon_t(\psi, \gamma_0)) \\
 &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta)} \partial_\rho m(Z_{t-1}, \rho) (\partial_\rho m(Z_{t-1}, \rho))^\top \phi''_f[\varepsilon_t(\psi, \gamma_0)].
 \end{aligned}$$

On remplace ψ par $\tilde{\psi}_n$ dans $\partial_\rho^2 \Delta_n(\psi, \gamma_0, \beta)$ et on le multiplie par $1/\sqrt{n}$. On a

$$\begin{aligned}
 D_{n,1,1}(\tilde{\psi}_n, \gamma_0, \beta) &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \tilde{\theta}_n)} \partial_\rho^2 m(Z_{t-1}, \tilde{\rho}_n) \phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \\
 &\quad + \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \tilde{\theta}_n)} \partial_\rho m(Z_{t-1}, \tilde{\rho}_n) (\partial_\rho m(Z_{t-1}, \tilde{\rho}_n))^\top \phi''_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \\
 &= \Delta_{n1} + \Delta_{n2}.
 \end{aligned}$$

On a

$$|\Delta_{n1}| \leq \frac{1}{\tau^2} \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]|.$$

Et puisque ϕ'_f est lipschitzienne, alors

$$\begin{aligned}
 \left| \phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right| &\leq \left| \phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \right| + |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &\leq c_\phi \left| \varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0) \right| + |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]|.
 \end{aligned} \tag{3.6}$$

Pour majorer $\phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]$, on a besoin de majorer $\varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)$. Pour cela, on a, pour tout $\gamma \in \mathbb{R}^{k+1}$,

$$\varepsilon_t(\psi_n, \gamma) - \varepsilon_t(\psi_0, \gamma) = \frac{X_t - m(Z_{t-1}, \rho_n) - \gamma^\top \omega(t)}{\sigma(Z_{t-1}, \theta_n)} - \varepsilon_t(\psi_0, \gamma).$$

En ajoutant et retranchant $m(Z_{t-1}, \rho_0)\sigma^{-1}(Z_{t-1}, \theta_n)$, on a

$$\begin{aligned} \varepsilon_t(\psi_n, \gamma) - \varepsilon_t(\psi_0, \gamma) &= \frac{X_t - m(Z_{t-1}, \rho_0) - \gamma^\top \omega(t)}{\sigma(Z_{t-1}, \theta_n)} + \frac{m(Z_{t-1}, \rho_0) - m(Z_{t-1}, \rho_n)}{\sigma(Z_{t-1}, \theta_n)} \\ &\quad - \varepsilon_t(\psi_0, \gamma). \end{aligned}$$

En multipliant et en divisant le premier terme de l'égalité précédente par $\sigma(Z_{t-1}, \theta_0)$, on obtient,

$$\begin{aligned} \varepsilon_t(\psi_n, \gamma) - \varepsilon_t(\psi_0, \gamma) &= \varepsilon_t(\psi_0, \gamma) \frac{\sigma(Z_{t-1}, \theta_0)}{\sigma(Z_{t-1}, \theta_n)} + \frac{m(Z_{t-1}, \rho_0) - m(Z_{t-1}, \rho_n)}{\sigma(Z_{t-1}, \theta_n)} - \varepsilon_t(\psi_0, \gamma) \\ &= \left[\frac{\sigma(Z_{t-1}, \theta_0)}{\sigma(Z_{t-1}, \theta_n)} - 1 \right] \varepsilon_t(\psi_0, \gamma) + \frac{m(Z_{t-1}, \rho_0) - m(Z_{t-1}, \rho_n)}{\sigma(Z_{t-1}, \theta_n)}. \end{aligned}$$

Par un développement de Taylor à l'ordre 1 de la fonction $m(Z_{t-1}, \cdot)$ au point ρ_0 , on a

$$\varepsilon_t(\psi_n, \gamma) - \varepsilon_t(\psi_0, \gamma) = \left[\frac{\sigma(Z_{t-1}, \theta_0)}{\sigma(Z_{t-1}, \theta_n)} - 1 \right] \varepsilon_t(\psi_0, \gamma) + (\rho_0 - \rho_n)^\top \frac{\partial_\rho m(Z_{t-1}, \dot{\rho}_n)}{\sigma(Z_{t-1}, \theta_n)},$$

où $\dot{\rho}_n$ compris entre ρ_n et ρ_0 .

Ce qui implique que pour tout $\gamma \in \mathbb{R}^{k+1}$,

$$\varepsilon_t(\psi_n, \gamma) - \varepsilon_t(\psi_0, \gamma) = \frac{\sigma(Z_{t-1}, \theta_0) - \sigma(Z_{t-1}, \theta_n)}{\sigma(Z_{t-1}, \theta_n)} \varepsilon_t(\psi_0, \gamma) + (\rho_0 - \rho_n)^\top \frac{\partial_\rho m(Z_{t-1}, \dot{\rho}_n)}{\sigma(Z_{t-1}, \theta_n)}.$$

Encore une fois, par un développement de Taylor à l'ordre 1 de la fonction $\sigma(Z_{t-1}, \cdot)$ au point θ_0 , on a, pour tout $\gamma \in \mathbb{R}^{k+1}$,

$$\varepsilon_t(\psi_n, \gamma) - \varepsilon_t(\psi_0, \gamma) = (\theta_0 - \theta_n)^\top \frac{\partial_\theta \sigma(Z_{t-1}, \dot{\theta}_n)}{\sigma(Z_{t-1}, \theta_n)} \varepsilon_t(\psi_0, \gamma) + (\rho_0 - \rho_n)^\top \frac{\partial_\rho m(Z_{t-1}, \dot{\rho}_n)}{\sigma(Z_{t-1}, \theta_n)} \quad (3.7)$$

où $\dot{\psi}_n = (\dot{\rho}_n, \dot{\theta}_n)$ est comprise entre $\psi_n = (\rho_n, \theta_n)$ et $\psi_0 = (\rho_0, \theta_0)$.

Maintenant, revenant à l'inégalité (3.6), en majorant $\varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)$, on obtient,

$$\begin{aligned} \left| \phi'_f(\varepsilon_t(\tilde{\psi}_n, \gamma_0)) \right| &\leq c_\phi \left[\|\theta_0 - \tilde{\theta}_n\| \frac{\|\partial_\theta \sigma(Z_{t-1}, \hat{\theta}_n)\|}{\sigma(Z_{t-1}, \tilde{\theta}_n)} \|\varepsilon_t(\psi_0, \gamma_0)\| \right. \\ &\quad \left. + \|\rho_0 - \tilde{\rho}_n\| \frac{\|\partial_\rho m(Z_{t-1}, \hat{\rho}_n)\|}{\sigma(Z_{t-1}, \tilde{\theta}_n)} \right] \\ &\quad + |\phi'_f(\varepsilon_t(\psi_0, \gamma_0))|, \end{aligned} \quad (3.8)$$

avec $\widehat{\psi}_n$ comprise entre $\widetilde{\psi}_n$ et ψ_0 .

En utilisant l'hypothèse (B_0) , il s'en suit que

$$\left| \phi'_f[\varepsilon_t(\widetilde{\psi}_n, \gamma_0)] \right| \leq Cst \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| + Cst \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) + |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]|.$$

D'où,

$$\begin{aligned} |\Delta_{n1}| &\leq Cst \|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \\ &+ Cst \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) \\ &+ Cst \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]|. \end{aligned} \quad (3.9)$$

D'après le théorème ergodique,

$$\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|$$

$$\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} \mathbb{E}(|\varepsilon_1(\psi_0, \gamma_0)|) < \infty.$$

$$\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) \xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} < \infty,$$

et

$$\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]|$$

$$\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{1,j} \int_{\mathbb{R}} |\phi'_f(x)| f(x) dx < \infty.$$

Puisque $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ tend vers 0 en probabilité, alors, le membre de droite de l'inégalité (3.9) converge en probabilité vers

$$a = \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} \mathbb{E}(|\varepsilon_1(\psi_0, \gamma_0)|) + \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{1,j} \int_{\mathbb{R}} |\phi'_f(x)| f(x) dx + \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} < \infty.$$

Donc Δ_{1n} converge en probabilité vers une limite dans l'intervalle $[-a, a]$. Ce qui implique que Λ_{1n} est tendue.

Pour la convergence de Δ_{n2} , sous les hypothèses (B_0) et le fait que ϕ_f'' est bornée, on peut remarquer que

$$|\Delta_{n2}| \leq Cst \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}).$$

Or, par le théorème ergodique,

$$\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) \xrightarrow{p.s.} b = \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} < \infty.$$

Ce qui implique que Δ_{n2} tend en probabilité vers une limite finie qui est dans l'intervalle $[-b, b]$. On peut donc conclure que $D_{n,1,1}(\tilde{\psi}_n, \gamma_0, \beta)$ converge en probabilité vers une limite finie appartenant à l'intervalle $[-(a+b), a+b]$.

Étude du comportement asymptotique de $D_{n,2,2}(\tilde{\psi}_n, \gamma_0, \beta)$

Pour tout $\psi \in \Theta \times \tilde{\Theta}$, on a

$$\begin{aligned} \partial_\theta^2 \Delta_n(\psi, \gamma_0, \beta) &= \frac{\partial}{\partial \theta} \left(\frac{\partial \Delta_n(\psi, \gamma_0, \beta)}{\partial \theta} \right) \\ &= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial}{\partial \theta} \left\{ \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \phi_f'[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \right\} \\ &\quad - \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial}{\partial \theta} \left\{ \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \phi_f[\varepsilon_t(\psi, \gamma_0)] \right\}. \end{aligned} \quad (3.10)$$

Or,

$$\begin{aligned} &\partial_\theta \left\{ \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \phi_f'[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \right\} \\ &= \partial_\theta \left[\frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \right] \phi_f'[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) + \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \frac{\partial}{\partial \theta} \left\{ \phi_f'[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \right\} \\ &= \frac{1}{\sigma^4(Z_{t-1}, \theta)} \left\{ \partial_\theta^2 \sigma(Z_{t-1}, \theta) \sigma^2(Z_{t-1}, \theta) - 2\sigma(Z_{t-1}, \theta) \partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top \right\} \\ &\quad \times \phi_f'[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \\ &\quad + \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \left\{ \phi_f''[\varepsilon_t(\psi, \gamma_0)] \partial_\theta \varepsilon_t(\psi, \tilde{\gamma}_n) \varepsilon_t(\psi, \gamma_0) + \partial_\theta \varepsilon_t(\psi, \gamma_0) \phi_f'[\varepsilon_t(\psi, \gamma_0)] \right\} \\ &= \frac{1}{\sigma^4(Z_{t-1}, \theta)} \left\{ \partial_\theta^2 \sigma(Z_{t-1}, \theta) \sigma^2(Z_{t-1}, \theta) - 2\sigma(Z_{t-1}, \theta) \partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top \right\} \\ &\quad \times \phi_f'[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \\ &\quad + \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \left\{ -\frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma(Z_{t-1}, \theta)} \varepsilon_t^2(\psi, \tilde{\gamma}_n) \phi_f''[\varepsilon_t(\psi, \gamma_0)] \right. \\ &\quad \left. - \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma(Z_{t-1}, \theta)} \phi_f'[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \right\} \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{\sigma^4(Z_{t-1}, \theta)} \left\{ \partial_\theta^2 \sigma(Z_{t-1}, \theta) \sigma^2(Z_{t-1}, \theta) - 2\sigma(Z_{t-1}, \theta) \partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top \right\} \\
 &\times \phi'_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \\
 &- \frac{\partial_\theta \sigma(Z_{t-1}, \theta) (\partial_\theta \sigma(Z_{t-1}, \theta))^\top}{\sigma^3(Z_{t-1}, \theta)} \left\{ \varepsilon_t^2(\psi, \gamma_0) \phi''_f[\varepsilon_t(\psi, \gamma_0)] + \varepsilon_t(\psi, \gamma_0) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \right\}.
 \end{aligned}$$

Et

$$\begin{aligned}
 &\partial_\theta \left\{ \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \phi_f[\varepsilon_t(\psi, \gamma_0)] \right\} \\
 &= \partial_\theta \left[\frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \right] \phi_f[\varepsilon_t(\psi, \gamma_0)] + \phi'_f[\varepsilon_t(\psi, \gamma_0)] \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \frac{\partial \varepsilon_t}{\partial \theta}(\psi, \gamma_0) \\
 &= \frac{1}{\sigma^4(Z_{t-1}, \theta)} \left\{ \partial_\theta^2 \sigma(Z_{t-1}, \theta) \sigma^2(Z_{t-1}, \theta) - 2\sigma(Z_{t-1}, \theta) \partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top \right\} \\
 &\times \phi_f[\varepsilon_t(\psi, \gamma_0)] - \phi'_f[\varepsilon_t(\psi, \gamma_0)] \frac{\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top}{\sigma^3(Z_{t-1}, \theta)} \varepsilon_t(\psi, \gamma_0).
 \end{aligned}$$

On remplace

$$\partial_\theta \left\{ \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \right\}$$

et

$$\partial_\theta \left\{ \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \phi_f[\varepsilon_t(\psi, \gamma_0)] \right\}$$

par leurs expressions dans (3.10), et on obtient, pour tout $\psi \in \Theta \times \tilde{\Theta}$,

$$\begin{aligned}
 \partial_\theta^2 \Delta_n(\psi, \gamma_0) &= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \left\{ \frac{\partial_\theta^2 \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} - \frac{2\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top}{\sigma^3(Z_{t-1}, \theta)} \right\} \\
 &\times \phi'_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \\
 &- \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top}{\sigma^3(Z_{t-1}, \theta)} \left\{ \phi''_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t^2(\psi, \gamma_0) \right. \\
 &+ \left. \phi'_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \right\} \\
 &- \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \left\{ \frac{\partial_\theta^2 \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} - \frac{2\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top}{\sigma^3(Z_{t-1}, \theta)} \right\} \\
 &\times \phi_f[\varepsilon_t(\psi, \gamma_0)] \\
 &+ \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top}{\sigma^3(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0).
 \end{aligned}$$

On remplace ψ par $\tilde{\psi}_n$ dans $\partial_\theta^2 \Delta_n(\psi, \gamma_n)$ et on le multiplie par $1/\sqrt{n}$, on a

$$\begin{aligned}
 D_{n,2,2}(\tilde{\psi}_n, \gamma_0, \beta) &= \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \varepsilon_t(\tilde{\psi}_n, \gamma_0) \\
 &\times \left\{ \frac{4\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n) [\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n)]^\top}{\sigma^3(Z_{t-1}, \tilde{\theta}_n)} - \frac{\partial_\theta^2 \sigma(Z_{t-1}, \tilde{\theta}_n)}{\sigma^2(Z_{t-1}, \tilde{\theta}_n)} \right\}
 \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n) (\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n))^\top}{\sigma^3(Z_{t-1}, \tilde{\theta}_n)} \phi_f''[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \varepsilon_t^2(\tilde{\psi}_n, \gamma_0) \\
& + \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \left\{ \frac{2\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n) [\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n)]^\top}{\sigma^3(Z_{t-1}, \tilde{\theta}_n)} - \frac{\partial_\theta^2 \sigma(Z_{t-1}, \tilde{\theta}_n)}{\sigma^2(Z_{t-1}, \tilde{\theta}_n)} \right\} \\
& \times \phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \\
& = \Delta_{n3} + \Delta_{n4} + \Delta_{n5}.
\end{aligned}$$

Pour étudier la convergence de Δ_{n3} , Δ_{n4} et Δ_{n5} , nous avons besoin de quelques inégalités préliminaires. Elles concernent la majoration de $\phi_f'[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]$, $\varepsilon_t(\tilde{\psi}_n, \gamma_0)$ et $\phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]$ par des quantités appropriées.

Commençons par $\phi_f'[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]$.

On a

$$\left| \phi_f'[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right| \leq \left| \phi_f'[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] - \phi_f'[\varepsilon_t(\psi_0, \gamma_0)] \right| + \left| \phi_f'[\varepsilon_t(\psi_0, \gamma_0)] \right|.$$

En utilisant le fait que ϕ_f' est lipschitzienne, on a

$$\left| \phi_f'[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right| \leq c_\phi \left| \varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0) \right| + \left| \phi_f'[\varepsilon_t(\psi_0, \gamma_0)] \right|.$$

À l'aide de l'égalité (1.7), on a

$$\begin{aligned}
\left| \phi_f'[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right| & \leq Cst \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| + Cst \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) \\
& + \left| \phi_f'[\varepsilon_t(\psi_0, \gamma_0)] \right|.
\end{aligned} \tag{3.11}$$

Trouvons une majoration pour $|\varepsilon_t(\tilde{\psi}_n, \gamma_0)|$:

$$\begin{aligned}
\left| \varepsilon_t(\tilde{\psi}_n, \gamma_0) \right| & \leq \left| \varepsilon_t(\tilde{\psi}_n, \tilde{\gamma}_n) - \varepsilon_t(\psi_n, \gamma_0) \right| + |\varepsilon_t(\psi_n, \gamma_0)| \\
& \leq Cst \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_n, \gamma_0)| + Cst \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) + |\varepsilon_t(\psi_n, \gamma_0)| \\
& \leq Cst \left[\|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) + 1 \right] |\varepsilon_t(\psi_n, \gamma_0)| + Cst \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}).
\end{aligned} \tag{3.12}$$

Or,

$$|\varepsilon_t(\psi_n, \gamma_0)| \leq |\varepsilon_t(\psi_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)| + |\varepsilon_t(\psi_0, \gamma_0)|.$$

Une fois de plus, à l'aide de l'égalité (3.7), on a

$$\begin{aligned}
|\varepsilon_t(\psi_n, \gamma_0)| & \leq \frac{1}{\tau} \|\theta_0 - \theta_n\| \times \|\partial \sigma(Z_{t-1}, \tilde{\theta}_n)\| \times |\varepsilon_t(\psi_0, \gamma_0)| + \frac{1}{\tau} \|\rho_0 - \rho_n\| \\
& \times \|\partial m(Z_{t-1}, \tilde{\rho}_n)\| + |\varepsilon_t(\psi_0, \gamma_0)| \\
& \leq \frac{1}{\tau} \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| + \frac{1}{\tau} \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) + |\varepsilon_t(\psi_0, \gamma_0)|.
\end{aligned}$$

On majore $|\varepsilon_t(\psi_n, \gamma_0)|$ dans l'inégalité (3.12), on obtient

$$\begin{aligned}
 \left| \varepsilon_t(\tilde{\psi}_n, \gamma_0) \right| &\leq Cst \left[\left| \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) + 1 \right| \left[\left| \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right. \right. \right. \\
 &\quad \left. \left. + \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) + |\varepsilon_t(\psi_0, \gamma_0)| \right] + Cst \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) \right. \\
 &\leq Cst \left[\left| \|\theta_0 - \theta_n\| \times |\varepsilon_t(\psi_0, \gamma_0)| + \|\rho_0 - \rho_n\| \right| \vartheta(Z_{t-1}) \right. \\
 &\quad \left. + Cst \|\theta_0 - \theta_n\|^2 \vartheta^2(Z_{t-1}) \times |\varepsilon_t(\psi_0, \gamma_0)| \right. \\
 &\quad \left. + Cst \|\theta_0 - \theta_n\| \times \|\rho_0 - \rho_n\| \vartheta^2(Z_{t-1}) + Cst \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) \times |\varepsilon_t(\psi_0, \gamma_0)| \right. \\
 &\quad \left. + Cst |\varepsilon_t(\psi_0, \gamma_0)| \right]. \tag{3.13}
 \end{aligned}$$

On déduit des inégalités précédentes que

$$\begin{aligned}
 \left| \phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \varepsilon_t(\tilde{\psi}_n, \gamma_0) \right| &\leq Cst \|\rho_0 - \rho_n\| \times |\varepsilon_t(\psi_0, \gamma_0)| \vartheta(Z_{t-1}) \\
 &\quad + Cst \|\rho_0 - \rho_n\| \times \|\theta_0 - \theta_n\| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \\
 &\quad + Cst \|\rho_0 - \rho_n\|^2 \vartheta^2(Z_{t-1}) \\
 &\quad + Cst \|\theta_0 - \theta_n\|^3 \vartheta^3(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \\
 &\quad + Cst \|\theta_0 - \theta_n\|^2 \times \|\rho_0 - \rho_n\| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \\
 &\quad + Cst \|\theta_0 - \theta_n\|^2 \vartheta^2(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \\
 &\quad + Cst \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \\
 &\quad + Cst \|\rho_0 - \rho_n\|^2 \|\theta_0 - \theta_n\| \vartheta^3(Z_{t-1}) \\
 &\quad + Cst \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &\quad + Cst \|\theta_0 - \theta_n\|^2 \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &\quad + Cst \|\rho_0 - \rho_n\| \times \|\theta_0 - \theta_n\| \vartheta^2(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &\quad + Cst \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &\quad + Cst |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]|. \tag{3.14}
 \end{aligned}$$

Trouvons une majoration pour $\phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]$. On a :

$$\left| \phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right| \leq \left| \phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] - \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right| + |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|.$$

Puisque ϕ'_f est lipschitzienne, alors

$$\begin{aligned}
 \left| \phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right| &\leq Cst \left| \varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0) \right| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &\quad + \frac{1}{2} \left| \varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0) \right|^2 \left| \phi''_f[\tilde{\varepsilon}_t(\tilde{\psi}_n, \gamma_0)] \right| + |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|,
 \end{aligned}$$

où $\tilde{\varepsilon}_t(\tilde{\psi}_n, \gamma_0)$ est comprise entre $\varepsilon_t(\tilde{\psi}_n, \gamma_0)$ et $\varepsilon_t(\psi_0, \gamma_0)$.

En majorant $|\varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)|$ par l'utilisation de (3.7) et en utilisant le fait que ϕ''_f

est bornée on a :

$$\begin{aligned}
 |\phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]| &\leq Cst \|\theta_0 - \theta_n\| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &+ \|\rho_0 - \rho_n\| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
 &+ Cst \|\theta_0 - \theta_n\|^2 \vartheta^2(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \\
 &+ Cst \|\rho_0 - \rho_n\|^2 \vartheta^2(Z_{t-1}) + |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|. \tag{3.15}
 \end{aligned}$$

Ainsi, en utilisant les majorations ci-dessus, on a

$$\begin{aligned}
 |\Delta_{n3}| &\leq \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \left[\frac{4}{\tau^3} \vartheta^2(Z_{t-1}) + \frac{1}{\tau^2} \vartheta^2(Z_{t-1}) \right] |\phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]| \times |\varepsilon_t(\tilde{\psi}_n, \gamma_0)| \\
 &\leq Cst \|\rho_0 - \rho_n\| \times \|\theta_0 - \theta_n\|^2 \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^5(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right\} \\
 &+ Cst \|\rho_0 - \rho_n\|^2 \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) \right\} \\
 &+ Cst \|\rho_0 - \rho_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right\} \\
 &+ Cst \|\theta_0 - \theta_n\| \times \|\rho_0 - \rho_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right\} \\
 &+ Cst \|\theta_0 - \theta_n\|^3 \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^5(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \right\} \\
 &+ Cst \|\theta_0 - \theta_n\|^2 \|\rho_0 - \rho_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^5(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right\} \\
 &+ Cst \|\theta_0 - \theta_n\|^2 \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \right\} \\
 &+ Cst \|\theta_0 - \theta_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \right\} \\
 &+ Cst \|\rho_0 - \rho_n\|^2 \|\theta_0 - \theta_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^5(Z_{t-1}) \right\} \\
 &+ Cst \|\rho_0 - \rho_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\} \\
 &+ Cst \|\theta_0 - \theta_n\|^2 \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^5(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\}
 \end{aligned}$$

$$\begin{aligned}
& + Cst \|\rho_0 - \rho_n\| \times \|\theta_0 - \theta_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\} \\
& + Cst \|\theta_0 - \theta_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\} \\
& + Cst \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\}.
\end{aligned}$$

L'application du théorème ergodique à chacun des termes entre accolades dans le membre de droite de l'inégalité précédente montre que ces termes convergent presque sûrement vers des limites finies. Du fait que $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ convergent en probabilité vers 0 on a facilement que

$$\Delta_{n3} \xrightarrow{P} 0.$$

Pour étudier la convergence de Δ_{n4} , on utilise l'inégalité (3.13), on a

$$\begin{aligned}
|\Delta_{n4}| & \leq Cst \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t^2(\tilde{\psi}_n, \gamma_0)| \\
& \leq Cst \|\theta_0 - \theta_n\| \times \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) \\
& + Cst \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \\
& + Cst \|\theta_0 - \theta_n\|^2 \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \\
& + Cst \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \\
& + Cst \|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|.
\end{aligned}$$

Une fois de plus, on a par le théorème ergodique et le fait que $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ convergent en probabilité vers 0, que Δ_{n4} est majoré par une quantité qui converge en probabilité vers 0, ce qui implique que Δ_{n4} converge en probabilité vers 0 et est tendue.

Étudions à présent la convergence de Δ_{n5} . En procédant comme précédemment, on a

$$\begin{aligned}
|\Delta_{n5}| & \leq \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \left[\frac{2}{\tau^3} \vartheta^2(Z_{t-1}) + \frac{1}{\tau^2} \vartheta^2(Z_{t-1}) \right] \left| \phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right| \\
& \leq Cst \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) \left| \phi_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \right|.
\end{aligned}$$

En utilisant l'inégalité (3.15), on obtient,

$$\begin{aligned}
|\Delta_{n5}| &\leq Cst\|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
&+ Cst\|\theta_0 - \theta_n\|^2 \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \\
&+ \|\rho_0 - \rho_n\|^2 \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) + \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|.
\end{aligned}$$

Une fois de plus, en appliquant le théorème ergodique, du fait que $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ convergent en probabilité vers 0 et que $\int_{\mathbb{R}} \phi_f(x) f(x) dx = 0$, il s'en suit que Δ_{n5} tend vers aussi en probabilité vers 0. On conclut donc que $D_{n,2,2}(\tilde{\psi}_n, \gamma_0, \beta)$ converge en probabilité vers 0. Il est donc tendu.

Étude du comportement asymptotique de $D_{n,1,2}(\tilde{\psi}_n, \gamma_0, \beta)$

Commençons par une décomposition de $D_{n,1,2}(\tilde{\psi}_n, \gamma_0, \beta)$ en des sommes de termes dont nous étudions plus facilement le comportement asymptotique.

Pour tout $\psi \in \Theta \times \tilde{\Theta}$, on a

$$\begin{aligned}
\partial_{\rho\theta}^2 \Delta_n(\psi, \gamma_0, \beta) &= \frac{\partial}{\partial \rho} \left(\frac{\partial \Lambda_n(\psi, \gamma_0, \beta)}{\partial \theta} \right) \\
&= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \frac{\partial}{\partial \rho} \left\{ \phi'_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \right\} \\
&- \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)] \partial_\rho \varepsilon_t(\psi, \gamma_0) \\
&= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^2(Z_{t-1}, \theta)} \\
&\times \left\{ \phi''_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) \partial_\rho \varepsilon_t(\psi, \gamma_0) + \partial_\rho \varepsilon_t(\psi, \gamma_0) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \right\} \\
&+ \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta)}{\sigma^3(Z_{t-1}, \theta)} \partial_\rho m(Z_{t-1}, \rho) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \\
&= \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\rho m(Z_{t-1}, \rho)]^\top}{\sigma^3(Z_{t-1}, \theta)} \\
&\times \left\{ \phi''_f[\varepsilon_t(\psi, \gamma_0)] \varepsilon_t(\psi, \gamma_0) + \phi'_f[\varepsilon_t(\psi, \gamma_0)] \right\} \\
&+ \frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\rho m(Z_{t-1}, \rho)]^\top}{\sigma^3(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)]
\end{aligned}$$

$$\begin{aligned}
 &= \frac{2}{n} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\rho m(Z_{t-1}, \rho)]^\top}{\sigma^3(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)] \\
 &+ \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \theta) [\partial_\rho m(Z_{t-1}, \rho)]^\top}{\sigma^3(Z_{t-1}, \theta)} \varepsilon_t(\psi, \gamma_0) \phi''_f[\varepsilon_t(\psi, \gamma_0)].
 \end{aligned}$$

Donc, en multipliant cette égalité par $1/\sqrt{n}$ et en remplaçant ψ par $\tilde{\psi}_n$, il vient

$$\begin{aligned}
 D_{n,1,2}(\tilde{\psi}_n, \gamma_0, \beta) &= \frac{2}{n} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n) [\partial_\rho m(Z_{t-1}, \tilde{\rho}_n)]^\top}{\sigma^3(Z_{t-1}, \tilde{\theta}_n)} \phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \\
 &+ \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n) [\partial_\rho m(Z_{t-1}, \tilde{\rho}_n)]^\top}{\sigma^3(Z_{t-1}, \tilde{\theta}_n)} \varepsilon_t(\tilde{\psi}_n, \gamma_0) \phi''_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \\
 &= 2\Delta_{n6} + \Delta_{n7}.
 \end{aligned}$$

Étudions la convergence de chacun de ces termes.

Commençons par Δ_{n6} , en utilisant (3.11), on a :

$$\begin{aligned}
 |\Delta_{n6}| &\leq \frac{1}{\tau^3} \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\phi'_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)]| \right\} \\
 &\leq Cst \|\theta_0 - \theta_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right\} \\
 &+ Cst \|\rho_0 - \rho_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \right\} \\
 &+ Cst \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\}.
 \end{aligned}$$

Comme dans les cas précédents, il est facile de voir que par le théorème ergodique, les termes entre accolades dans le membre de droite de la dernière inégalité convergent presque sûrement vers une limite finie $c = \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} \int_{\mathbb{R}} |\phi'_f(x)| f(x) dx < \infty$. Étant donné que $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ convergent en probabilité vers 0, Δ_{n6} converge en probabilité vers une limite appartenant à l'intervalle $[-c, c]$.

Pour le terme Δ_{n7} , en utilisant l'inégalité (3.13), on a facilement les majorations suivantes :

$$|\Delta_{n7}| \leq Cst \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) \varepsilon_t(\tilde{\psi}_n, \gamma_0) \right)$$

$$\begin{aligned}
&\leq Cst \|\theta_0 - \theta_n\| \times \|\rho_0 - \rho_n\| \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) \right) \\
&+ Cst \|\rho_0 - \rho_n\| \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \right) \\
&+ Cst \|\theta_0 - \theta_n\| \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right) \\
&+ Cst \|\theta_0 - \theta_n\|^2 \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^4(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right) \\
&+ Cst \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right).
\end{aligned}$$

Toujours par des arguments analogues à ceux déjà utilisés précédemment, par le théorème ergodique et le fait que $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ convergent en probabilité vers 0, on peut voir que Δ_{n7} converge en probabilité vers 0. Donc en conclusion, $D_{n,1,2}(\tilde{\psi}_n, \gamma_0, \beta)$ converge en probabilité vers une limite finie appartenant à $[-d, d]$, où $d = 2c < \infty$.

Pour l'étude de $D_{n,2,1}(\tilde{\psi}_n, \gamma_0, \beta)$, on peut tout simplement observer que

$$D_{n,1,2}(\tilde{\psi}_n, \gamma_0, \beta) = \left(D_{n,2,1}(\tilde{\psi}_n, \gamma_0, \beta) \right)^\top.$$

En effet, pour tout $\psi \in \Theta \times \tilde{\Theta}$, on a

$$\begin{aligned}
\partial_{\theta\rho}^2 \Delta_n(\psi, \gamma_0, \beta) &= \frac{\partial}{\partial \theta} \left(\frac{\partial \Delta_n(\psi, \gamma_0, \beta)}{\partial \rho} \right) \\
&= \frac{\partial}{\partial \theta} \left\{ -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \partial_\rho m(Z_{t-1}, \rho) \phi'_f[\varepsilon_t(\psi, \gamma_0)] \right\} \\
&= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial}{\partial \theta} \left\{ \frac{\phi'_f[\varepsilon_t(\psi, \gamma_0)]}{\sigma^2(Z_{t-1}, \theta)} \right\} \partial_\rho m(Z_{t-1}, \rho) \\
&= \frac{1}{n} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\rho m(Z_{t-1}, \rho) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top}{\sigma^3(Z_{t-1}, \theta)} \varepsilon_t(\psi, \gamma_0) \phi''_f[\varepsilon_t(\psi, \gamma_0)] \\
&+ \frac{2}{n} \sum_{t=1}^n \beta^\top \omega(t) \frac{\partial_\rho m(Z_{t-1}, \rho) [\partial_\theta \sigma(Z_{t-1}, \theta)]^\top}{\sigma^3(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma_0)] \\
&= \left(\frac{\partial^2 \Delta_n(\psi, \gamma_0, \beta)}{\partial \rho \partial \theta} \right)^\top.
\end{aligned}$$

On obtient à la fin que $\left\| \partial_\psi^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta) \right\|_M / \sqrt{n}$ converge en probabilité vers une variable aléatoire positive. D'où,

$$(\psi_0 - \psi_n)^\top \partial_\psi^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta) (\psi_0 - \psi_n) = o_P(1).$$

Retour à la démonstration de la proposition. D'après ce qui précède on a :

$$\begin{aligned}\Delta_n(\psi_0, \gamma_0, \beta) &= \Delta_n(\psi_n, \gamma_0, \beta) + (\psi_0 - \psi_n)^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) + o_P(1) \\ &= \Delta_n(\psi_n, \gamma_0, \beta) + (\psi_0 - \psi_{N(n)})^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) \\ &\quad + (\psi_{N(n)} - \psi_n)^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) + o_P(1).\end{aligned}$$

On a facilement,

$$(\psi_0 - \psi_{N(n)})^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) = \sqrt{n}(\psi_0 - \psi_{N(n)})^\top \frac{1}{\sqrt{n}} \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta).$$

Or,

$$\sqrt{n}(\psi_0 - \psi_{N(n)})^\top = \sqrt{N(n)}(\psi_0 - \psi_{N(n)})^\top \frac{\sqrt{n}}{\sqrt{N(n)}} = o_P(1).$$

Montrons à présent que $\partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) / \sqrt{n}$ converge en probabilité vers un vecteur aléatoire.

Étude de $\partial_\rho \Delta_n(\psi_n, \gamma_0, \beta) / \sqrt{n}$

On a

$$\begin{aligned}\frac{1}{\sqrt{n}} \partial_\rho \Delta_n(\psi_n, \gamma_0, \beta) &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\rho m(Z_{t-1}, \rho_n) \phi'_f[\varepsilon_t(\psi_n, \tilde{\gamma}_n)] \\ &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\rho m(Z_{t-1}, \rho_n) \{ \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \} \\ &\quad - \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\rho m(Z_{t-1}, \rho_n) \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \\ &= \Delta_n^{(1)} - \Delta_n^{(2)},\end{aligned}$$

où

$$\begin{aligned}\Delta_n^{(1)} &:= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\rho m(Z_{t-1}, \rho_n) \{ \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \} \\ \Delta_n^{(2)} &:= \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\rho m(Z_{t-1}, \rho_n) \phi'_f[\varepsilon_t(\psi_0, \gamma_0)].\end{aligned}$$

Puisque ϕ'_f est lipschitzienne, alors en utilisant (3.7), on a

$$|\Delta_n^{(1)}| \leq \frac{c_\phi}{\tau^2} \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)|$$

$$\begin{aligned}
&\leq Cst\|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) \frac{\|\partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n)\|}{\sigma(Z_{t-1}, \theta_n)} |\varepsilon_t(\psi_0, \gamma_0)| \\
&+ Cst\|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) \frac{\|\partial_\rho m(Z_{t-1}, \tilde{\rho}_n)\|}{\sigma(Z_{t-1}, \theta_n)} \\
&\leq Cst\|\theta_0 - \theta_n\| \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \right) \\
&+ Cst\|\rho_0 - \rho_n\| \left(\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) \right) \\
&= Cst\|\theta_0 - \theta_n\| \Delta_n^{(11)} + Cst\|\rho_0 - \rho_n\| \Delta_n^{(12)},
\end{aligned}$$

où

$$\begin{aligned}
\Delta_n^{(11)} &:= \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|, \\
\Delta_n^{(12)} &:= \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}).
\end{aligned}$$

Le théorème ergodique entraîne que $\Delta_n^{(12)} \xrightarrow{p.s.} 0$ et

$$\Delta_n^{(12)} = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) \xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j}.$$

Par conséquent, lorsque $n \rightarrow \infty$,

$$\Delta_n^{(1)} \xrightarrow{P} 0.$$

Pour l'étude de la convergence de $\Delta_n^{(2)}$, on utilise une simple majoration. On a

$$\begin{aligned}
|\Delta_n^{(2)}| &= \left| \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\rho m(Z_{t-1}, \rho_n) \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \right| \\
&\leq \frac{1}{\tau^2 n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]|.
\end{aligned}$$

D'après théorème ergodique,

$$\begin{aligned}
\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\
&\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{1,j} \int_{\mathbb{R}} |\phi'_f(x)| f(x) dx < \infty.
\end{aligned}$$

Donc $\Delta_n^{(2)}$ converge en probabilité vers une variable aléatoire. Ce qui implique que $\partial_\rho \Delta_n(\psi_n, \gamma_n) / \sqrt{n}$ converge en probabilité vers une variable aléatoire.

Étude de $\partial_\theta \Delta_n(\psi_n, \gamma_0, \beta) / \sqrt{n}$

On a la décomposition ci-dessous :

$$\begin{aligned} \frac{1}{\sqrt{n}} \partial_\theta \Delta_n(\psi_n, \gamma_0, \beta) &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \\ &\quad - \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\theta \sigma(Z_{t-1}, \theta_n) \phi_f[\varepsilon_t(\psi_n, \gamma_0)] \\ &= \Delta_n^{(3)} - \Delta_n^{(4)}. \end{aligned}$$

Nous étudions la convergence de $\Delta_n^{(3)}$ et $\Delta_n^{(4)}$. Nous décomposons d'abord $\Delta_n^{(3)}$ en des sommes de termes faciles à étudier.

$$\begin{aligned} \Delta_n^{(3)} &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi'_f[\varepsilon_t(\psi_n, \tilde{\gamma}_n)] \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \\ &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \\ &\quad \times \{ \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] + \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \} \\ &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \{ \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \} \\ &\quad - \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \\ &= \Delta_n^{(31)} - \Delta_n^{(32)}, \end{aligned}$$

où

$$\begin{aligned} \Delta_n^{(31)} &:= -\frac{1}{n} \sum_{t=1}^n \left\{ \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \tilde{\gamma}_n) \{ \phi'_f[\varepsilon_t(\psi_n, \tilde{\gamma}_n)] - \phi'_f[\varepsilon_t(\psi_0, \tilde{\gamma}_n)] \} \right\}, \\ \Delta_n^{(32)} &:= \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \phi'_f[\varepsilon_t(\psi_0, \gamma_0)]. \end{aligned}$$

Commençons par l'étude de la convergence de $\Delta_n^{(31)}$.

Puisque ϕ'_f est lipschitzienne, alors

$$\begin{aligned} |\Delta_n^{(31)}| &\leq \frac{c_\phi}{\tau^2} \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \times |\varepsilon_t(\psi_n, \gamma_0)| \times |\varepsilon_t(\psi_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)| \vartheta(Z_{t-1}) \} \\ &\leq Cst \|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n \{ \|\beta^\top \omega(t)\| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_n, \gamma_0)| \times |\varepsilon_t(\psi_0, \gamma_0)| \} \end{aligned}$$

$$+ Cst \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_n, \gamma_0)|.$$

Ainsi, d'après l'inégalité (1.7) on a

$$\begin{aligned} |\Delta_n^{(31)}| &\leq Cst \|\theta_0 - \theta_n\|^2 \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|^2 \\ &+ Cst \|\rho_0 - \rho_n\| \times \|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \\ &+ Cst \|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \\ &+ Cst \|\rho_0 - \rho_n\|^2 \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \\ &+ Cst \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|. \end{aligned}$$

D'après théorème ergodique,

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^3(Z_{t-1}) \varepsilon_t^2(\psi_0, \gamma_0) \\ &\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} < \infty, \\ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^3(Z_{t-1}) \\ &\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{3,j} < \infty, \\ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|^2 &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|^2 \\ &\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} \mathbb{E}[\varepsilon_1(\psi_0, \gamma_0)^2] < \infty, \\ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^3(Z_{t-1}) \\ &\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{3,j} < \infty, \end{aligned}$$

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_j|}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \\ &\xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \chi_{2,j} \mathbb{E}[\varepsilon_t(\psi_0, \gamma_0)] < \infty. \end{aligned}$$

Et on sait que $\|\rho_0 - \rho_n\| \xrightarrow{P} 0$ et $\|\theta_0 - \theta_n\| \xrightarrow{P} 0$. Donc,

$$\Delta_n^{(31)} \xrightarrow{P} 0.$$

Étude du comportement asymptotique de $\Delta_n^{(32)}$

On a

$$\Delta_n^{(32)} = \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_{\theta\sigma}(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \phi'_f[\varepsilon_t(\psi_0, \gamma_0)].$$

Ce qui implique que

$$\begin{aligned} |\Delta_n^{(32)}| &\leq Cst \|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \} \\ &+ Cst \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\ &+ Cst \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \}. \end{aligned}$$

À l'aide du théorème ergodique et du fait que $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ convergent en probabilité vers 0, $\Delta_n^{(32)}$ converge en probabilité vers une variable aléatoire. Alors, $\Delta_n^{(3)}$ converge en probabilité vers une variable aléatoire.

Étude du comportement asymptotique de $\Delta_n^{(4)}$

Par un développement de Taylor à l'ordre 2 de ϕ_f au voisinage de point $\varepsilon_t(\psi_n, \gamma_0)$ on a,

$$\begin{aligned} \Delta_n^{(4)} &= \frac{1}{n} \sum_{t=1}^n \left\{ \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_{\theta\sigma}(Z_{t-1}, \theta_n) [\varepsilon_t(\psi_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)] \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \right\} \\ &+ \frac{1}{2n} \sum_{t=1}^n \left\{ \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_{\theta\sigma}(Z_{t-1}, \theta_n) [\varepsilon_t(\psi_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)]^2 \phi''_f[\tilde{\varepsilon}_t(\psi_n, \gamma_0)] \right\} \\ &+ \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_{\theta\sigma}(Z_{t-1}, \theta_n) \phi_f[\varepsilon_t(\psi_0, \gamma_0)], \end{aligned}$$

où $\tilde{\varepsilon}_t(\psi_n, \gamma_0)$ est comprise entre $\varepsilon_t(\psi_n, \gamma_0)$ et $\varepsilon_t(\psi_0, \gamma_0)$.

On se rappelant que pour tout $\theta \in \tilde{\Theta}$ et $z \in \mathbb{R}^p$, $\sigma(z, \theta) > \tau > 0$, on déduit que

$$\begin{aligned} |\Delta_n^{(4)}| &\leq \frac{1}{\tau^2} \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \} \\ &+ \frac{1}{2\tau^2} \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\varepsilon_t(\psi_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)|^2 |\phi''_f[\tilde{\varepsilon}_t(\psi_n, \gamma_0)]| \} \\ &+ \frac{1}{\tau^2} \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) \phi_f[\varepsilon_t(\psi_0, \gamma_0)]. \end{aligned}$$

En utilisant (3.7), on obtient la majoration suivante :

$$\begin{aligned} |\Delta_n^{(4)}| &\leq Cst \|\theta_0 - \theta_n\| \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \} \\ &+ Cst \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \\ &+ Cst \|\theta_0 - \theta_n\|^2 \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\phi''_f[\tilde{\varepsilon}_t(\psi_n, \tilde{\gamma}_n)]| \times |\varepsilon_t(\psi_0, \gamma_0)|^2 \} \\ &+ Cst \|\rho_0 - \rho_n\|^2 \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\phi''_f[\tilde{\varepsilon}_t(\psi_n, \tilde{\gamma}_n)]| \\ &+ \frac{1}{\tau^2} \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) \phi_f[\varepsilon_t(\psi_0, \gamma_0)]. \end{aligned}$$

Puisque ϕ''_f est bornée, on a la nouvelle majoration suivante :

$$\begin{aligned} |\Delta_n^{(4)}| &\leq Cst \|\gamma_0 - \gamma_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\} \\ &+ Cst \|\theta_0 - \theta_n\| \left\{ \frac{1}{n} \sum_{t=1}^n \{ |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \times |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \} \right\} \\ &+ Cst \|\rho_0 - \rho_n\| \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\phi'_f[\varepsilon_t(\psi_0, \gamma_0)]| \right\} \\ &+ Cst \|\theta_0 - \theta_n\|^2 \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)|^2 \right\} \\ &+ Cst \|\rho_0 - \rho_n\|^2 \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^3(Z_{t-1}) \right\} \\ &+ \frac{1}{\tau^2} \left\{ \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right\}. \end{aligned}$$

Une fois de plus, à l'aide du théorème ergodique et du fait que $\|\rho_0 - \rho_n\|$ et $\|\theta_0 - \theta_n\|$ convergent en probabilité vers 0, on a que $\Lambda_n^{(4)}$ converge en probabilité vers une variable

aléatoire. Par suite, $\partial_\theta \Delta_n(\psi_n, \gamma_0, \beta) / \sqrt{n}$ converge en probabilité vers une variable aléatoire, et donc $\partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) / \sqrt{n}$ converge en probabilité vers un vecteur aléatoire. Nous venons ainsi de prouver que

$$(\psi_0 - \psi_{N(n)})^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) = \sqrt{n}(\psi_0 - \psi_{N(n)})^\top \frac{1}{\sqrt{n}} \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) = o_P(1).$$

Ce qui a pour conséquence que

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_n, \gamma_0, \beta) + (\psi_{N(n)} - \psi_n)^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) + o_P(1).$$

Lemme 3.3.1. *Soit ψ_n un estimateur consistant et asymptotiquement normal de ψ_0 . Alors, $\psi_{N(n)}$ est asymptotiquement dans l'espace tangent à la courbe de $\Delta_n(\psi, \gamma_0, \beta)$ en ψ_n , Γ_n défini comme suit :*

$$\Gamma_n := \{x \in \mathbb{R}^{l+q}, \Delta_n(x, \gamma_0, \beta) = \Delta_n(\psi_n, \gamma_0, \beta) + (x - \psi_n) \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta)\}, \quad (3.16)$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$.

Démonstration. Par un développement de Taylor à l'ordre 2 de $\Delta_n(\psi, \gamma_0, \beta)$ au voisinage de point ψ_n on a,

$$\begin{aligned} \Delta_n(\psi_{N(n)}, \gamma_0, \beta) &= \Delta_n(\psi_n, \gamma_0, \beta) + (\psi_{N(n)} - \psi_n)^\top \partial_\psi \Delta_n(\psi_n, \gamma_0, \beta) \\ &\quad + \frac{1}{2} (\psi_{N(n)} - \psi_n)^\top \partial_\psi^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) (\psi_{N(n)} - \psi_n), \end{aligned}$$

où $\tilde{\psi}_{N(n)}$ est entre $\psi_{N(n)}$ et ψ_n .

Pour montrer que $\psi_{N(n)}$ est asymptotiquement dans Γ_n , il suffit de montrer que $(\psi_{N(n)} - \psi_n)^\top \partial_\psi^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) (\psi_{N(n)} - \psi_n) = o_P(1)$.

On a,

$$\begin{aligned} \sqrt{n}(\psi_{N(n)} - \psi_n) &= \sqrt{n}(\psi_{N(n)} - \psi_0) + \sqrt{n}(\psi_0 - \psi_n) \\ &= \sqrt{N(n)}(\psi_{N(n)} - \psi_0) \frac{\sqrt{n}}{\sqrt{N(n)}} + \sqrt{n}(\psi_0 - \psi_n) \\ &= o_P(1) + \sqrt{n}(\psi_0 - \psi_n). \end{aligned}$$

Ce implique que $\sqrt{n}(\psi_{N(n)} - \psi_n)$ converge en loi vers une loi normale. De plus,

$$\psi_{N(n)} - \psi_n = \psi_{N(n)} - \psi_0 + \psi_0 - \psi_n = o_P(1) + o_P(1) = o_P(1).$$

Donc pour montrer que $(\psi_{N(n)} - \psi_n)^\top \partial_\psi^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) (\psi_{N(n)} - \psi_n) = o_P(1)$, il suffit de montrer que $\partial_\psi^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) / \sqrt{n}$ converge en probabilité vers un vecteur aléatoire.

On a :

$$\frac{1}{\sqrt{n}} \partial_\psi^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) = \frac{1}{\sqrt{n}} \begin{pmatrix} \partial_\rho^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) & \partial_{\rho\theta}^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) \\ \partial_{\theta\rho}^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) & \partial_\theta^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) \end{pmatrix},$$

où $\tilde{\psi}_{N(n)}$ est entre $\psi_{N(n)}$ et ψ_n , $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ et ψ_n un estimateur asymptotiquement normal de ψ_0 .

On a facilement que

$$\tilde{\psi}_{N(n)} - \psi_0 = o_P(1).$$

Ainsi, comme on a montré précédemment que $\left\| \partial_{\psi}^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta) \right\|_M / \sqrt{n}$ converge en probabilité vers une variable aléatoire positive, où $\tilde{\psi}_n$ entre ψ_n et ψ_0 , on peut montrer également que $\left\| \partial_{\psi}^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) \right\|_M / \sqrt{n}$ converge en probabilité vers une variable aléatoire positive. Par conséquent,

$$(\psi_{N(n)} - \psi_n)^\top \partial_{\psi}^2 \Delta_n(\tilde{\psi}_{N(n)}, \gamma_0, \beta) (\psi_{N(n)} - \psi_n) = o_P(1).$$

□

D'après le lemme précédent, $\psi_{N(n)}$ est asymptotiquement dans Γ_n défini précédemment.

On prend donc $x = \psi_{N(n)}$. On a alors

$$\Delta_n(\psi_{N(n)}, \gamma_n) = \Delta_n(\psi_n, \gamma_0, \beta) + (\psi_{N(n)} - \psi_n)^\top \partial_{\psi} \Delta_n(\psi_n, \gamma_0, \beta) + o_P(1).$$

Donc

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_{N(n)}, \gamma_0, \beta) + o_P(1).$$

ii) Pour tout $\beta \in \mathbb{R}^{k+1}$, on a

$$\begin{aligned} & \hat{\mu}_n(\hat{\psi}_n, \gamma_0, \beta) - \mu(\psi_0, \gamma_0, \beta) \\ &= \sum_{j=1}^{k+1} \beta_j^2 \left[\hat{\alpha}_j \hat{\mu}_{j,2}(\hat{\psi}_n, \gamma_0) - \alpha_j \mu_{j,2}(\psi_0, \gamma_0) \right] \\ &= \sum_{j=1}^{k+1} \beta_j^2 \left[(\hat{\alpha}_j - \alpha_j) \hat{\mu}_{j,2}(\hat{\psi}_n, \gamma_0) + \alpha_j [\hat{\mu}_{j,2}(\hat{\psi}_n, \gamma_0) - \mu_{j,2}(\psi_0, \gamma_0)] \right] \\ &= \sum_{j=1}^{k+1} \beta_j^2 \left\{ (\hat{\alpha}_j - \alpha_j) [\hat{\mu}_{j,2}(\hat{\psi}_n, \gamma_0) - \mu_{j,2}(\psi_0, \gamma_0)] + (\hat{\alpha}_j - \alpha_j) \mu_{j,2}(\psi_0, \gamma_0) \right\} \\ &+ \sum_{j=1}^{k+1} \beta_j^2 \left\{ \alpha_j [\hat{\mu}_{j,2}(\hat{\psi}_n, \gamma_0) - \mu_{j,2}(\psi_0, \gamma_0)] \right\}. \end{aligned}$$

On a pour tout $j = 1, \dots, k+1$, $\hat{\alpha}_j - \alpha_j \rightarrow 0$ lorsque $n \rightarrow \infty$, et puisque la fonction $\sigma(\cdot, \cdot)$ est continue et majorée par une constante finie, d'après le théorème de convergence de Lebesgue, lorsque $n \rightarrow \infty$,

$$\hat{\mu}_n(\hat{\psi}_n, \gamma_0, \beta) - \mu(\psi_0, \gamma_0, \beta) \rightarrow 0.$$

□

Théorème 3.3.1. (*Optimalité*)

Supposons que les hypothèses (A') , (A_0) - (A_5) et (B_0) - (B_4) sont satisfaites. Alors,

- i. Sous $H_\beta^{(n)}$, pour un seuil de signification $\alpha \in (0, 1)$, la puissance asymptotique du test basé sur $\mathcal{T}_n(\psi_{N(n)}, \gamma_0, \beta)$ est $1 - \Phi(u_\alpha - \varpi(\mu(\psi_0, \gamma_0, \beta)))$, où Φ est la fonction de répartition de la loi normale centrée réduite et u_α son quantile d'ordre $1 - \alpha$.
- ii. Le test basé sur $\mathcal{T}_n(\psi_{N(n)}, \gamma_0, \beta)$ est localement asymptotiquement optimal.

Démonstration. On a

$$\mathcal{T}_n(\psi_{N(n)}, \gamma_0, \beta) = \frac{\Delta_n(\psi_{N(n)}, \gamma_0, \beta)}{\widehat{\varpi}_n(\psi_{N(n)}, \beta)}$$

et lorsque $n \rightarrow \infty$,

$$\widehat{\mu}_n(\psi_{N(n)}, \gamma_0, \beta) \longrightarrow \mu(\psi_0, \gamma_0, \beta).$$

D'après la Proposition 3.3.2, on a, sous H_0 ,

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_{N(n)}, \gamma_0, \beta) + o_P(1).$$

Donc, sous H_0 ,

$$\begin{aligned} \mathcal{T}_n(\psi_{N(n)}, \gamma_0, \beta) &= \frac{\Delta_n(\psi_0, \gamma_0, \beta)}{\widehat{\varpi}_n(\psi_{N(n)}, \gamma_0, \beta)} + \frac{1}{\widehat{\varpi}_n(\psi_{N(n)}, \gamma_0, \beta)} \times o_P(1) \\ &= \frac{\Delta_n(\psi_0, \gamma_0, \beta)}{\varpi(\psi_0, \gamma_0, \beta)} \times \frac{\varpi(\psi_0, \gamma_0, \beta)}{\widehat{\varpi}_n(\psi_{N(n)}, \gamma_0, \beta)} + \frac{1}{\widehat{\varpi}_n(\psi_{N(n)}, \gamma_0, \beta)} \times o_P(1) \end{aligned}$$

Puisque sous H_0 , lorsque $n \rightarrow \infty$, $\widehat{\varpi}_n(\psi_{N(n)}, \gamma_0, \beta) \longrightarrow \varpi(\psi_0, \gamma_0, \beta)$, par contiguïté, le résultat reste valable sous $H_\beta^{(n)}$. Par le Troisième lemme de Le Cam, sous $H_\beta^{(n)}$, lorsque $n \rightarrow \infty$,

$$\frac{\Delta_n(\psi_0, \gamma_0, \beta)}{\varpi(\psi_0, \gamma_0, \beta)} \xrightarrow{\text{loi}} \mathcal{N}(\varpi(\psi_0, \gamma_0, \beta), 1).$$

Ce qui conduit à conclure que sous $H_\beta^{(n)}$, lorsque $n \rightarrow \infty$,

$$\mathcal{T}_n(\psi_{N(n)}, \gamma_0, \gamma_0, \beta) \xrightarrow{\text{loi}} \mathcal{N}(\varpi(\psi_0, \gamma_0, \beta), 1),$$

et que la puissance locale asymptotique du test construit est conservée. Le remplacement de $\Delta_n(\psi_0, \gamma_0, \beta)$ par sa version estimée n'a pas d'effet sur cette puissance. Ainsi, les suites de tests $\mathcal{T}_n(\psi_{N(n)}, \gamma_0, \beta)$ et $\mathcal{T}_n(\psi_0, \gamma_0, \beta)$ sont localement asymptotiquement équivalentes, d'où l'optimalité du test. \square

3.3.2 Cas où γ_0 est inconnu

On peut d'abord remarquer que si $\widehat{\gamma}_{0,n}$ et $\widetilde{\gamma}_n$ les estimateurs du maximum de vraisemblance de γ_0 et de $\gamma_n = \gamma_0 + \frac{\beta}{\sqrt{n}}$ respectivement, alors, on a

$$\widetilde{\gamma}_{0,n} = \widehat{\gamma}_{0,n} + \frac{\beta}{\sqrt{n}}.$$

En effet, asymptotiquement, en probabilité : $\widetilde{\gamma}_n = \gamma_0 + \frac{\beta}{\sqrt{n}}$. Or asymptotiquement, en probabilité, on a $\widehat{\gamma}_{0,n} = \gamma_0$. D'où asymptotiquement, en probabilité, $\widetilde{\gamma}_n = \widehat{\gamma}_{0,n} + \frac{\beta}{\sqrt{n}}$.

Proposition 3.3.3. *Sous les hypothèses (A'), (A₀)-(A₅) et (B₀)-(B₄), pour tout estimateur consistant et asymptotiquement normal $\widehat{\gamma}_{0,n}$ de γ_0 , on a pour tout $n \geq 1$,*

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_0, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1),$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$.

Démonstration. Par un développement de Taylor à l'ordre 2 au point $\widehat{\gamma}_{0,n}$ on a,

$$\begin{aligned} \Delta_n(\psi_0, \gamma_0, \beta) &= \Delta_n(\psi_0, \widehat{\gamma}_{0,n}, \beta) - (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Delta_n(\psi_0, \widehat{\gamma}_{0,n}, \beta) \\ &\quad + \frac{1}{2} (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,n} - \gamma_0), \end{aligned}$$

où $\widetilde{\gamma}_{0,n}$ est comprise entre $\widehat{\gamma}_{0,n}$ et γ_0 .

On a

$$(\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta) \times (\gamma_0 - \widehat{\gamma}_{0,n}) = o_P(1).$$

En effet,

$$\begin{aligned} |(\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta) (\gamma_0 - \widehat{\gamma}_{0,n})| &\leq \|\sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n})\| \\ &\quad \times \frac{1}{\sqrt{n}} \|\partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta)\|_M \times \|(\gamma_0 - \widehat{\gamma}_{0,n})\|. \end{aligned}$$

Puisque $\sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n})$ converge en loi vers une loi normale et que $\gamma_0 - \widehat{\gamma}_{0,n}$ converge presque sûrement vers 0, alors, pour montrer que $(\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta) \times (\gamma_0 - \widehat{\gamma}_{0,n}) = o_P(1)$, il suffit de montrer que $\|\partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta)\|_M / \sqrt{n}$ converge en probabilité vers une variable aléatoire positive.

On a pour tout $(\psi, \gamma, \beta) \in \mathbb{R}^{l+q} \times \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$,

$$\Delta_n(\psi, \gamma, \beta) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma(Z_{t-1}, \theta)} \phi_f[\varepsilon_t(\psi, \gamma)].$$

D'où pour tout $(\psi, \gamma, \beta) \in \mathbb{R}^{l+q} \times \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$,

$$\partial_\gamma \Delta_n(\psi, \gamma, \beta) = -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta)} \phi'_f[\varepsilon_t(\psi, \gamma)]. \quad (3.17)$$

Ainsi, pour tout $(\psi, \gamma, \beta) \in \mathbb{R}^{l+q} \times \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$,

$$\partial_\gamma^2 \Delta_n(\psi, \gamma, \beta) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\omega(t) \omega(t)^\top \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta)} \phi''_f[\varepsilon_t(\psi, \gamma)]. \quad (3.18)$$

Par suite,

$$\frac{1}{\sqrt{n}} \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) = -\frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \omega(t)^\top \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta_0)} \phi''_f[\varepsilon_t(\psi_0, \tilde{\gamma}_{0,n})].$$

Puisque ϕ''_f est bornée, alors

$$\frac{1}{\sqrt{n}} \left\| \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) \right\|_M \leq Cst \frac{1}{n} \sum_{t=1}^n \frac{|\beta^\top \omega(t)|}{\sigma^3(Z_{t-1}, \theta_0)}.$$

En utilisant le théorème ergodique on obtient,

$$\frac{1}{n} \sum_{t=1}^n \frac{|\beta^\top \omega(t)|}{\sigma^3(Z_{t-1}, \theta_0)} \xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \int_{\mathbb{R}^p} \frac{1}{\sigma^3(x, \theta_0)} dF_j(x) < \infty.$$

Par conséquent, $\left\| \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) \right\|_M / \sqrt{n}$ converge p.s. vers une variable aléatoire positive.

Ce qui implique que $(\gamma_0 - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) (\gamma_0 - \hat{\gamma}_{0,n}) = o_P(1)$. Ainsi,

$$\begin{aligned} \Delta_n(\psi_0, \gamma_0, \beta) &= \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + (\gamma_0 - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + o_P(1) \\ &= \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + (\gamma_0 - \hat{\gamma}_{0,N})^\top \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + (\hat{\gamma}_{0,N} \\ &\quad - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + o_P(1). \end{aligned}$$

Nous commençons par étudier la convergence de $(\gamma_0 - \hat{\gamma}_{0,N})^\top \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta)$. Pour cela, on a

$$(\gamma_0 - \hat{\gamma}_{0,N})^\top \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) = \sqrt{n} (\gamma_0 - \hat{\gamma}_{0,N})^\top \frac{1}{\sqrt{n}} \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta).$$

Or,

$$\sqrt{n} (\gamma_0 - \hat{\gamma}_{0,N})^\top = \sqrt{N(n)} (\gamma_0 - \hat{\gamma}_{0,N})^\top \frac{\sqrt{n}}{\sqrt{N(n)}} = o_P(1).$$

Montrons à présent que $\partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) / \sqrt{n}$ converge en probabilité vers un vecteur aléatoire.

On a,

$$\begin{aligned} \frac{1}{\sqrt{n}} \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) &= -\frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \phi'_f[\varepsilon_t(\psi_0, \hat{\gamma}_{0,n})] \\ &= \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \\ &\quad - (\hat{\gamma}_{0,n} - \gamma_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta_0)} \phi''_f[\varepsilon_t(\psi_0, \tilde{\gamma}_{0,n})], \end{aligned}$$

où $\tilde{\gamma}_{0,n}$ est comprise entre γ_0 et $\hat{\gamma}_{0,n}$. Puisque ϕ''_f est bornée et que $\hat{\gamma}_{0,n} - \gamma_0$ converge presque sûrement vers 0, alors le théorème ergodique nous permet de conclure que

$$(\hat{\gamma}_{0,n} - \gamma_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta_0)} \phi''_f[\varepsilon_t(\psi_0, \tilde{\gamma}_{0,n})] \xrightarrow{p.s.} 0.$$

En utilisant encore une fois le théorème ergodique, on a que $\frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)]$ converge presque sûrement vers une variable aléatoire. Il en résulte que $\partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}) / \sqrt{n}$ converge en probabilité vers un vecteur aléatoire.

Par conséquent,

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + o_P(1).$$

Montrons maintenant que $\hat{\gamma}_{0,N}$ est asymptotiquement dans l'espace tangent à la courbe de $\Delta_n(\psi_0, \gamma, \beta)$ en $\hat{\gamma}_{0,n}$.

Par un développement de Taylor d'ordre 2 de $\Delta_n(\psi_0, \gamma, \beta)$ au point $\hat{\gamma}_{0,n}$ on a

$$\begin{aligned} \Delta_n(\psi_0, \hat{\gamma}_{0,N}, \beta) &= \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) + (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_0, \hat{\gamma}_{0,n}, \beta) \\ &\quad + \frac{1}{2} (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) \times (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n}), \end{aligned}$$

où $\tilde{\gamma}_{0,n}$ comprise entre $\hat{\gamma}_{0,n}$ et $\hat{\gamma}_{0,N}$.

Pour montrer que $\hat{\gamma}_{0,N}$ est asymptotiquement dans l'espace tangent à la courbe de $\Delta_n(\psi_0, \gamma, \beta)$ en $\hat{\gamma}_{0,n}$, il suffit de montrer que $(\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) \times (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n}) = o_P(1)$.

On a

$$\begin{aligned} &(\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) \times (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n}) \\ &= \sqrt{n} (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \frac{1}{\sqrt{n}} \partial_\gamma^2 \Delta_n(\psi_0, \tilde{\gamma}_{0,n}, \beta) \times (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n}). \end{aligned}$$

Or,

$$\sqrt{n} (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top = \sqrt{N(n)} (\hat{\gamma}_{0,N} - \gamma_0) \frac{\sqrt{n}}{\sqrt{N(n)}} + \sqrt{n} (\gamma_0 - \hat{\gamma}_{0,n}) = o_P(1) + \sqrt{n} (\gamma_0 - \hat{\gamma}_{0,n}).$$

Ce qui implique que $\sqrt{n}(\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top$ converge en loi vers une loi normale. Donc pour montrer que $(\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta)(\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n}) = o_P(1)$, il suffit de montrer comme précédemment que $\partial_\gamma^2 \Delta_n(\psi_0, \widetilde{\gamma}_{0,n}, \beta) / \sqrt{n}$ converge en probabilité vers un vecteur aléatoire. On aura ensuite que

$$\Delta_n(\psi_0, \widehat{\gamma}_{0,N}, \beta) = \Delta_n(\psi_0, \widehat{\gamma}_{0,n}, \beta) + (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_0, \widehat{\gamma}_{0,n}, \beta) + o_P(1).$$

Alors, on conclut que

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_0, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1),$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$. \square

Proposition 3.3.4. *Sous les hypothèses (A') , (A_0) - (A_5) et (B_0) - (B_4) , pour tout estimateur consistant et asymptotiquement normal $(\psi_n, \widehat{\gamma}_{0,n})$ de (ψ_0, γ_0) on a pour tout $n \geq 1$,*

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1),$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$.

Démonstration. La démonstration fait usage du lemme ci-dessous que nous énonçons et démontrons d'abord.

Lemme 3.3.2. *Sous les hypothèses (A') , (A_0) - (A_5) et (B_0) - (B_4) , sous H_0 , pour tout estimateurs consistant et asymptotiquement normal $\widehat{\gamma}_{0,n}$ de γ_0 , on a pour tout $n \geq 1$,*

$$\Delta_n(\psi_n, \gamma_0, \beta) = \Delta_n(\psi_n, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1),$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$.

Démonstration. Par un développement de Taylor à l'ordre 2 au point $\widehat{\gamma}_{0,n}$ on a,

$$\begin{aligned} \Delta_n(\psi_n, \gamma_0, \beta) &= \Delta_n(\psi_n, \widehat{\gamma}_{0,n}, \beta) - (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Delta_n(\psi_n, \widehat{\gamma}_{0,n}, \beta) \\ &\quad + \frac{1}{2} (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta) \times (\widehat{\gamma}_{0,n} - \gamma_0), \end{aligned}$$

où $\widetilde{\gamma}_{0,n}$ est comprise entre $\widehat{\gamma}_{0,n}$ et γ_0 .

On a $(\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta) \times (\gamma_0 - \widehat{\gamma}_{0,n}) = o_P(1)$. En effet,

$$\begin{aligned} &|(\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta) \times (\gamma_0 - \widehat{\gamma}_{0,n})| \\ &\leq \|\sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n})\| \times \frac{1}{\sqrt{n}} \|\partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta)\|_M \times \|(\gamma_0 - \widehat{\gamma}_{0,n})\|. \end{aligned}$$

Puisque $\sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n})$ converge en loi vers une loi normale et que $\gamma_0 - \widehat{\gamma}_{0,n}$ converge presque sûrement vers 0, alors, pour montrer que $(\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta)(\gamma_0 - \widehat{\gamma}_{0,n}) = o_P(1)$,

il suffit d'étudier la convergence de $\partial_\gamma^2 \Delta_n(\psi_n, \tilde{\gamma}_{0,n}, \beta) / \sqrt{n}$.

On a

$$\frac{1}{\sqrt{n}} \partial_\gamma^2 \Delta_n(\psi_n, \tilde{\gamma}_{0,n}, \beta) = \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \omega(t)^\top \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta_n)} \phi_f''[\varepsilon_t(\psi_n, \tilde{\gamma}_{0,n})].$$

Par l'hypothèse (B_0) , on a

$$\frac{1}{\sqrt{n}} \left\| \partial_\gamma^2 \Delta_n(\psi_n, \tilde{\gamma}_{0,n}, \beta) \right\|_M \leq Cst \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| = Cst \|\beta\| < \infty.$$

Par conséquent, $\left\| \partial_\gamma^2 \Delta_n(\psi_n, \tilde{\gamma}_{0,n}, \beta) \right\|_M / \sqrt{n}$ converge presque sûrement vers une variable aléatoire positive. Ce qui implique que $(\gamma_0 - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \tilde{\gamma}_{0,n}, \beta) (\gamma_0 - \hat{\gamma}_{0,n}) = o_P(1)$. Ainsi,

$$\begin{aligned} \Delta_n(\psi_n, \gamma_0, \beta) &= \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) + (\gamma_0 - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) + o_P(1) \\ &= \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) + (\gamma_0 - \hat{\gamma}_{0,N})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) \\ &\quad + (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) + o_P(1). \end{aligned}$$

On commence par étudier la convergence de $(\gamma_0 - \hat{\gamma}_{0,N})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta)$. On a

$$(\gamma_0 - \hat{\gamma}_{0,N})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) = \sqrt{n} (\gamma_0 - \hat{\gamma}_{0,N})^\top \frac{1}{\sqrt{n}} \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta).$$

Or,

$$\sqrt{n} (\gamma_0 - \hat{\gamma}_{0,N})^\top = \sqrt{N(n)} (\gamma_0 - \hat{\gamma}_{0,N})^\top \frac{\sqrt{n}}{\sqrt{N(n)}} = o_P(1).$$

Montrons à présent que $\partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) / \sqrt{n}$ converge en probabilité vers un vecteur aléatoire.

On a,

$$\begin{aligned} \frac{1}{\sqrt{n}} \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t) \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi_f'[\varepsilon_t(\psi_n, \hat{\gamma}_{0,n})] \\ &= \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi_f'[\varepsilon_t(\psi_n, \gamma_0)] \\ &\quad - (\hat{\gamma}_{0,n} - \gamma_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta_n)} \phi_f''[\varepsilon_t(\psi_n, \tilde{\gamma}_{0,n})], \end{aligned}$$

où $\tilde{\gamma}_{0,n}$ est comprise entre γ_0 et $\hat{\gamma}_{0,n}$.

Sous les hypothèses (B_0) et (B_4) et du fait que $\hat{\gamma}_{0,n} - \gamma_0$ converge presque sûrement vers 0, le théorème ergodique permet de conclure que

$$(\hat{\gamma}_{0,n} - \gamma_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta_n)} \phi_f''[\varepsilon_t(\psi_n, \tilde{\gamma}_{0,n})] \xrightarrow{p.s.} 0.$$

Maintenant, on a

$$\begin{aligned}
 \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] &= \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \{ \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \} \\
 &- \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \\
 &= (\psi_n - \psi_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi''_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \partial_{\psi} \varepsilon_t(\tilde{\psi}_n, \gamma_0) \\
 &- \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \\
 &= (\theta_n - \theta_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi''_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \partial_{\theta} \varepsilon_t(\tilde{\psi}_n, \gamma_0) \\
 &+ (\rho_n - \rho_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi''_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \partial_{\rho} \varepsilon_t(\tilde{\psi}_n, \gamma_0) \\
 &- \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \\
 &= A_{1n} + A_{2n} - A_{3n}.
 \end{aligned}$$

Sous l'hypothèse (B_0) , en utilisant le théorème ergodique, on peut montrer facilement que A_{3n} converge presque sûrement vers une variable aléatoire.

Pour l'étude de A_{2n} , on peut écrire

$$A_{2n} = (\rho_n - \rho_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi''_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \frac{\partial_{\rho} m(Z_{t-1}, \tilde{\rho}_n)}{\sigma(Z_{t-1}, \tilde{\theta}_n)}.$$

Ce qui implique que

$$|A_{2n}| \leq Cst \|\rho_n - \rho_0\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}).$$

Or, par le théorème ergodique et du fait que $\psi_n - \psi_0 = o_P(1)$, on a

$$\|\rho_n - \rho_0\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta(Z_{t-1}) \xrightarrow{p.s.} 0.$$

Par conséquent,

$$A_{2n} \xrightarrow{p.s.} 0.$$

Pour étudier le terme A_{1n} , on peut écrire le développement

$$A_{1n} = -(\theta_n - \theta_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi''_f[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \varepsilon_t(\tilde{\psi}_n, \gamma_0) \partial_{\theta} \sigma(Z_{t-1}, \tilde{\theta}_n) \quad (3.19)$$

$$\begin{aligned}
&= -(\theta_n - \theta_0) \frac{1}{n} \sum_{t=1}^n \left\{ \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi_f''[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n) [\varepsilon_t(\tilde{\psi}_n, \gamma_0) - \varepsilon_t(\psi_0, \gamma_0)] \right\} \\
&- (\theta_n - \theta_0) \frac{1}{n} \sum_{t=1}^n \left\{ \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi_f''[\varepsilon_t(\tilde{\psi}_n, \gamma_0)] \partial_\theta \sigma(Z_{t-1}, \tilde{\theta}_n) \varepsilon_t(\psi_0, \gamma_0) \right\} \\
&= A_{1n1} + A_{1n2}.
\end{aligned}$$

En utilisant l'égalité (3.7), il est clair que le terme A_{1n1} est majoré par

$$\begin{aligned}
&Cst \|\theta_n - \theta_0\| \times \|\theta_0 - \theta_n\|^2 \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}) |\varepsilon_t(\psi_0, \gamma_0)| \\
&+ Cst \|\theta_n - \theta_0\| \times \|\rho_0 - \rho_n\| \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \vartheta^2(Z_{t-1}).
\end{aligned}$$

En utilisant encore une fois le théorème ergodique et du fait que $\|\theta_0 - \theta_n\|$ et que $\|\rho_0 - \rho_n\|$ convergent presque sûrement vers 0, il s'en suit que A_{1n1} converge presque sûrement vers 0.

De même, par une simple majoration et par le théorème ergodique, A_{1n2} converge presque sûrement vers 0. Il en découle que

$$A_{1n} \xrightarrow{p.s.} 0.$$

On conclut que $\sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi_f'[\varepsilon_t(\psi_n, \gamma_0)] / \sqrt{n}$ converge en probabilité vers une limite finie. Ce qui implique que $\partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) / \sqrt{n}$ converge en probabilité vers une variable aléatoire.

D'où

$$(\gamma_0 - \hat{\gamma}_{0,N})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) = o_P(1).$$

Ainsi,

$$\Delta_n(\psi_n, \gamma_0, \beta) = \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) + (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) + o_P(1).$$

Montrons à présent que $\hat{\gamma}_{0,N}$ est asymptotiquement dans l'espace tangent à la courbe de $\Delta_n(\psi_n, \gamma, \beta)$ en $\hat{\gamma}_{0,n}$.

Par un développement de Taylor d'ordre 2 de $\Delta_n(\psi_n, \gamma, \beta)$ au point $\hat{\gamma}_{0,n}$ on a,

$$\begin{aligned}
\Delta_n(\psi_n, \hat{\gamma}_{0,N}, \beta) &= \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) + (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_n, \hat{\gamma}_{0,n}, \beta) \\
&+ \frac{1}{2} (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \tilde{\gamma}_{0,n}, \beta) (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n}),
\end{aligned}$$

où $\tilde{\gamma}_{0,n}$ comprise entre $\hat{\gamma}_{0,n}$ et $\hat{\gamma}_{0,N}$.

Pour montrer que $\hat{\gamma}_{0,N}$ est asymptotiquement dans l'espace tangent à la courbe de $\Delta_n(\psi_n, \gamma, \beta)$ en $\hat{\gamma}_{0,n}$, il suffit de montrer que

$$(\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \tilde{\gamma}_{0,n}, \beta) (\hat{\gamma}_{0,N} - \hat{\gamma}_{0,n}) = o_P(1).$$

On a

$$\begin{aligned} & (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n}) \\ &= \sqrt{n} (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top \frac{1}{\sqrt{n}} \partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n}). \end{aligned}$$

Or,

$$\sqrt{n} (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top = \sqrt{N(n)} (\widehat{\gamma}_{0,N} - \gamma_0) \frac{\sqrt{n}}{\sqrt{N(n)}} + \sqrt{n} (\gamma_0 - \widehat{\gamma}_{0,n}) = o_P(1) + \sqrt{n} (\gamma_0 - \widehat{\gamma}_{0,n}).$$

Ce qui implique que $\sqrt{n} (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top$ converge en loi vers une loi normale. Donc pour montrer que $(\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n}) = o_P(1)$, il suffit de montrer comme précédemment que $\partial_\gamma^2 \Delta_n(\psi_n, \widetilde{\gamma}_{0,n}, \beta) / \sqrt{n}$ converge en probabilité vers une variable aléatoire. On aura ensuite que

$$\Delta_n(\psi_n, \widehat{\gamma}_{0,N}, \beta) = \Delta_n(\psi_n, \widehat{\gamma}_{0,n}, \beta) + (\widehat{\gamma}_{0,N} - \widehat{\gamma}_{0,n})^\top \partial_\gamma \Delta_n(\psi_n, \widehat{\gamma}_{0,n}, \beta) + o_P(1).$$

Alors, on conclut que

$$\Delta_n(\psi_n, \gamma_0, \beta) = \Delta_n(\psi_n, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1),$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$. \square

Revenons maintenant à la preuve de la proposition. On a, d'après le lemme précédent, pour tout $n \geq 1$,

$$\Delta_n(\psi_n, \gamma_0, \beta) = \Delta_n(\psi_n, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1).$$

En particulier pour $\psi_{N(n)}$, on a

$$\Delta_n(\psi_{N(n)}, \gamma_0, \beta) = \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1).$$

Ainsi,

$$\begin{aligned} & \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) - \Delta_n(\psi_0, \gamma_0, \beta) \\ &= \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) - \Delta_n(\psi_{N(n)}, \gamma_0, \beta) + \Delta_n(\psi_{N(n)}, \gamma_0, \beta) - \Delta_n(\psi_0, \gamma_0, \beta) \\ &= \Delta_n(\psi_{N(n)}, \gamma_0, \beta) - \Delta_n(\psi_0, \gamma_0, \beta) + o_P(1). \end{aligned}$$

Or, d'après la Proposition 3.3.4, on a

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_{N(n)}, \gamma_0, \beta) + o_P(1).$$

On en déduit que

$$\Delta_n(\psi_0, \gamma_0, \beta) = \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) + o_P(1).$$

\square

Soit $N(n)$ une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$ lorsque $n \rightarrow \infty$. Notre statistique du test pour ce problème est la suivante :

$$\mathcal{T}_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) = \frac{\Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)}{\widehat{\varpi}_n(\psi_{N(n)}, \beta)}.$$

Théorème 3.3.2. (*Optimalité*) *Supposons que les conditions (A'), (A₀)-(A₅) et (B₀)-(B₄) sont satisfaites. Alors,*

- i. Sous $H_\beta^{(n)}$, pour un seuil de signification $\alpha \in (0, 1)$, la puissance locale asymptotique du test basé sur $\mathcal{T}_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)$ est égale à $1 - \Phi(u_\alpha - \varpi(\psi_0, \gamma_0, \beta))$, où Φ est la densité d'une loi normale centrée réduite et u_α son quantile d'ordre $(1 - \alpha)$.*
- ii. Le test basé sur $\mathcal{T}_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)$ est localement asymptotiquement optimal.*

Démonstration. En utilisant la Proposition 3.3.4, la preuve est similaire à celle du Théorème 3.3.1. □

3.4 Conclusion

Dans ce chapitre, nous avons montré théoriquement l'optimalité asymptotique locale du test construit lorsque les fonctions T et V sont connues et lorsqu'elles sont connues mais dépendent de paramètres inconnus. Dans ce dernier cas, les paramètres sont estimés et la démonstration est basée essentiellement sur l'équivalence de la suite centrale et sa version estimée. Pour chacun des cas étudiés, un résultat de normalité asymptotique locale est établi et l'expression explicite de la puissance locale du test construit est donnée. Les résultats de ce chapitre sont illustrés par des simulations et sont appliqués à des données réelles. Les résultats sont présentés et commentés au chapitre 5.

Chapitre 4

Détection des ruptures et leurs localisations par un test du type Cramér-Von-Mises

4.1 Introduction

Dans ce chapitre, nous considérons le modèle (3.1) du chapitre précédent. Pour $\gamma_0 \in \mathbb{R}^{k+1}$, le problème étudié ici est celui de tester

$$H_0 : \gamma = \gamma_0 \text{ contre } H^{(n)} = \bigcup_{\beta \in \mathbb{R}^{k+1}} H_{\beta}^{(n)},$$

où $H_{\beta}^{(n)} : \gamma = \gamma_0 + \beta/\sqrt{n} = \gamma_n$, $n > 1$, $\gamma_0 \in \mathbb{R}^{k+1}$ et $\beta \in \mathbb{R}^{k+1}$ dépend des instants de rupture $t^k = (t_1, \dots, t_k)$.

En d'autres termes, on teste un modèle CHARN contre un ensemble de modèles CHARN, contrairement au chapitre précédent où l'on testait un modèle CHARN contre un autre modèle CHARN. Pour cette étude, nous considérons un test du type Cramer-Von-Mises basé sur le processus de vraisemblance. Ce type de test est étudié par exemple dans Csörgő (1983), Henze et Wagner (1997), Meintanis *et al.* (2015) et Ngatchou-Wandji (2009).

Dans ces articles, plusieurs techniques sont proposées pour trouver la loi asymptotique du test sous l'hypothèse alternative. Dans notre étude, nous utilisons essentiellement les techniques de Ngatchou-Wandji (2009). Nous considérons le cas où les fonctions T et V sont de formes connues mais dépendent de paramètres inconnus. Nous traitons le cas où le paramètre γ_0 est connu et le cas où il est inconnu. Pour cela, nous commençons d'abord à étudier la convergence faible du processus de statistique centrale dans le paragraphe 4.3 après avoir présenté la méthode de la détection de rupture.

4.2 Détection des ruptures

Dans ce chapitre, nous ne trouvons pas l'expression explicite de la puissance du test comme dans le chapitre précédent. Ici nous trouvons que sous l'hypothèse alternative la statistique de test construite converge en loi vers une somme pondérée de v.a indépendantes du loi Khi-deux décentrée.

Dans la littérature, il existe plusieurs méthodes numériques pour estimer cette dernière loi. Nous utilisons la méthode proposée par Imhof (1961) pour estimer les quantiles de cette loi. Ainsi, nous obtenons la région de rejet du test. Plus particulièrement, on note \mathcal{F} la fonction de répartition de $Q = \sum_{i=1}^J \lambda_i \chi_{\lambda_i}^2$, où pour tout $i = 1, \dots, J$, $\chi_{\lambda_i}^2$ suit une loi Khi-deux de paramètre de décentralisation λ_i et c_α est le quantile à chercher. Nous prenons une séquence de 0 à 20 au pas de 0.1 et nous calculons \mathcal{F} en tout point de cette séquence.

Nous nous arrêtons dès que nous trouvons un réel x tel que $|\mathcal{F}(x) - \alpha| \simeq 0$. Alors c_α est le réel x trouvé. Nous faisons cette approximation sur plusieurs jeux des données. Puis, nous prenons la moyenne comme estimateur de c_α . Avec ceci, nous pouvons estimer la puissance locale du test.

Ensuite, pour connaître la localisation du point de rupture, nous procédons de la même façon que le chapitre précédent en utilisant la puissance approximée numériquement.

4.3 Convergence faible du processus de statistique centrale

On étudie la convergence faible du processus de statistique centrale indexé par β . Cette étude est faite dans un espace de Fréchet. Ici, l'espace $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$ des fonctions définies sur \mathbb{R}^{k+1} à valeurs dans \mathbb{R} , continues, muni de la métrique ρ définie pour tout, $x, y \in \mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$ par

$$\rho(x, y) = \sum_{j=1}^{\infty} 2^{-j} \frac{\rho_j(x, y)}{1 + \rho_j(x, y)},$$

où pour tout $j \geq 1$,

$$\rho_j(x, y) = \sup_{\|t\| < j} |x(t) - y(t)|$$

est la norme sup usuelle.

Avec cette métrique, $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$ est un espace de Fréchet séparable, et la convergence dans cet espace correspond aux convergences uniformes sur tout ensemble compact : $\lim_{n \rightarrow \infty} \rho(x_n, x) = 0$ si et seulement si pour tout $j \geq 1$, $\lim_{n \rightarrow \infty} \rho_j(x_n, x) = 0$. Ainsi, $\lim_{n \rightarrow \infty} \rho(X_n, Y_n) = 0$ si et seulement si pour tout $j \geq 1$, $\lim_{n \rightarrow \infty} \rho_j(X_n, Y_n) = 0$, où X_n

et Y_n sont deux variables aléatoires dans l'espace $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$.

Soit $\Delta_n = \{\Delta_n(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$ et $\Delta = \{\Delta(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$, où pour tout $\beta \in \mathbb{R}^{k+1}$,

$$\Delta_n(\psi_0, \gamma_0, \beta) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma(Z_{t-1}, \theta)} \phi_f[\varepsilon_t(\psi, \gamma)],$$

et Δ un processus gaussien de moyenne nulle et de noyau de covariance K défini pour tout $\beta_1 = (\beta_{1,1}, \dots, \beta_{1,k+1}) \in \mathbb{R}^{k+1}$ et $\beta_2 = (\beta_{2,1}, \dots, \beta_{2,k+1}) \in \mathbb{R}^{k+1}$ par

$$K(\psi_0, \beta_1, \beta_2) = I(f) \sum_{j=1}^{k+1} \alpha_j \beta_{1,j} \beta_{2,j} \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_t(x) = I(f) \langle \alpha \odot \varrho, \beta_1 \odot \beta_2 \rangle, \quad (4.1)$$

où $\alpha = (\alpha_1, \dots, \alpha_{k+1})$, $\varrho = \left(\int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_1(x), \dots, \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_{k+1}(x) \right)^\top$ et pour tout $\psi = (\rho^\top, \theta^\top)^\top \in \mathbb{R}^{l+q}$, $\gamma \times \mathbb{R}^{k+1}$,

$$\varepsilon_t(\psi, \gamma) = \frac{X_t - m(Z_{t-1}, \rho) - \gamma^\top \omega(t)}{\sigma(Z_{t-1}, \theta)}, \quad t \in \mathbb{Z}.$$

4.3.1 Cas où γ_0 est connue

Proposition 4.3.1. *Sous les hypothèses (A'), (A₀)-(A₅) et (B₀)-(B₄) du chapitre précédent,*

- i. *Sous H_0 , $\{\Delta_n(\psi_0, \gamma_0, \cdot), n \geq 1\}$ converge faiblement vers le processus $\Delta(\psi_0, \gamma_0, \cdot)$ dans $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$.*
- ii. *Sous $H^{(n)}$, $\{\Delta_n(\psi_0, \gamma_0, \cdot), n \geq 1\}$ converge faiblement vers le processus $\Delta(\psi_0, \gamma_0, \cdot) + C(\cdot)$ dans $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$,*

où C la fonction définie pour tout $\beta \in \mathbb{R}^{k+1}$ par

$$C(\beta) = I(f) \sum_{j=1}^{k+1} \alpha_j \beta_j^2 \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_j(x) = I(f) \langle \alpha \odot \varrho, \beta \odot \beta \rangle,$$

avec $\beta = (\beta_1, \dots, \beta_{k+1})$.

Démonstration. i) Soit $(\psi, \gamma) \in \mathbb{R}^{l+q} \times \mathbb{R}^{k+1}$, on étudie les lois finies-dimensionnelles et la tension du processus $\{\Delta_n(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$.

Lemme 4.3.1. (Lois finies-dimensionnelles)

Soit $d \in \mathbb{N}^*$ et $\beta_1, \beta_2, \dots, \beta_d \in \mathbb{R}^{k+1}$. Sous (A'), (A₀)-(A₅) et (B₀)-(B₄), sous H_0 , lorsque $n \rightarrow \infty$,

$$(\Delta_n(\psi_0, \gamma_0, \beta_1), \Delta_n(\psi_0, \gamma_0, \beta_2), \dots, \Delta_n(\psi_0, \gamma_0, \beta_d))$$

converge en loi vers le vecteur aléatoire $(\Delta(\psi_0, \gamma_0, \beta_1), \dots, \Delta(\psi_0, \gamma_0, \beta_d))$, où $\{\Delta(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$ est un processus gaussien de moyenne nulle et de noyau de covariance K défini par (4.1).

Démonstration. Soit $(\delta_1, \delta_2, \dots, \delta_d) \in \mathbb{R}^d$, montrons que lorsque $n \rightarrow \infty$,

$$A_n = \sum_{i=1}^d \delta_i \Delta_n(\psi_0, \gamma_0, \beta_i)$$

converge en loi vers une variable aléatoire gaussienne.

On a pour tout $i = 1, \dots, d$,

$$\Delta_n(\psi_0, \gamma_0, \beta_i) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)].$$

On va utiliser le Corollaire 3.1 de Hall et Heyde (1980). Pour cela, on a besoin d'abord de vérifier ses conditions. On a,

$$\begin{aligned} A_n &= \sum_{i=1}^d \delta_i \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \sum_{i=1}^d \delta_i \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)]. \end{aligned}$$

On définit, pour tout $t = 1, \dots, n$ et pour tout $j = 1, \dots, n$,

$$C_{n,t} = \frac{1}{\sqrt{n}} \sum_{i=1}^d \delta_i \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)],$$

$$B_{n,j} = \sum_{t=1}^j C_{n,t}.$$

Il est facile de voir que pour tout $j = 1, \dots, n$, $\mathbb{E}(B_{n,j}) = 0$. Nous montrons à présent que $\{(B_{n,j}, G_j), j = 1, \dots, n\}$, où $G_{t-1} = \sigma(Z_1, \dots, Z_t)$ est la σ -algèbre engendrée par Z_1, \dots, Z_t , $t \in \mathbb{Z}$, est une suite de martingales.

En effet, soient $j_1, j_2 \in \mathbb{R}$ tels que $j_1 < j_2$ on a :

$$\begin{aligned} \mathbb{E}(B_{n,j_2} | G_{j_1}) &= \mathbb{E}\left(\sum_{t=1}^{j_2} C_{n,t} | G_{j_1}\right) \\ &= \mathbb{E}\left(\sum_{t=1}^{j_1} C_{n,t} | G_{j_1}\right) + \mathbb{E}\left(\sum_{t=j_1+1}^{j_2} C_{n,t} | G_{j_1}\right). \end{aligned}$$

Puisque $G_{j_1} = \sigma(Z_1, \dots, Z_{j_1})$, alors pour tout $t = 1, \dots, j_1$, $C_{n,t}$ est G_{j_1} -mesurable. Ainsi,

$$\begin{aligned} \mathbb{E}\left(B_{n,j_2} \middle| G_{j_1}\right) &= \sum_{t=1}^{j_1} C_{n,t} + \sum_{t=j_1+1}^{j_2} \mathbb{E}(C_{n,t} | G_{j_1}) \\ &= B_{n,j_1} + \frac{1}{\sqrt{n}} \sum_{t=j_1+1}^{j_2} \sum_{i=1}^d \mathbb{E} \left\{ \delta_i \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \middle| G_{j_1} \right\} \\ &= B_{n,j_1} + \frac{1}{\sqrt{n}} \sum_{t=j_1+1}^{j_2} \sum_{i=1}^d \delta_i \mathbb{E} \left\{ \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \middle| G_{j_1} \right\} \mathbb{E} \left\{ \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \middle| G_{j_1} \right\}. \quad (4.2) \end{aligned}$$

Or, pour tout $i = 1, \dots, d$, ε_i indépendant de G_{i-1} . Donc

$$\mathbb{E}\left(B_{n,j_2} \middle| G_{j_1}\right) = B_{n,j_1} + \frac{1}{\sqrt{n}} \sum_{t=j_1+1}^{j_2} \sum_{i=1}^d \delta_i \beta_i^\top \omega(t) \mathbb{E} \left\{ \frac{1}{\sigma(Z_{t-1}, \theta_0)} \middle| G_{j_1} \right\} \mathbb{E} \left\{ \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right\}.$$

D'après la remarque 1, pour tout $t \in \mathbb{Z}$, pour tout $i = 1, \dots, d$, $\mathbb{E} \left\{ \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right\} = 0$. Ce qui implique que le deuxième terme de l'égalité (4.2) est nul. Il s'ensuit que

$$\mathbb{E}\left(B_{n,j_2} \middle| G_{j_1}\right) = B_{n,j_1}.$$

On a aussi

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(C_{n,t}^2 | G_{t-1}) &= \sum_{t=1}^n \mathbb{E} \left\{ \left[\frac{1}{\sqrt{n}} \sum_{i=1}^d \delta_i \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right]^2 \middle| G_{t-1} \right\} \\ &= \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left\{ \sum_{i=1}^d \delta_i^2 \frac{[\beta_i^\top \omega(t)]^2}{\sigma^2(Z_{t-1}, \theta_0)} \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \middle| G_{t-1} \right\} \\ &\quad + \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left\{ \sum_{i,j=1, i \neq j}^d \delta_i \delta_j \frac{[\beta_i^\top \omega(t)][\beta_j^\top \omega(t)]}{\sigma^2(Z_{t-1}, \theta_0)} \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \middle| G_{t-1} \right\} \\ &= \frac{1}{n} \sum_{t=1}^n \sum_{i=1}^d \delta_i^2 \frac{[\beta_i^\top \omega(t)]^2}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \middle| G_{t-1} \right\} \\ &\quad + \frac{1}{n} \sum_{t=1}^n \sum_{i,j=1, i \neq j}^d \delta_i \delta_j \frac{\beta_i^\top \omega(t) \beta_j^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \middle| G_{t-1} \right\}. \end{aligned}$$

Puisque pour tout $i = 1, \dots, d$, ε_i indépendant de G_{i-1} . Alors,

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(C_{n,t}^2 | G_{t-1}) &= \frac{1}{n} \sum_{t=1}^n \sum_{i=1}^d \delta_i^2 \frac{[\beta_i^\top \omega(t)]^2}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \right\} \\ &\quad + \frac{1}{n} \sum_{t=1}^n \sum_{i,j=1, i \neq j}^d \delta_i \delta_j \frac{\beta_i^\top \omega(t) \beta_j^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \right\} \end{aligned}$$

$$\begin{aligned}
 &= \sum_{i=1}^d \delta_i^2 \frac{1}{n} \sum_{t=1}^n \frac{[\beta_i^\top \omega(t)]^2}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \} \\
 &+ \sum_{i,j=1, i \neq j}^d \delta_i \delta_j \frac{1}{n} \sum_{t=1}^n \frac{\beta_i^\top \omega(t) \beta_j^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \} \\
 &= \sum_{i=1}^d \delta_i^2 \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{\beta_{ij}^2}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \} \\
 &+ \sum_{i,j=1, i \neq j}^d \delta_i \delta_j \sum_{\ell=1}^{k+1} \frac{n_\ell(n)}{n} \frac{\beta_{i\ell} \beta_{j\ell}}{n_\ell(n)} \sum_{t=t_{\ell-1}}^{t_\ell} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \mathbb{E} \{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \}.
 \end{aligned}$$

D'après le théorème ergodique, lorsque $n \rightarrow \infty$,

$$\sum_{t=1}^n \mathbb{E}(C_{n,t}^2 | G_{t-1}) \xrightarrow{p.s.} \zeta,$$

où

$$\begin{aligned}
 \zeta &= \sum_{i=1}^d \delta_i^2 \sum_{j=1}^{k+1} \alpha_j \beta_{ij}^2 \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_j(x) I(f) \\
 &+ \sum_{i,j=1, i \neq j}^d \delta_i \delta_j \sum_{\ell=1}^{k+1} \alpha_\ell \beta_{i\ell} \beta_{j\ell} \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_\ell(x) I(f) < \infty.
 \end{aligned}$$

Maintenant on montre la condition de Lindeberg.

Soit $\epsilon > 0$, à l'aide de l'inégalité de Hölder conditionnelle, on a

$$\begin{aligned}
 \sum_{t=1}^n \mathbb{E}(C_{n,t}^2 \mathbf{1}_{|C_{n,t}| > \epsilon} | G_{t-1}) &\leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}}(|C_{n,t}|^3 | G_{t-1}) \mathbb{E}^{\frac{1}{3}}(\mathbf{1}_{|C_{n,t}| > \epsilon} | G_{t-1}) \\
 &= \sum_{t=1}^n \mathbb{E}(|C_{n,t}|^3 | G_{t-1})^{\frac{2}{3}} P^{\frac{1}{3}}(|C_{n,t}| > \epsilon | G_{t-1}).
 \end{aligned}$$

D'après l'inégalité de Markov conditionnelle,

$$\begin{aligned}
 &\sum_{t=1}^n \mathbb{E}(C_{n,t}^2 \mathbf{1}_{|C_{n,t}| > \epsilon} | G_{t-1}) \\
 &\leq \frac{1}{\epsilon} \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}}(|C_{n,t}|^3 | G_{t-1}) \mathbb{E}^{\frac{1}{3}}(|C_{n,t}| | G_{t-1}) \\
 &\leq Cst \sum_{t=1}^n \mathbb{E}(|C_{n,t}|^3 | G_{t-1}) \\
 &= Cst \frac{1}{n^{\frac{3}{2}}} \sum_{t=1}^n \mathbb{E} \left\{ \left| \sum_{i=1}^d \delta_i \frac{\beta_i^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \right|^3 \middle| G_{t-1} \right\}
 \end{aligned}$$

$$\begin{aligned}
&\leq Cst \frac{1}{n^{\frac{3}{2}}} \sum_{t=1}^n \mathbb{E} \left\{ \sum_{i=1}^d |\delta_i|^3 \frac{|\beta_i^\top \omega(t)|^3}{\sigma^3(Z_{t-1}, \theta_0)} |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|^3 \middle| G_{t-1} \right\} \\
&\leq Cst \times n^{-\frac{3}{2}} \sum_{t=1}^n \sum_{i=1}^d |\delta_i|^3 \frac{|\beta_i^\top \omega(t)|^3}{\sigma^3(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|^3 \middle| G_{t-1} \right\} \\
&\leq Cst \times n^{-\frac{3}{2}} \sum_{i=1}^d \sum_{t=1}^n |\delta_i|^3 \frac{|\beta_i^\top \omega(t)|^3}{\sigma^3(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|^3 \right\} \\
&= Cst \frac{1}{\sqrt{n}} \sum_{i=1}^d |\delta_i|^3 \frac{1}{n} \sum_{t=1}^n \frac{|\beta_i^\top \omega(t)|^3}{\sigma^3(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|^3 \right\} \\
&= Cst \frac{1}{\sqrt{n}} \sum_{i=1}^d |\delta_i|^3 \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_{ij}|^3}{n_j(n)} \\
&\times \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^3(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|^3 \right\}. \tag{4.3}
\end{aligned}$$

D'après le théorème ergodique, presque sûrement,

$$\begin{aligned}
&\lim_{n \rightarrow \infty} \sum_{i=1}^d |\delta_i|^3 \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{|\beta_{ij}|^3}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^3(Z_{t-1}, \theta_0)} \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\psi_0, \gamma_0)]|^3 \right\} \\
&= \sum_{i=1}^d |\delta_i|^3 \sum_{j=1}^{k+1} \alpha_j |\beta_{ij}|^3 \int_{\mathbb{R}^p} \frac{1}{\sigma^3(x, \theta_0)} dF_j(x) \int_{\mathbb{R}} |\phi_f(x)|^3 f(x) dx < \infty.
\end{aligned}$$

Donc, le membre à droite de (4.3) converge presque sûrement vers 0. Par conséquent, lorsque $n \rightarrow \infty$,

$$\sum_{t=1}^n \mathbb{E} (C_{n,t}^2 \mathbb{1}_{|C_{n,t}| > \epsilon} | G_{t-1}) \xrightarrow{p.s.} 0.$$

Les conditions du Corollaire 3.1 de Hall et Heyde (1980) sont ainsi vérifiées. Il s'ensuit que sous H_0 , lorsque $n \rightarrow \infty$,

$$A_n \xrightarrow{loi} \mathcal{N}(0, \zeta),$$

et on peut ainsi conclure que sous H_0 , lorsque $n \rightarrow \infty$,

$$(\Delta_n(\psi_0, \gamma_0, \beta_1), \Delta_n(\psi_0, \gamma_0, \beta_2), \dots, \Delta_n(\psi_0, \gamma_0, \beta_d))$$

converge en loi vers un vecteur gaussien d'espérance nulle et de matrice de covariance

$$V = \begin{pmatrix} V_{11} & V_{12} & \dots & V_{1d} \\ V_{21} & V_{22} & \dots & V_{2d} \\ \vdots & & & \\ V_{p1} & V_{p2} & \dots & V_{dd} \end{pmatrix}, \tag{4.4}$$

où

$$V_{ij} = I(f) \sum_{t=1}^{k+1} \alpha_t \beta_{it} \beta_{jt} \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_t(x) = I(f) \langle \alpha \odot \varrho, \beta_i \odot \beta_j \rangle, \quad i, j = 1, \dots, d,$$

et pour tout $i = 1, \dots, d$, $\beta_i = (\beta_{i,1}, \dots, \beta_{i,k+1})$. \square

Lemme 4.3.2. *Sous les hypothèses du lemme précédent, le processus $\{\Delta_n(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$ est tendue.*

Démonstration. Soient \mathcal{X} un compact de \mathbb{R}^{k+1} et $(\beta_1, \beta_2) \in \mathcal{X} \times \mathcal{X}$, on a

$$\Delta_n(\psi_0, \gamma_0, \beta_2) - \Delta_n(\psi_0, \gamma_0, \beta_1) = \sum_{t=1}^n \frac{(\beta_2 - \beta_1)^\top}{\sqrt{n}} \frac{\omega(t)}{\sigma(Z_{t-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)].$$

D'où l'on peut écrire

$$\begin{aligned} & [\Delta_n(\psi_0, \gamma_0, \beta_2) - \Delta_n(\psi_0, \gamma_0, \beta_1)]^2 \\ &= \frac{1}{n} \sum_{t,s=1, t \neq s}^n \left\{ \frac{(\beta_2 - \beta_1)^\top \omega(t)}{\sigma(Z_{t-1}, \theta_0)} \frac{(\beta_2 - \beta_1)^\top \omega(s)}{\sigma(Z_{s-1}, \theta_0)} \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \phi_f[\varepsilon_s(\psi_0, \gamma_0)] \right\} \\ &+ \frac{1}{n} \sum_{t=1}^n \frac{[(\beta_2 - \beta_1)^\top \omega(t)]^2}{\sigma^2(Z_{t-1}, \theta_0)} \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \\ &\leq \frac{1}{\tau^2} \frac{1}{n} \sum_{t,s=1, t \neq s}^n \left\{ |(\beta_2 - \beta_1)^\top \omega(t)| \times |(\beta_2 - \beta_1)^\top \omega(s)| \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \phi_f[\varepsilon_s(\psi_0, \gamma_0)] \right\} \\ &+ \frac{1}{\tau^2} \frac{1}{n} \sum_{t=1}^n \left\{ \left| |(\beta_2 - \beta_1)^\top \omega(t)| \right|^2 \times \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \right\}. \end{aligned}$$

En prenant les espérances, on a :

$$\begin{aligned} & \mathbb{E} \left\{ [\Delta_n(\psi_0, \gamma_0, \beta_2) - \Delta_n(\psi_0, \gamma_0, \beta_1)]^2 \right\} \\ &\leq \frac{1}{\tau^2} \frac{1}{n} \sum_{t,s=1, t \neq s}^n \left| |(\beta_2 - \beta_1)^\top \omega(t)| \times |(\beta_2 - \beta_1)^\top \omega(s)| \right. \\ &\quad \times \mathbb{E} [\phi_f(\varepsilon_t(\psi_0, \gamma_0)) \phi_f(\varepsilon_s(\psi_0, \gamma_0))] \\ &+ \frac{1}{\tau^2} \frac{1}{n} \sum_{t=1}^n \left| |(\beta_2 - \beta_1)^\top \omega(t)| \right|^2 \times \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \right\} \\ &\leq \frac{(k+1)^2}{\tau^2} \|\beta_2 - \beta_1\|^2 \frac{1}{n} \sum_{t,s=1, t \neq s}^n \mathbb{E} [\phi_f[\varepsilon_t(\psi_0, \gamma_0)] \phi_f[\varepsilon_s(\psi_0, \gamma_0)]] \\ &+ \frac{(k+1)^2}{\tau^2} \|\beta_2 - \beta_1\|^2 \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left\{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \right\}. \end{aligned}$$

Puisque pour tout $t \in \mathbb{Z}$, les ε_t sont indépendantes, il s'ensuit que

$$\begin{aligned} & \mathbb{E} \{ [\Delta_n(\psi_0, \gamma_0, \beta_2) - \Delta_n(\psi_0, \gamma_0, \beta_1)]^2 \} \\ & \leq \frac{(k+1)^2}{\tau^2} \|\beta_2 - \beta_1\|^2 \frac{1}{n} \sum_{t,s=1, t \neq s}^n \mathbb{E} \{ \phi_f[\varepsilon_t(\psi_0, \gamma_0)] \} \mathbb{E} \{ \phi_f[\varepsilon_s(\psi_0, \gamma_0)] \} \\ & + \frac{(k+1)^2}{\tau^2} \|\beta_2 - \beta_1\|^2 \frac{1}{n} \sum_{t=1}^n \mathbb{E} \{ \phi_f^2[\varepsilon_t(\psi_0, \gamma_0)] \}. \end{aligned}$$

Mais on sait que $\int_{\mathbb{R}} \phi_f(x) f(x) dx = 0$. C'est-à-dire que pour tout $t \in \mathbb{Z}$, $\mathbb{E} [\phi_f(\varepsilon_t(\psi_0, \gamma_0))] = 0$. Par suite,

$$\begin{aligned} \mathbb{E} \{ [\Delta_n(\psi_0, \gamma_0, \beta_2) - \Delta_n(\psi_0, \gamma_0, \beta_1)]^2 \} & \leq \frac{(k+1)^2}{\tau^2} \int_{\mathbb{R}} \phi_f^2(x) f(x) dx \|\beta_2 - \beta_1\|^2 \\ & = \frac{(k+1)^2}{\tau^2} I(f) \|\beta_2 - \beta_1\|^2. \end{aligned}$$

Ce qui implique que

$$\mathbb{E} \{ [\Delta_n(\psi_0, \gamma_0, \beta_2) - \Delta_n(\psi_0, \gamma_0, \beta_1)]^2 \} \leq \|F(\beta_2) - F(\beta_1)\|^2,$$

avec $F(\beta) = \beta(k+1)I(f)^{1/2}/\tau$, $\beta \in \mathcal{X}$.

D'après le Théorème 12.3 Billingsley (1968) énoncé pour $\mathbb{C}([0, 1], \mathbb{R})$, le processus $\{\Delta_n(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathcal{X}}$ est tendu. \square

À l'aide des deux lemmes précédents et en adaptant le Théorème 8.1 de Billingsley (1968), sous les hypothèses (A') , (A_0) - (A_5) et (B_0) - (B_4) , sous H_0 , $\{\Delta_n(\psi_0, \gamma_0, \cdot), n \geq 1\}$ converge faiblement vers le processus gaussien $\Delta(\psi_0, \gamma_0, \cdot)$ dans $\mathbb{C}(\mathcal{X}, \mathbb{R})$. Puisque \mathcal{X} est arbitraire, sous H_0 , $\{\Delta_n(\psi_0, \gamma_0, \cdot), n \geq 1\}$ converge faiblement vers $\Delta(\psi_0, \gamma_0, \cdot)$ dans $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$.

ii) D'après le lemme précédent, sous H_0 , le processus $\{\Delta_n(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$ est tendu. Puisque H_0 et $H_\beta^{(n)}$ sont contiguës pour tout $\beta \in \mathbb{R}^{k+1}$, H_0 et $H^{(n)} = \bigcup_{\beta \in \mathbb{R}^{k+1}} H_\beta^{(n)}$ sont contiguës. Il découle alors du Théorème 4.2.4 de Sen (1981), sous $H^{(n)}$, $\{\Delta_n(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$ est tendu.

D'après le lemme 4.1.1, pour tout $i = 1, \dots, d$, $d \in \mathbb{N}^*$ et pour tout $\beta_1, \dots, \beta_d \in \mathbb{R}^{k+1}$, sous H_0 , le vecteur aléatoire $(\Delta_n(\psi_0, \gamma_0, \beta), \Delta_n(\psi_0, \gamma_0, \beta_1), \Delta_n(\psi_0, \gamma_0, \beta_2), \dots, \Delta_n(\psi_0, \gamma_0, \beta_d))$ converge en loi vers un vecteur gaussien de moyenne $(\mu(\psi_0, \beta)/2, 0, 0, \dots, 0)$ et de matrice de covariance

$$\begin{pmatrix} \mu(\psi_0, \beta) & C^\top \\ C & V \end{pmatrix},$$

où C est un vecteur de d lignes telle que la i ème ligne est une fonction définie pour tout

$\beta \in \mathbb{R}^{k+1}$ par

$$\begin{aligned} C_i(\beta) &= \lim_{n \rightarrow \infty} \text{Cov}(\Delta_n(\psi_0, \gamma_0, \beta_i), \Delta_n(\psi_0, \gamma_0, \beta)) \\ &= I(f) \sum_{l=1}^{k+1} \alpha_l \beta_{i,l} \beta^l \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_l(x) \\ &= I(f) \langle \alpha \odot \varrho, \beta_i \odot \beta \rangle, \text{ pour } l = 1, \dots, d, \end{aligned}$$

où $\beta_i = (\beta_{i,1}, \dots, \beta_{i,k+1})$, $\beta = (\beta^1, \dots, \beta^{k+1})$ et V la matrice définie par (4.4).

D'après le troisième lemme de Le Cam, sous $H^{(n)}$,

$$(\Delta_n(\psi_0, \gamma_0, \beta_1), \Delta_n(\psi_0, \gamma_0, \beta_2), \dots, \Delta_n(\psi_0, \gamma_0, \beta_d)) \xrightarrow{\text{loi}} \mathcal{N}(C, V).$$

Par conséquent, sous $H^{(n)}$, $\{\Delta_n(\psi_0, \gamma_0, \beta)\}_{\beta \in \mathbb{R}^{k+1}}$ tend faiblement vers un processus gaussien dans $C(\mathbb{R}^{k+1}, \mathbb{R})$. \square

4.3.2 Cas où γ_0 est inconnu

Proposition 4.3.2. *Sous les hypothèses (A'), (A₀)-(A₅) et (B₀)-(B₄), pour toute suite d'estimateurs consistants et asymptotiquement normaux $\{(\psi_n, \widehat{\gamma}_{0,n})\}_{n \geq 1}$ de (ψ_0, γ_0) on a pour tout $\beta \in \mathbb{R}^{k+1}$ et pour tout $n \geq 1$,*

$$\boldsymbol{\rho}(\Delta_n(\psi_0, \gamma_0, \beta), \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)) = o_P(1),$$

où $\{N(n)\}_{n \geq 1}$ est une sous-suite de $\{1, \dots, n\}$ telle que $n/N(n) \rightarrow 0$, lorsque $n \rightarrow \infty$.

Démonstration. Soit $\beta \in \mathbb{R}^{k+1}$, pour montrer que $\boldsymbol{\rho}(\Delta_n(\psi_0, \gamma_0, \beta), \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)) = o_P(1)$, il suffit de montrer que pour tout $j \in \mathbb{N}^*$,

$$\boldsymbol{\rho}_j(\Delta_n(\psi_0, \gamma_0, \beta), \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)) = o_P(1).$$

On a pour tout $j \in \mathbb{N}^*$,

$$\begin{aligned} &\boldsymbol{\rho}_j(\Delta_n(\psi_0, \gamma_0, \beta), \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)) \\ &= \sup_{\|\beta\| < j} |\Delta_n(\psi_0, \gamma_0, \beta), \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)| \\ &= \sup_{\|\beta\| < j} |\Delta_n(\psi_0, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \gamma_0, \beta) + \Delta_n(\psi_{N(n)}, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)| \\ &\leq \sup_{\|\beta\| < j} |\Delta_n(\psi_0, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \gamma_0, \beta)| + \sup_{\|\beta\| < j} |\Delta_n(\psi_{N(n)}, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta)|. \end{aligned}$$

On a pour tout $\beta \in \mathbb{R}^{k+1}$,

$$\Delta_n(\psi_0, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \gamma_0, \beta)$$

$$= (\psi_{N(n)} - \psi_0) \partial_\psi \Delta_n(\psi_0, \gamma_0, \beta) + \frac{1}{2} (\psi_0 - \psi_{N(n)})^\top \partial_\psi^2 \Delta_n^2(\tilde{\psi}_n, \gamma_0, \beta) (\psi_0 - \psi_{N(n)}),$$

où $\tilde{\psi}_n$ est comprise entre ψ_0 et $\psi_{N(n)}$.

Ce qui implique que

$$\begin{aligned} & \sup_{\|\beta\| < j} |\Delta_n(\psi_0, \gamma_0, \beta) - \Delta_n(\psi_{N(n)} - \gamma_0, \beta)| \\ & \leq \|\sqrt{n}(\psi_0 - \psi_{N(n)})\| \sup_{\|\beta\| < j} \left\| \frac{1}{\sqrt{n}} \partial_\psi \Delta_n(\psi_0, \gamma_0, \beta) \right\| \\ & + \frac{1}{2} \|\sqrt{n}(\psi_0 - \psi_{N(n)})\| \times \|\psi_0 - \psi_{N(n)}\| \sup_{\|\beta\| < j} \left\| \frac{1}{\sqrt{n}} \partial_\psi^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta) \right\|. \end{aligned}$$

Pour tout $\beta \in \mathbb{R}^{k+1}$, on a d'après (3.4)

$$\frac{1}{\sqrt{n}} \partial_\rho \Delta_n(\psi_0, \gamma_0, \beta) = -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \partial_\rho m(Z_{t-1}, \rho_0) \phi'_f(\varepsilon_t(\psi_0, \gamma_0)).$$

Ceci est égal à $\Delta_n^{(2)} = \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\rho m(Z_{t-1}, \rho_n) \phi'_f[\varepsilon_t(\psi_0, \gamma_0)]$, en remplaçant ψ_n par ψ_0 . D'où, en procédant comme pour la convergence de $|\Delta_n^{(2)}|$, on montre facilement que $\sup_{\|\beta\| < j} \|(1/\sqrt{n}) \partial_\rho \Delta_n(\psi_0, \gamma_0, \beta)\|$ converge en probabilité vers une limite finie.

On a aussi, par (3.5),

$$\begin{aligned} \partial_\theta \Delta_n(\psi_0, \gamma_0, \beta) &= -\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \partial_\theta \sigma(Z_{t-1}, \theta_0) \varepsilon_t(\psi_0, \gamma_0) \\ &- \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_0)} \partial_\theta \sigma(Z_{t-1}, \theta_0) \phi_f[\varepsilon_t(\psi_0, \gamma_0)]. \end{aligned}$$

Ceci vaut

$$\begin{aligned} \Delta_n^{(3)} - \Delta_n^{(4)} &= -\frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \phi'_f[\varepsilon_t(\psi_n, \gamma_0)] \partial_\theta \sigma(Z_{t-1}, \theta_n) \varepsilon_t(\psi_n, \gamma_0) \\ &- \frac{1}{n} \sum_{t=1}^n \frac{\beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_n)} \partial_\theta \sigma(Z_{t-1}, \theta_n) \phi_f[\varepsilon_t(\psi_n, \gamma_0)] \end{aligned}$$

en remplaçant ψ_n par ψ_0 . Ici aussi, en procédant comme pour la convergence de $|\Delta_n^{(3)}|$ et $|\Delta_n^{(4)}|$, montre que $\sup_{\|\beta\| < j} \|(1/\sqrt{n}) \partial_\theta \Delta_n(\psi_0, \gamma_0, \beta)\|$ converge en probabilité vers une variable aléatoire positive. Par conséquent, $\sup_{\|\beta\| < j} \|(1/\sqrt{n}) \partial_\psi \Delta_n(\psi_0, \gamma_0, \beta)\|$ converge en probabilité vers une variable aléatoire positive.

On a pour tout $\beta \in \mathbb{R}^{k+1}$,

$$\frac{1}{\sqrt{n}} \partial_\psi^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta) = \begin{pmatrix} D_{n,1,1}(\tilde{\psi}_n, \gamma_0, \beta) & D_{n,1,2}(\tilde{\psi}_n, \gamma_0, \beta) \\ D_{n,2,1}(\tilde{\psi}_n, \gamma_0, \beta) & D_{n,2,2}(\tilde{\psi}_n, \gamma_0, \beta) \end{pmatrix}.$$

Dans la démonstration de la Proposition 3.3.2, on a montré que pour tout $l, m = 1, 2$, $\sum_{\|\beta\| < j} \|D_{n,l,m}\|$ converge en probabilité vers une variable aléatoire. Il en découle que $\sup_{\|\beta\| < j} \|(1/\sqrt{n})\partial_\psi^2 \Delta_n(\tilde{\psi}_n, \gamma_0, \beta)\|$ converge en probabilité vers une variable aléatoire positive. Et puisque que $\sqrt{n}(\psi_0 - \psi_{N(n)}) = o_P(1)$ et $\psi_0 - \psi_{N(n)} = o_P(1)$, on a

$$\sup_{\|\beta\| < j} |\Delta_n(\psi_0, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \gamma_0, \beta)| = o_P(1). \quad (4.5)$$

Par ailleurs, pour tout $\beta \in \mathbb{R}^{k+1}$,

$$\begin{aligned} & \Delta_n(\psi_{N(n)}, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \hat{\gamma}_{0,N(n)}, \beta) \\ &= -(\hat{\gamma}_{0,N(n)} - \gamma_0)^\top \partial_\gamma \Delta_n(\psi_{N(n)}, \gamma_0, \beta) \\ &+ \frac{1}{2}(\hat{\gamma}_{0,N(n)} - \gamma_0)^\top \partial_\gamma^2 \Delta_n(\psi_{N(n)}, \tilde{\gamma}_{N(n)}, \beta)(\hat{\gamma}_{0,N(n)} - \gamma_0), \end{aligned}$$

où $\tilde{\gamma}_{0,N(n)}$ est comprise entre $\hat{\gamma}_{0,N(n)}$ et γ_0 .

Donc,

$$\begin{aligned} & \sup_{\|\beta\| < j} |\Delta_n(\psi_{N(n)}, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \hat{\gamma}_{0,N(n)}, \beta)| \\ & \leq \|\sqrt{n}(\hat{\gamma}_{0,N(n)} - \gamma_0)\| \sup_{\|\beta\| < j} \left\| \frac{1}{\sqrt{n}} \partial_\gamma \Delta_n(\psi_{N(n)}, \gamma_0, \beta) \right\| \\ & + \frac{1}{2} \|\sqrt{n}(\hat{\gamma}_{0,N(n)} - \gamma_0)\| \times \|\hat{\gamma}_{0,N(n)} - \gamma_0\| \sup_{\|\beta\| < j} \left\| \frac{1}{\sqrt{n}} \partial_\gamma^2 \Delta_n(\psi_{N(n)}, \tilde{\gamma}_{0,N(n)}, \beta) \right\| \end{aligned}$$

Pour tout $\beta \in \mathbb{R}^{k+1}$, d'après (3.17), on a

$$\begin{aligned} & \frac{1}{\sqrt{n}} \partial_\gamma \Delta_n(\psi_{N(n)}, \gamma_0, \beta) \\ &= -\frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_{N(n)})} \phi'_f[\varepsilon_t(\psi_{N(n)}, \gamma_0)] \\ &= -\frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_{N(n)})} \phi'_f[\varepsilon_t(\psi_{N(n)}, \gamma_0)] \\ &= -\frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_{N(n)})} \{ \phi'_f[\varepsilon_t(\psi_{N(n)}, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] + \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \} \\ &= -\frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_{N(n)})} \{ \phi'_f[\varepsilon_t(\psi_{N(n)}, \gamma_0)] - \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \} \\ &- \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_{N(n)})} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)] \\ &= -(\psi_{N(n)} - \psi_0) \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_{N(n)})} \phi''_f[\varepsilon_t(\tilde{\psi}_{N(n)}, \gamma_0)] \\ &- \frac{1}{n} \sum_{t=1}^n \frac{\omega(t) \beta^\top \omega(t)}{\sigma^2(Z_{t-1}, \theta_{N(n)})} \phi'_f[\varepsilon_t(\psi_0, \gamma_0)], \end{aligned}$$

où $\tilde{\psi}_{N(n)}$ est entre $\psi_{N(n)}$ et ψ_0 .

Par de simples majorations et en utilisant l'hypothèse (B_0) , du fait que lorsque $\psi_{N(n)} - \psi_0$ converge en probabilité vers 0, par le théorème ergodique, on obtient que lorsque $n \rightarrow \infty$, $\sup_{\|\beta\| < j} \|(1/\sqrt{n})\partial_\gamma \Delta_n(\psi_{N(n)}, \gamma_0, \beta)\|$ converge en probabilité vers 0.

Pour tout $\beta \in \mathbb{R}^{k+1}$, on a d'après (3.18)

$$\frac{1}{\sqrt{n}} \partial_\gamma^2 \Delta_n(\psi_{N(n)}, \tilde{\gamma}_{0,N(n)}, \beta) = \frac{1}{n} \sum_{t=1}^n \frac{\omega(t)\omega(t)^\top \beta^\top \omega(t)}{\sigma^3(Z_{t-1}, \theta_{N(n)})} \phi_f''[\varepsilon_t(\psi_{N(n)}, \tilde{\gamma}_{0,N(n)})].$$

Ce qui implique que

$$\sup_{\|\beta\| < j} \left\| \frac{1}{\sqrt{n}} \partial_\gamma^2 \Delta_n(\psi_{N(n)}, \tilde{\gamma}_{0,N(n)}, \beta) \right\| \leq Cst \sup_{\|\beta\| < j} \frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)|.$$

Or, d'après le théorème ergodique, pour tout $\beta \in \mathbb{R}^{k+1}$, lorsque $n \rightarrow \infty$,

$$\frac{1}{n} \sum_{t=1}^n |\beta^\top \omega(t)| \xrightarrow{p.s.} \sum_{j=1}^{k+1} \alpha_j |\beta_j| \int_{\mathbb{R}^p} \frac{1}{\sigma^3(x, \theta_0)} dF_j(x) < \infty.$$

D'où, $\sup_{\|\beta\| < j} \|(1/\sqrt{n})\partial_\gamma^2 \Delta_n(\psi_{N(n)}, \tilde{\gamma}_{0,N(n)}, \beta)\|$ converge en probabilité vers une limite finie. De plus, on sait que $\sqrt{n}(\hat{\gamma}_{N(n)} - \gamma_0) = o_P(1)$ et $\hat{\gamma}_{N(n)} - \gamma_0 = o_P(1)$. Par conséquent, lorsque $n \rightarrow \infty$,

$$\sup_{\|\beta\| < j} |\Delta_n(\psi_{N(n)}, \gamma_0, \beta) - \Delta_n(\psi_{N(n)}, \hat{\gamma}_{0,N(n)}, \beta)| \xrightarrow{P} 0. \quad (4.6)$$

Par (4.5) et (4.6), on a pour tout $j \in \mathbb{N}^*$,

$$\rho_j(\Delta_n(\psi_0, \gamma_0, \beta), \Delta_n(\psi_{N(n)}, \hat{\gamma}_{0,N(n)}, \beta)) = o_P(1).$$

□

D'après les Propositions 4.3.1 et 4.3.2 on a le corollaire suivant :

Corollaire 4.3.1. *Sous les hypothèses (A') , (A_0) - (A_5) , (B_0) - (B_4) , on a*

- i. *Sous H_0 , $\{\Delta_n(\psi_{N(n)}, \hat{\gamma}_{0,N(n)}, \cdot), n \geq 1\}$ converge faiblement vers le processus $\Delta(\psi_0, \gamma_0, \cdot)$ dans $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$.*
- ii. *Sous $H^{(n)}$, $\{\Delta_n(\psi_{N(n)}, \hat{\gamma}_{0,N(n)}, \cdot), n \geq 1\}$ converge faiblement vers le processus $\Delta(\psi_0, \gamma_0, \cdot) + C(\cdot)$ dans $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$,*

où C est la fonction définie pour tout $\beta = (\beta_1, \dots, \beta_{k+1}) \in \mathbb{R}^{k+1}$ par

$$C(\beta) = I(f) \sum_{j=1}^{k+1} \alpha_j \beta_j^2 \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_j(x) = I(f) \langle \alpha \odot \varrho, \beta \odot \beta \rangle.$$

Dans ce qui suit, les techniques utilisées sont celles de Ngatchou-Wandji (2009), Meintanis *et al.* (2015) ou encore de Henze et Wagner (1997).

On considère η une mesure positive et $g(\beta)$ sa densité par rapport à la mesure de Lebesgue, à support compact dans \mathbb{R}^{k+1} . Soit $\mathcal{L}_2 = L_2(\eta)$ l'ensemble des fonctions h de \mathbb{R}^{k+1} telles que $\int_{\mathbb{R}^{k+1}} h^2(\beta) d\eta(\beta) < \infty$. Pour $h_1, h_2, h \in \mathcal{L}_2$, on pose $\langle h_1, h_2 \rangle = \int_{\mathbb{R}^{k+1}} h_1(\beta) h_2(\beta) d\eta(\beta)$ et $\|h\|_{\mathcal{L}_2} = \langle h, h \rangle^{\frac{1}{2}}$, respectivement le produit scalaire et la norme dans \mathcal{L}_2 . Soit pour tout $(\beta_1, \beta_2) \in \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$,

$$\begin{aligned} K(\psi_0, \beta_1, \beta_2) &= I(f) \sum_{j=1}^{k+1} \alpha_j \beta_{1,j} \beta_{2,j} \int_{\mathbb{R}^p} \frac{1}{\sigma^2(x, \theta_0)} dF_j(x) \\ &\approx I(f) \sum_{j=1}^{k+1} \alpha_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_n)}. \end{aligned}$$

où $\beta_i = (\beta_{i,1}, \dots, \beta_{i,k+1})^\top$, $i = 1, 2$.

Lemme 4.3.3. *K est un noyau de Mercer.*

Démonstration. Soit $m \in \mathbb{N}$ et a_1, a_2, \dots, a_m des nombres arbitraires dans \mathbb{R} . On a

$$\begin{aligned} \text{Var} \left(\sum_{i=1}^m a_i \Delta(\psi_0, \gamma_0, \beta_i) \right) &= \text{Cov} \left(\sum_{i=1}^m a_i \Delta(\psi_0, \gamma_0, \beta_i), \sum_{j=1}^m a_j \Delta(\psi_0, \gamma_0, \beta_j) \right) \\ &= \sum_{i,j=1}^m a_i a_j \text{Cov} (\Delta(\psi_0, \gamma_0, \beta_i), \Delta(\psi_0, \gamma_0, \beta_j)) \\ &= \sum_{i,j=1}^m a_i a_j K(\psi_0, \beta_i, \beta_j). \end{aligned}$$

D'où, pour tout $m \in \mathbb{N}$ et a_1, a_2, \dots, a_m des nombres arbitraires dans \mathbb{R} ,

$$\sum_{i,j=1}^m a_i a_j K(\psi_0, \beta_i, \beta_j) \geq 0.$$

Par conséquent, K est un noyau semi-défini positif. Ainsi, K est un noyau de Mercer. \square

D'après Riesz et Szökefalvi-Nagy (1972), l'opérateur linéaire défini dans \mathcal{L}_2 par

$$\nabla_k h(\beta_1) = \int_{\mathbb{R}^{k+1}} K(\psi_0, \beta_1, \beta_2) h(\beta_2) d\eta(\beta_2)$$

admet des valeurs propres λ_j , $j \geq 1$, telles que $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$, et des fonctions propres e_1, e_2, \dots , qui forment une base orthonormée de \mathcal{L}_2 .

Soit pour tout $j \geq 1$, le nombre réel

$$\pi_j = \langle C, e_j \rangle = \int_{\mathbb{R}^{k+1}} C(\beta) e_j(\beta) d\eta(\beta).$$

Théorème 4.3.1. *Sous les hypothèses (A'), (A₀)-(A₅) et (B₀)-(B₄) on a, lorsque $n \rightarrow \infty$,*

i. *Sous H_0 ,*

$$\Psi_n = \int_{\mathbb{R}^{k+1}} \Delta_n^2(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) d\eta(\beta) \xrightarrow{\text{loi}} \sum_{j=1}^{\infty} \lambda_j Z_j^2,$$

où pour tout $j \geq 1$, les Z_j sont des variables aléatoires indépendantes et identiquement distribuées et chacune suit la loi normale centrée réduite.

ii. *Sous $H^{(n)}$, pour tout $\beta \in \mathbb{R}^{k+1}$,*

$$\Psi_n = \int_{\mathbb{R}^{k+1}} \Delta_n^2(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) d\eta(\beta) \xrightarrow{\text{loi}} \sum_{j=1}^{\infty} \lambda_j Z_j^{*2},$$

où pour tout $j \geq 1$, les Z_j^ sont des variables aléatoires indépendantes et de loi gaussienne de moyenne $\pi_j/\sqrt{\lambda_j}$, $\lambda_j > 0$, et de variance 1.*

Démonstration. i) D'après le Corollaire 4.2.1, sous H_0 , $\Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \cdot)$ converge faiblement vers le processus $\Delta(\psi_0, \gamma_0, \cdot)$ dans $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$, de noyau de covariance K . Sous nos hypothèses, $\Delta(\psi_0, \gamma_0, \cdot)$ est un élément de \mathcal{L}_2 . Donc, d'après la représentation de Karhunen-Loève, on a

$$\Delta(\psi_0, \gamma_0, \beta) = \sum_{j=1}^{\infty} N_j e_j(\beta), \quad \beta \in \mathbb{R}^{k+1},$$

où pour tout $j \geq 1$, $N_j = \langle \Delta(\psi_0, \gamma_0, \cdot), e_j \rangle$ est une variable aléatoire, de loi normale standard de variance λ_j .

Ainsi, $\|\Delta(\psi_0, \gamma_0, \cdot)\|_{\mathcal{L}_2}^2 = \sum_{j=1}^{\infty} N_j^2$ et on sait que $\mathbb{E}(N_j^2) = \lambda_j \geq 0$, $\forall j \geq 1$. D'où, pour $\lambda_j > 0$, on peut observer que les $Z_j = N_j/\sqrt{\lambda_j}$, $j \geq 1$ sont des variables aléatoires gaussiennes centrées réduites.

On a,

$$\Psi = \int_{\mathbb{R}^{k+1}} \Delta^2(\psi_0, \gamma_0, \beta) d\eta(\beta) = \|\Delta(\psi_0, \gamma_0, \cdot)\|_{\mathcal{L}_2}^2 = \sum_{j=1}^{\infty} \lambda_j Z_j^2.$$

Par conséquent,

$$\Psi_n = \int_{\mathbb{R}^{k+1}} \Delta_n^2(\psi_{N(n)}, \widehat{\gamma}_{0,N(n)}, \beta) d\eta(\beta) \xrightarrow{\text{loi}} \Psi = \int_{\mathbb{R}^{k+1}} \Delta^2(\psi_0, \gamma_0, \beta) d\eta(\beta) = \sum_{j=1}^{\infty} \lambda_j Z_j^2.$$

ii) D'après le Corollaire 4.1.1, sous $H^{(n)}$, lorsque $n \rightarrow \infty$,

$\{\Delta_n(\psi_{N(n)}, \widehat{\gamma}_{0,n}, \cdot), n \geq 1\}$ converge faiblement vers le processus

$$\Delta(\psi_0, \gamma_0, \cdot) + C(\cdot) \text{ dans } \mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R}).$$

Pour les mêmes raisons que dans la preuve de la partie i), la décomposition du processus $\Delta(\psi_0, \gamma_0, \beta) + C(\beta)$ est donnée par

$$\Delta(\psi_0, \gamma_0, \beta) + C(\beta) = \sum_{j=1}^{\infty} \tilde{N}_j e_j(\beta), \quad \beta \in \mathbb{R}^{k+1},$$

où pour tout $j \geq 1$, $\tilde{N}_j = N_j + \pi_j$. Il est facile de voir que pour tout $j \geq 1$, $\tilde{N}_j \sim \mathcal{N}(\pi_j, \lambda_j)$.
Donc,

$$\|\Delta(\psi_0, \gamma_0, \cdot) + C(\cdot)\|_{\mathcal{L}_2}^2 = \sum_{j=1}^{\infty} \tilde{N}_j^2 = \sum_{j \geq 1} \lambda_j Z_j^{*2},$$

où pour tout $j \geq 1$, $Z_j^* \sim \mathcal{N}(\pi_j/\sqrt{\lambda_j}, 1)$.

Il en résulte que sous $H^{(n)}$, lorsque $n \rightarrow \infty$,

$$\Psi_n = \int_{\mathbb{R}^{k+1}} \Delta_n^2(\psi_{N(n)}, \hat{\gamma}_{0, N(n)}, \beta) d\eta(\beta) \xrightarrow{loi} \sum_{j=1}^{\infty} \lambda_j Z_j^{*2},$$

où pour tout $j \geq 1$, $Z_j^* \sim \mathcal{N}(\pi_j/\sqrt{\lambda_j}, 1)$. □

Soit pour tout $(\beta_1, \beta_2) \in \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$,

$$K_n(\psi_n, \beta_1, \beta_2) = I(f) \sum_{j=1}^{k+1} \hat{\alpha}_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_n)}.$$

Lemme 4.3.4. (Convergence ponctuelle)

$K_n(\psi_n, \beta_1, \beta_2)$ est un estimateur consistant de $K(\psi_0, \beta_1, \beta_2)$.

Démonstration. Pour tout $(\beta_1, \beta_2) \in \mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$, on a

$$\begin{aligned} & K_n(\psi_n, \beta_1, \beta_2) - K(\psi_0, \beta_1, \beta_2) \\ \approx & I(f) \left[\sum_{j=1}^{k+1} \hat{\alpha}_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_n)} - \sum_{j=1}^{k+1} \alpha_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \right] \\ = & I(f) \sum_{j=1}^{k+1} \hat{\alpha}_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_n)} - I(f) \sum_{j=1}^{k+1} \hat{\alpha}_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \end{aligned}$$

$$\begin{aligned}
& + I(f) \sum_{j=1}^{k+1} \widehat{\alpha}_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} - I(f) \sum_{j=1}^{k+1} \alpha_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \\
& = I(f) \left\{ \sum_{j=1}^{k+1} \widehat{\alpha}_j \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \left[\frac{1}{\sigma^2(Z_{t-1}, \theta_n)} - \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \right] \right\} \\
& + I(f) \sum_{j=1}^{k+1} (\widehat{\alpha}_j - \alpha_j) \beta_{1,j} \beta_{2,j} \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \\
& = \frac{I(f)}{n} \sum_{j=1}^{k+1} \widehat{\alpha}_j \beta_{1,j} \beta_{2,j} \frac{n}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \left[\frac{\sigma^2(Z_{t-1}, \theta_0) - \sigma^2(Z_{t-1}, \theta_n)}{\sigma^2(Z_{t-1}, \theta_0) \sigma^2(Z_{t-1}, \theta_n)} \right] \\
& + \frac{I(f)}{n} \sum_{j=1}^{k+1} (\widehat{\alpha}_j - \alpha_j) \beta_{1,j} \beta_{2,j} \frac{n}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \\
& = \frac{I(f)}{n} \sum_{j=1}^{k+1} \beta_{1,j} \beta_{2,j} \sum_{t=t_{j-1}}^{t_j} \left[\frac{\sigma^2(Z_{t-1}, \theta_0) - \sigma^2(Z_{t-1}, \theta_n)}{\sigma^2(Z_{t-1}, \theta_0) \sigma^2(Z_{t-1}, \theta_n)} \right] \\
& + \frac{I(f)}{n} \sum_{j=1}^{k+1} \left(1 - \frac{\alpha_j}{\widehat{\alpha}_j} \right) \beta_{1,j} \beta_{2,j} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)} \\
& = I(f)(K^1 + K^2),
\end{aligned}$$

où

$$K^1 := \frac{1}{n} \sum_{j=1}^{k+1} \beta_{1,j} \beta_{2,j} \sum_{t=t_{j-1}}^{t_j} \left[\frac{\sigma^2(Z_{t-1}, \theta_0) - \sigma^2(Z_{t-1}, \theta_n)}{\sigma^2(Z_{t-1}, \theta_0) \sigma^2(Z_{t-1}, \theta_n)} \right]$$

et

$$K^2 := \frac{1}{n} \sum_{j=1}^{k+1} \left(1 - \frac{\alpha_j}{\widehat{\alpha}_j} \right) \beta_{1,j} \beta_{2,j} \sum_{t=t_{j-1}}^{t_j} \frac{1}{\sigma^2(Z_{t-1}, \theta_0)}.$$

Avant d'étudier le comportement asymptotique de K_n , on peut d'abord remarquer que par un développement de Taylor d'ordre 2 de $\sigma^2(Z_{t-1}, \cdot)$ en θ_0 , on peut trouver $\widetilde{\theta}$ comprise entre θ_n et θ_0 , tel que

$$\begin{aligned}
& \sigma^2(Z_{t-1}, \theta_0) - \sigma^2(Z_{t-1}, \theta_n) \\
& = 2(\theta_0 - \theta_n) \sigma(Z_{t-1}, \theta_0) \partial_{\theta} \sigma(Z_{t-1}, \theta_0) \\
& + (\theta_0 - \theta_n) \partial_{\theta} \left(\sigma(Z_{t-1}, \widetilde{\theta}) \partial_{\theta} \sigma(Z_{t-1}, \widetilde{\theta}) \right) (\theta_0 - \theta_n)^{\top} \\
& = 2(\theta_0 - \theta_n) \sigma(Z_{t-1}, \theta_0) \partial_{\theta} \sigma(Z_{t-1}, \theta_0) \\
& + (\theta_0 - \theta_n) \left[(\partial_{\theta} \sigma(Z_{t-1}, \widetilde{\theta}))^2 + \partial_{\theta}^2 \sigma(Z_{t-1}, \widetilde{\theta}) \sigma(Z_{t-1}, \widetilde{\theta}) \right] (\theta_0 - \theta_n)^{\top}.
\end{aligned}$$

Donc,

$$\begin{aligned} |\sigma^2(Z_{t-1}, \theta_0) - \sigma^2(Z_{t-1}, \theta_n)| &\leq 2 \cdot \|\theta_0 - \theta_n\| \cdot \vartheta^2(Z_{t-1}) + \|\theta_0 - \theta_n\|^2 \cdot \vartheta^2(Z_{t-1}) \\ &\leq Cst [\|\theta_0 - \theta_n\| + \|\theta_0 - \theta_n\|^2] \vartheta^2(Z_{t-1}). \end{aligned}$$

Ainsi, on peut écrire les inégalités ci-dessous.

$$\begin{aligned} |K^1| &\leq \frac{1}{\tau^4} \frac{1}{n} \sum_{j=1}^{k+1} |\beta_{1,j} \beta_{2,j}| \sum_{t=t_{j-1}}^{t_j} |\sigma^2(Z_{t-1}, \theta_0) - \sigma^2(Z_{t-1}, \theta_n)| \\ &\leq \frac{1}{\tau^4} (\|\theta_0 - \theta_n\| + \|\theta_0 - \theta_n\|^2) \sum_{j=1}^{k+1} \widehat{\alpha}_j |\beta_{1,j} \beta_{2,j}| \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) \\ &\leq \frac{1}{\tau^4} (\|\theta_0 - \theta_n\| + \|\theta_0 - \theta_n\|^2) \sum_{j=1}^{k+1} |\widehat{\alpha}_j - \alpha_j| \cdot |\beta_{1,j} \beta_{2,j}| \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) \\ &\quad + \frac{1}{\tau^4} (\|\theta_0 - \theta_n\| + \|\theta_0 - \theta_n\|^2) \sum_{j=1}^{k+1} \alpha_j |\beta_{1,j} \beta_{2,j}| \frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}). \end{aligned}$$

D'après le théorème ergodique, lorsque $n \rightarrow +\infty$,

$$\frac{1}{n_j(n)} \sum_{t=t_{j-1}}^{t_j} \vartheta^2(Z_{t-1}) \xrightarrow{p.s.} \int_{\mathbb{R}^p} \vartheta^2(z) dF_j(z) < +\infty.$$

et on a, pour tout $j = 1, 2, \dots, k+1$, $\widehat{\alpha}_j - \alpha_j$ tend vers 0, $\theta_0 - \theta_n$ tend en probabilité vers 0. Par conséquent, lorsque $n \rightarrow \infty$,

$$K^1 \xrightarrow{P} 0.$$

Montrons maintenant que $n \rightarrow \infty$, $K^1 \xrightarrow{P} 0$.

On a les majorations suivantes :

$$\begin{aligned} |K^2| &\leq \frac{1}{\tau^2} \frac{1}{n} \sum_{j=1}^{k+1} \left| 1 - \frac{\alpha_j}{\widehat{\alpha}_j} \right| \cdot |\beta_{1,j} \beta_{2,j}| (t_j - t_{j-1} + 1) \\ &= \frac{1}{\tau^2} \frac{1}{n} \sum_{j=1}^{k+1} n_j(n) \left| 1 - \frac{\alpha_j}{\widehat{\alpha}_j} \right| \cdot |\beta_{1,j} \beta_{2,j}| \\ &= \frac{1}{\tau^2} \sum_{j=1}^{k+1} \widehat{\alpha}_j \left| 1 - \frac{\alpha_j}{\widehat{\alpha}_j} \right| \cdot |\beta_{1,j} \beta_{2,j}|. \end{aligned}$$

Le membre de droite de la dernière inégalité tend vers 0 en probabilité car pour tout $j = 1, 2, \dots, k+1$, $1 - \frac{\alpha_j}{\widehat{\alpha}_j}$ tend vers 0.

Par conséquent,

$$K_n(\psi_n, \beta_1, \beta_2) - K(\psi_0, \beta_1, \beta_2) \xrightarrow{P} 0.$$

□

Remarque 7. On a pour tout Ξ un compact de $\mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$,

$$\sup_{(\beta_1, \beta_2) \in \Xi} |K_n(\psi_n, \beta_1, \beta_2) - K(\psi_0, \beta_1, \beta_2)| \xrightarrow{P} 0.$$

Donc la convergence est uniforme sur tout compact de $\mathbb{R}^{k+1} \times \mathbb{R}^{k+1}$ et elle l'est donc aussi au sens de la distance ρ .

De ce qui procède, on a par le théorème de Lebesgue, que lorsque $n \rightarrow \infty$,

$$\int_{\mathbb{R}^{k+1}} K_n(\psi_n, \cdot, \beta) e_j(\beta) d\eta(\beta) \xrightarrow{P} \int_{\mathbb{R}^{k+1}} K(\psi_0, \cdot, \beta) e_j(\beta) d\eta(\beta).$$

Ainsi, pour n assez grand, on peut écrire

$$\int_{\mathbb{R}^{k+1}} K_n(\psi_n, \cdot, \beta) e_j(\beta) d\eta(\beta) = \lambda_j e_j(\cdot).$$

Cette intégrale peut être estimée par la méthode de Monte Carlo comme dans Ngatchou-Wandji (2009). Pour un entier arbitraire positif m , on échantillonne β_1, \dots, β_m à partir de la distribution associée à la mesure de probabilité η et on écrit

$$\frac{1}{m} \sum_{i=1}^m K_n(\psi_n, \beta_j, \beta_i) e_\ell(\beta_i) = \lambda_\ell e_\ell(\beta_j), \quad \ell \geq 1.$$

Puisque les fonctions propres e_ℓ sont inconnues, il n'est pas facile de calculer les λ_ℓ . Cependant, pour de grandes valeurs de m et pour $j, \ell = 1, \dots, m$, on peut écrire

$$\frac{1}{m} \sum_{i=1}^m K_n(\psi_n, \beta_j, \beta_i) e_\ell(\beta_i) = \lambda_\ell e_\ell(\beta_j),$$

Ce qui nous montre que λ_ℓ et les vecteurs $(e_\ell(\beta_1), \dots, e_\ell(\beta_m))^\top$ peuvent être obtenus en tant que valeurs propres et vecteurs propres de la matrice aléatoire $m \times m$ suivante :

$$\left(\frac{1}{m} K_n(\psi_n, \beta_i, \beta_j) : 1 \leq i, j \leq m \right).$$

Maintenant, pour estimer la loi limite de Ψ_n , plusieurs méthodes sont possibles. Tout d'abord, on commence par tronquer la série. Ensuite, on calcule la distribution de la série tronquée. C'est-à-dire qu'on considère un entier J grand et on approxime la loi de $\sum_{j \geq 1} \lambda_j Z_j^2$ par celle de $\sum_{j=1}^J \lambda_j Z_j^2$.

Plusieurs travaux se sont intéressés au calcul de la distribution d'une somme finie de variables aléatoires du Khi-deux indépendantes centrées ou non-centrées. Par exemple, Kotz

et al. (1967a,b) proposent les fonctions g_J et h_J ci-dessous pour approximer les distributions de $\sum_{j=1}^J \lambda_j Z_j^2$ et $\sum_{j=1}^J \lambda_j Z_j^{*2}$ respectivement :

$$g_J(\lambda, y) = \sum_{j=0}^{\infty} c_j \frac{1}{b} g(J + 2j, \frac{y}{b})$$

et

$$h_J(\lambda, \varrho, y) = \sum_{j=0}^{\infty} a_j \frac{1}{b} h(J + 2j, \varrho, \frac{y}{b}),$$

où $g(J, y)$ est la distribution de la loi χ^2 centrée et de degré de liberté J , $c_0 = \prod_{j=1}^J (b/\sqrt{\lambda_j})^{1/2}$, pour $j \geq 1$, $c_j = (1/j) \prod_{r=0}^{j-1} d_{j-r} c_r$, $d_j = (1/2) \sum_{j=1}^J (1 - (b/\lambda_j))^j$, $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_J)$, $J \in \mathbb{N}$, $h(J, \varrho, y)$ est la distribution de la loi du χ^2 non-centrée, de degré de liberté J et de paramètre de décentrage $\varrho = \sum_{j=1}^J (\pi_j/\sqrt{\lambda_j})$, $a_0 = c_0$, pour tout $j \geq 1$, $a_j = (1/j) \prod_{r=0}^{j-1} b_{j-r} a_r$, $b_1 = (1/2) \sum_{j=1}^J [1 - (\pi_j/\lambda_j)^2] d_j$, $d_j = 1 - b/\sqrt{\lambda_j}$, pour $j \leq 2$, $b_j = (j/2) \sum_{i=1}^J (\pi_i^2/\lambda_i) d_i^{j-1} + (1/2) \sum_{i=1}^J (1 - j(\pi_i^2/\lambda_i)) d_i^j$ et b est un réel positif obtenu de sorte que pour tout $j \geq 1$, $c_j \geq 0$ et $\sum_{j=0}^{\infty} c_j = 1$.

Selon Ruben (1962), le meilleur choix de b est $2\lambda_1 \lambda_J / (\lambda_1 + \lambda_J)$. Les fonctions de répartition de ces distributions respectives sont alors

$$G(\lambda, y) = \sum_{j=1}^{\infty} c_j^c G(J + 2j, \frac{y}{b}),$$

$$H(\lambda, \varrho, y) = \sum_{j=1}^{\infty} a_j^c H(J + 2j, \varrho, \frac{y}{b}),$$

où $G(J, \cdot)$ et $H(J, \varrho, \cdot)$ sont respectivement la fonction de répartition de la loi χ^2 centrée et de degré de liberté J et la fonction de répartition de la loi χ^2 de paramètre de décentrage ϱ et de degré de liberté J .

Kotz *et al.* (1967a,b) montrent que l'approximation de la distribution de $\sum_{j=1}^J \lambda_j Z_j^2$ est meilleure que l'estimateur de la distribution de $\sum_{j=1}^J \lambda_j Z_j^{*2}$ pour une valeur très élevée de y . Cela signifie que l'erreur d'approximation pour le premier cas tend vers 0 plus vite que pour le deuxième.

Gurland (1956) estime la fonction de répartition de $\sum_{j=1}^J \lambda_j Z_j^{*2}$. Il propose l'approximation suivante,

$$F(x) = \frac{1}{2} \sum_{j=0}^{\infty} \frac{a_j}{\lambda^j} \sum_{i=1}^j (-1)^i \binom{j}{i} G_{n+2i} \left(\frac{x}{\lambda} \right),$$

où

$$G_z(x) = \frac{1}{2^{\frac{z}{2}} \Gamma(\frac{z}{2})} \int_0^x e^{-\frac{y}{2} u^{\frac{z}{2}-1}} du, \quad z \in \mathbb{N}$$

et a_j est défini par (5) dans Gurland (1956).

Ensuite, nous avons le travail Imhof (1961), qui établit que pour tout $x > 0$,

$$\mathbb{P}\left(\sum_{j=1}^J \lambda_j \chi_{2,\nu_j}^2 > x\right) = \sum_{j=1}^p \frac{1}{(\nu_j - 1)!} \left[\frac{\partial^{\nu_j - 1}}{\partial \lambda^{\nu_j - 1}} D_j(\lambda, x) \right]_{\lambda=\lambda_j},$$

où $D_j(\lambda, x) = \lambda^{J-1} e^{-\frac{x}{2\lambda}} \prod_{r=1, r \neq j}^J (\lambda - \lambda_r)^{-\nu_r}$ et χ_{2,ν_j}^2 est une variable aléatoire de loi khi-deux, de degré de liberté 2 et de paramètre de décentralisation ν_j .

Il prouve que si nous désignons $Q = \sum_{r=1}^J \lambda_r \chi_{h_r, \delta_r^2}^2$, on a pour tout $x \in \mathbb{R}$,

$$\mathbb{P}(Q > x) = \frac{1}{2} + \frac{1}{\pi} \int_0^{+\infty} \frac{\sin(\theta(u))}{u\rho(u)} du,$$

avec

$$\theta(u) = \frac{1}{2} \sum_{r=1}^J [h_r \tan^{-1}(\lambda_r u) + \delta_r^2 \lambda_r u (1 + \lambda_r^2 u^2)^{-1}] - \frac{1}{2} x u$$

et

$$\rho(u) = \prod_{r=1}^J (1 + \lambda_r^2 u^2)^{\frac{1}{2} h_r} \exp \left[\frac{\frac{1}{2} \sum_{r=1}^J (\delta_r \lambda_r u)^2}{(1 + \lambda_r^2 u^2)} \right].$$

Parmi les travaux les plus récents, nous avons Deheuvels et Martynov (1996), qui montrent que la statistique de Cramer-Von-Mises converge en loi vers $w^2 = \int_I \xi^2(t) dt$ pour un processus gaussien $\{\xi(t), t \in I\}$, $I = [0, 1]$. Ils ont présenté une méthode numérique pour évaluer $\mathbb{P}(w^2 \leq x)$, $x \in \mathbb{R}$. Ils ont proposé l'estimateur $w^{2,m} = \sum_{i=1}^m a_m \xi^2(t_i)$ de w^2 , où a_m et t_i sont déterminés dans l'article à partir du noyau de covariance et la moyenne du vecteur gaussien $\xi = (\xi(t_1), \xi(t_2), \dots, \xi(t_m))$, $m \in \mathbb{N}$. Ils ont montré que cet estimateur est consistant et qu'on peut approximer $\mathbb{P}(w^2 \leq x)$ par $\mathbb{P}(w^{2,m} \leq x)$. Ils ont aussi montré que

$$\begin{aligned} \mathbb{P}(w^{2,J} \leq x) &= \mathbb{P}(Q \leq x) \\ &= 1 - \frac{1}{\pi} \operatorname{artan} \left(\frac{1}{a} \right) - \frac{1}{\pi} \int_0^{+\infty} \left[\frac{\exp\left(\frac{-atx}{2} \sin(\tau(t) - \frac{tx}{2})\right)}{t\beta(t)} \right] dt, \quad a \in \mathbb{R}, \end{aligned}$$

où

$$\tau(t) = \sum_{j=1}^{+\infty} \left[\frac{s_j}{2} w_j(t) + \frac{u_j \delta_j^2 t}{2[(u_j - at)^2 + t^2]} \right], t \in \mathbb{R}_+$$

$$w_j(t) = \arctan \left(\frac{u_j - at}{t} \right), \quad j \geq 1, t \in \mathbb{R}_+$$

$$u_j = \frac{1}{\lambda_j}, \quad \beta(t) = \prod_{j=1}^J \left\{ \left[\left(1 - \frac{at}{\mu_j} \right) \mu_j \right]^{\frac{s_j}{4}} \exp \left\{ -\frac{\delta_j^2 t [a\mu_j - (a^2 + 1)t]}{2[(\mu_j - at)^2 + t^2]} \right\} \right\}, t \in \mathbb{R}_+$$

et

$$Q = \sum_{j=1}^J \lambda_j Z_j^2,$$

où les Z_j^2 sont des variables aléatoires iid de loi du Khi-deux décentrées de paramètre de décentrage s_j et de degrés de liberté δ_j . Cette formule énoncée par Deheuvels et Martynov (1996) est une généralisation de la formule de Imhof (1961).

4.4 Conclusion

Dans ce chapitre, nous avons étudié le cas où l'amplitude du saut est quelconque et inconnue. C'est un cadre plus général que celui du chapitre précédent. En effet, ici nous avons testé un modèle CHARN contre un ensemble de modèle CHARN, le test étudié est de type Cramer-Von-Mises basé sur le processus de vraisemblance. Nous avons trouvé la loi asymptotique de la statistique de test sous l'hypothèse nulle et sous l'hypothèse alternative. La loi asymptotique trouvée est celle d'une somme pondérée des v.a iid de loi du Khi-deux décentrée. L'expression exacte de la puissance locale de test n'est pas obtenue. Cependant, elle est approximée par des méthodes présentées à la fin du chapitre. Des simulations numériques restent à faire pour évaluer la performance du test construit. Celles-ci sont préalables à l'application aux données réelles.

Chapitre 5

Applications

5.1 Introduction

Dans ce chapitre, nous utilisons le logiciel R pour étudier les performances de nos résultats théoriques obtenus au chapitre 3. Nous les appliquons dans un premier temps aux données simulées suivant le modèle

$$\begin{aligned} X_t = & \left(\rho_1 + \rho_2 X_{t-1} e^{\rho_3 X_{t-1}^2} \right) X_{t-1} + \left(\gamma_0 + \frac{\beta}{\sqrt{n}} \right)^\top \omega(t) \\ & + \left(\theta_1 + \theta_2 X_{t-1}^2 e^{-\theta_3 X_{t-1}^2} \right)^{\frac{1}{2}} \varepsilon_t, t \in \mathbb{Z}, \end{aligned} \quad (5.1)$$

où les ρ_j , θ_j et γ_0 sont des paramètres prenant certaines valeurs à préciser, n est la taille de l'échantillon, $(\varepsilon_t)_t$ est un bruit blanc standard de densité f et β est une valeur arbitraire entre -10 et 10 .

Les niveaux nominaux considérés sont $\alpha = 0.01$, $\alpha = 0.05$ et $\alpha = 0.1$ et le nombre de réplifications est $N = 5000$. Nous commençons par étudier la puissance locale du test construit au chapitre 3. Puis, nos méthodes de détection de ruptures. Dans un second temps, nous appliquons nos résultats théoriques aux données de santé, et aux données financières. Nous prenons pour ces applications $\alpha = 0.05$.

5.2 Simulation

5.2.1 Étude de la puissance

5.2.1.1 Cas où $\gamma_0 = 0$

Nous commençons par calculer la puissance locale asymptotique pour $\gamma_0 = 0$. Nous considérons le cas d'une hypothèse nulle de non-rupture contre l'hypothèse alternative

d'un seul point de rupture ($k = 1$), de deux points de rupture ($k = 2$) et de trois points de rupture ($k = 3$). Nous commençons dans le paragraphe ci-dessous par le cas d'un seul point de rupture.

• **Un seul point de rupture ($k = 1$)**

Nous considérons le modèle (5.1) pour $\rho_1 = 0.5$, $\rho_2 = 0$, $\theta_1 = 1$, $\theta_2 = 0$, $n = 60$, $n_1(n) = n_2(n) = 30$. Les Figures 5.1 (a) et 5.1 (c) présentent le chronogramme des données simulées dans le cas où f est la densité gaussienne centrée réduite et $\beta = (0.5; 0.25)^\top$ et $\beta = (0.2; 3)^\top$ respectivement. On peut voir facilement que ces graphiques exhibent chacun un seul point de rupture à l'instant $t = 30$.

Soit $(\xi_t)_{t \in \mathbb{Z}}$ un bruit blanc de loi exponentielle de paramètre λ , $\lambda \in \mathbb{R}_+$. Alors $\varepsilon_t = \lambda(\xi_t - (1/\lambda))$ est une variable aléatoire d'espérance nulle et de variance 1. La densité de ε_t sera appelée densité exponentielle normalisée de paramètre λ .

Les Figures 5.2 (a) et 5.2 (b) correspondent à la puissance locale asymptotique du test pour $\alpha = 0.05$, $\beta \in [-10, 10]$, pour f la densité gaussienne standard, et la densité exponentielle normalisée de paramètre 1.25 respectivement (de même pour les Figures 5.3 (a) et 5.3 (b) pour $\alpha = 0.01$ et les Figures 5.4 (a) et 5.4 (b) pour $\alpha = 0.1$).

Les Figures 5.5 (a) et 5.5 (b) sont les graphiques des puissances correspondant au modèle (5.1) avec des paramètres de nuisances inconnus que nous l'avons estimé par la méthode du maximum de vraisemblance et $\alpha = 0.05$, respectivement pour f densité gaussienne standard, et f densité exponentielle normalisée de paramètre 1.25 (de même pour les Figures 5.6 (a) et 5.6 (b) pour $\alpha = 0.01$ et les Figures 5.7 (a) et 5.7 (b) pour $\alpha = 0.1$). Nous précisons que la valeur $\lambda = 1.25$ est totalement arbitraire.

On voit que sur l'ensemble de ces graphiques, la puissance tend vers 1 et la puissance correspondant au bruit exponentiel tend vers 1 plus vite que celle du bruit gaussien. Il semble, d'après les Figures 5.5 (a), 5.5 (b), 5.6 (a), 5.6 (b), 5.7 (a) et 5.7 (b), que l'estimation du paramètre de nuisance n'affecte pas beaucoup la puissance du test. Si on compare les Figures 5.5 et 5.6, on observe que lorsque $\alpha = 0.05$ la puissance tend vers 1 plus vite que dans le cas où $\alpha = 0.01$. Si on compare les Figures 5.6 et 5.7, on observe que lorsque $\alpha = 0.1$ la puissance tend vers 1 plus vite que dans le cas où $\alpha = 0.01$. Donc la puissance tend vers 1 plus vite quand le seuil α s'éloigne de 0.

Nous considérons ensuite un modèle CHARN dont la non-linéarité est exponentielle. Plus précisément, nous considérons le modèle (5.1) pour $\rho_1 = 0.5$, $\rho_2 = 0.2$, $\rho_3 = 50$, $\theta_1 = 0.1$, $\theta_2 = 0.0025$, $\theta_3 = 1$, $n = 60$, $n_1(n) = n_2(n) = 30$. Nous traçons le chronogramme des observations simulées pour $\beta = (0.2; 0.3)^\top$ sur la Figure 5.1 (d). Nous observons qu'il y a un point de rupture à l'instant $t = 30$.

Les Figures 5.2 (c) et 5.2 (d) présentent la puissance pour la densité normale standard, et la densité exponentielle normalisée de paramètre 1.25 respectivement pour $\alpha = 0.05$ (de même pour les Figures 5.3 (c) et 5.3 (d) pour $\alpha = 0.01$, et les Figures 5.4 (c) et 5.4 (d) pour $\alpha = 0.1$). On peut voir que dans les deux cas, elle tend vers 1. Donc notre test est bien consistant.

• **Deux points de rupture ($k = 2$)**

Dans le cas de deux points de rupture, nous avons considéré $n = 100$, $n_1(n) = n_2(n) = 30$, $n_3(n) = 40$, $\rho_1 = 0.5$, $\rho_2 = 0$, $\theta_1 = 1$, $\theta_2 = 0$ et $\beta = (2; 0.5; 3.5)^\top$. Le chronogramme des données simulées pour une densité gaussienne standard est présenté dans la Figure 5.1 (b). Ce graphique montre qu'il existe deux petites ruptures, la première est à l'instant $t_1 = 30$ et la deuxième est à l'instant $t_2 = 60$. Les résultats de la puissance correspondant au cas où f est la densité gaussienne centrée réduite et au cas où f est la densité exponentielle centrée réduite de paramètre 1.25 sont résumés dans la Table 5.1 ci-dessous pour $\alpha = 0.05$. D'après ce tableau, nous observons que pour $\alpha = 0.05$, la puissance du test tend vers 1. D'où la consistance de notre test.

β	$\begin{pmatrix} 0.2 \\ 0.25 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0.5 \\ 0.75 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.75 \\ 1 \\ 1.5 \end{pmatrix}$	$\begin{pmatrix} 0.75 \\ 1.2 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 2.2 \\ 1.5 \\ 2.2 \end{pmatrix}$	$\begin{pmatrix} 2.5 \\ 2 \\ 3 \end{pmatrix}$	$\begin{pmatrix} 2.75 \\ 2.2 \\ 3.2 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 2.5 \\ 3.8 \end{pmatrix}$	$\begin{pmatrix} 3.5 \\ 3 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 3.5 \\ 3 \\ 0.75 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 3.5 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 4.7 \\ 4 \\ 3 \end{pmatrix}$
Power ¹	0.119	0.271	0.447	0.607	0.831	0.955	0.975	0.994	0.999	0.953	0.999	1.000
Power ²	0.144	0.363	0.596	0.773	0.946	0.994	0.998	0.999	0.999	0.994	1.000	1.000

TABLE 5.1 – Puissance du test pour $k = 2$

• **Trois points de rupture ($k = 3$)**

Considérons maintenant le cas où on a trois ruptures, c'est-à-dire $k = 3$. Nous considérons des échantillons de taille $n = 200$, $n_1(n) = 60$, $n_2(n) = 40$ et $n_3(n) = n_4(n) = 50$. Nous générons les observations à partir du modèle (5.1) lorsque $\rho_1 = 0.5$, $\rho_2 = 0$, $\theta_1 = 1$, $\theta_2 = 0$ et $\beta = (2; 4; 0.5; 3)^\top$. Le chronogramme des données simulées dans le cas où f est la densité d'une loi normale centrée réduite est présenté dans la Figure 5.1 (e). Il est facile de voir que ce graphique exhibe trois points de ruptures, le premier à l'instant $t_1 = 30$, le deuxième à l'instant $t_2 = 100$ et le troisième à l'instant $t_3 = 150$, mais on ne peut pas les voir à l'œil nu car ils sont très petits.

Comme le cas où $k = 2$, la Table 5.2 pour un seuil $\alpha = 0.05$ nous montre que la puissance du test tend vers 1 lorsque β s'éloigne de $(0, 0, 0, 0)$.

β	$\begin{pmatrix} 0.2 \\ 0.25 \\ 0.5 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0.5 \\ 0.75 \\ 1 \\ 1.2 \end{pmatrix}$	$\begin{pmatrix} 0.75 \\ 1 \\ 1.5 \\ 2 \end{pmatrix}$	$\begin{pmatrix} 0.75 \\ 1.2 \\ 2 \\ 2.3 \end{pmatrix}$	$\begin{pmatrix} 2.2 \\ 1.5 \\ 2.2 \\ 2.8 \end{pmatrix}$	$\begin{pmatrix} 2.5 \\ 2 \\ 3 \\ 3 \end{pmatrix}$	$\begin{pmatrix} 2.75 \\ 2.2 \\ 3.2 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 3 \\ 2.5 \\ 3.8 \\ 1.75 \end{pmatrix}$	$\begin{pmatrix} 3.5 \\ 3 \\ 4 \\ 3.2 \end{pmatrix}$	$\begin{pmatrix} 3.5 \\ 3 \\ 2 \\ 5 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 3.5 \\ 4 \\ 4.5 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 5 \\ 2 \\ 7 \end{pmatrix}$
Power ¹	0.144	0.226	0.399	0.352	0.728	0.850	0.781	0.893	0.966	0.893	0.991	1.000
Power ²	0.179	0.299	0.537	0.474	0.880	0.956	0.916	0.975	0.997	0.975	1.000	1.000

TABLE 5.2 – Puissance du test pour $k = 3$

Remarque 8. Dans les applications ci-dessus, lorsque le bruit est de densité exponentielle normalisée, nous choisissons de prendre le paramètre de la loi exponentielle $\lambda = 1.25$. Nous faisons le travail pour d'autres valeurs de λ et nous obtenons des résultats similaires. C'est-à-dire la puissance du test tend vers 1 et pour toute valeur de λ , le test est plus puissant dans le cas exponentiel que dans le cas gaussien. De plus, lorsque le bruit est exponentiel normalisé, nous remarquons que si la valeur de λ augmente, la courbe de la puissance s'approche de plus en plus à 1. Autrement dit, quand λ tend vers l'infinie, la puissance tend vers 1.

5.2.1.2 γ_0 et f sont inconnus

En pratique γ_0 et f sont inconnus et doivent être estimés. Pour cela, nous estimons γ_0 par la méthode des moindres carrés et f par l'estimateur du noyau ou encore de Parzen-Rosenblatt (voir Parzen, 1962, (1.7)) défini par

$$\hat{f}_n(x) = \frac{1}{nh_n^2} \sum_{t=1}^n K\left(\frac{x - \hat{\varepsilon}_t}{h_n}\right), \quad x \in \mathbb{R},$$

où

$$\hat{\varepsilon}_t = \frac{X_t - m(Z_{t-1}, \rho_n) - \hat{\gamma}_n^\top \omega(t)}{\sigma(Z_{t-1}, \theta_n)},$$

avec $\psi_n = (\rho_n, \theta_n)$ et $\hat{\gamma}_n$ les estimateurs des moindres carrés des paramètres de nuisance $\psi_0 = (\rho_0, \theta_0)$ et γ_0 respectivement, h_n est le paramètre de lissage et K la fonction de poids (noyau, kernel) ayant les propriétés suivantes :

- (i) $K(x) \geq 0$ pour tout $x \in \mathbb{R}$ (positivité).
- (ii) $\int_{\mathbb{R}} K(x) dx = 1$ (densité).
- (iii) $\int_{\mathbb{R}} xK(x) dx = 0$ (symétrie).

La puissance du test est calculée sur la base de ces estimations. Nous considérons des échantillons de taille $n = 60$, $n_1(n) = n_2(n) = 30$ et nous générons les observations à

-
- 1. Cas où f est une densité de la loi normale centrée réduite.
 - 2. Cas où f est une densité de la loi exponentielle normalisée de paramètre 1.25.

partir du modèle (5.1) lorsque $\rho_1 = 0.5$, $\rho_2 = 0.2$, $\rho_3 = 50$, $\theta_1 = 0.1$, $\theta_2 = 0.0025$, $\theta_3 = 1$ et f est la densité gaussienne standard ou exponentielle normalisée. Nous obtenons les Figures 5.5 (c), 5.5 (d) pour $\alpha = 0.05$, les Figures 5.6 (c) et 5.6 (d) pour $\alpha = 0.01$ et les Figures 5.7 (c) et 5.7 (d) pour $\alpha = 0.1$. Nous observons que la puissance tend vers 1 et que le test est beaucoup plus puissant dans le cas exponentiel que dans le cas gaussien pour tout $\alpha = 0.05, 0.01, 0.1$.

Remarque 9. *Les Figures 5.5, 5.6 et 5.7 sont obtenus en prenant K un noyau Gaussien. En choisissant d'autres types de noyau comme par exemple noyau d'Epanechnikov, noyau Uniforme, noyau Quadratique etc., nous obtenons des résultats similaires.*

5.2.2 Détection des ruptures

Dans cette section, nous appliquons nos stratégies présentées au chapitre 3 à la détection des ruptures et à l'estimation de leurs localisations. Autrement dit, nous voulons savoir si notre test détecte bien les points de rupture lorsqu'il y en a. Plusieurs exemples sont étudiés.

À part le cas où nous avons simulé les données avec une seule rupture, la recherche aveugle du nombre de ruptures et l'estimation de leurs localisations n'ont été faites que dans la situation où les données ont été simulées avec une seule rupture, et pour les données réelles de la section 5.4.3. Dans les autres cas considérés, le nombre des ruptures a été supposé connu surtout pour les données réelles présentées dans le paragraphe 5.3 parce que leurs chronogrammes nous ont bien renseigné sur les zones susceptibles de contenir des ruptures.

Nous commençons par le cas d'une série d'observations stationnaires, c'est-à-dire le cas où le modèle considéré ne présente aucun risque ($k = 0$).

5.2.2.1 Aucun point de rupture dans les données

Nous commençons par calculer la puissance locale asymptotique dont l'expression est donnée dans les théorèmes du chapitre 3 dans le cas où le modèle ne présente aucun risque c'est-à-dire dans le cas où $k = 0$. Nous considérons le modèle (5.1) lorsque $n = 200$, $\gamma_0 = 0$ et f la densité gaussienne standard. Puis, nous représentons graphiquement la puissance dans le cas où $\rho_1 = \rho_2 = 0$, $\theta_1 = 1$ et $\theta_2 = 0$ et dans le cas où $\rho_1 = 0.5$, $\rho_2 = 0$, $\theta_1 = 1$, $\theta_2 = 0$. Les deux courbes sont construites sur les Figures 5.8, 5.9 et 5.10 dans le cas où $\alpha = 0.05$, $\alpha = 0.01$ et $\alpha = 0.1$ respectivement.

On observe que la puissance du test ne dépasse pas 0.053 lorsque $\alpha = 0.05$, ne dépasse pas 0.01 lorsque $\alpha = 0.01$ et ne dépasse pas 0.11 lorsque $\alpha = 0.1$. Donc pour tout

$\alpha = 0.05, 0.01, 0.1$, on accepte l'hypothèse nulle c'est-à-dire qu'on accepte l'hypothèse de non rupture, ce qui est bien le cas ici.

5.2.2.2 Des points de rupture dans les données

Nous considérons dans ce paragraphe le modèle (5.1) dans le cas où $n = 200$, $\rho_1 = 0.5$, $\rho_2 = 0$, $\theta_1 = 1$, $\theta_2 = 0$ et $\gamma_0 = 0$. Nous commençons dans la section ci-dessous l'étude du cas où $k = 1$.

- **Un seul point de rupture ($k = 1$)**

Nous traçons les courbes de la puissance du test en faisant varier les instants de rupture. Nous obtenons les Figures 5.11 (b) et 5.11 (d) pour un seuil $\alpha = 0.05$, les Figures 5.12 (b) et 5.12 (d) pour un seuil $\alpha = 0.01$, et les Figures 5.13 (b) et 5.13 (d) pour $\alpha = 0.1$. Par exemple dans la première figure (Figure 5.11 (b)), nous simulons des données présentant une rupture à l'instant $t_1 = 30$ c'est-à-dire $n_1(n) = 30$ et $n_2(n) = 170$ et nous traçons la puissance du test pour $\beta = (0; -0.5)^\top$, nous obtenons la courbe en marron. Ensuite, nous simulons une autre série des données présentant une rupture à l'instant $t_1 = 60$, c'est-à-dire $n_1(n) = 60$ et $n_2(n) = 140$ et nous traçons la courbe de la puissance pour $\beta = (0; 0.4)^\top$. Nous obtenons la courbe en bleu. Ainsi de suite.

Donc chaque couleur présente la courbe de la puissance d'une série des données. Il s'agit de plusieurs exemples simulés.

Dans les 3 cas considérés (voir les Figures 5.11, 5.12 et 5.13), on constate que la puissance du test est maximale à l'instant de rupture. Donc on conclut que la puissance du test considéré est un bon indicateur de l'instant auquel on a un point de rupture.

- **Deux points de rupture ($k = 2$)**

Dans ce cas nous traçons la courbe de la puissance asymptotique locale du test pour des différentes valeurs des instants de rupture t_1 et t_2 , et pour plusieurs valeurs de β . C'est-à-dire nous prenons plusieurs exemples simulés où chaque exemple présente deux ruptures. La première est à l'instant t_1 et la deuxième est à l'instant t_2 , et pour plusieurs choix de β .

Les Figures 5.14 (b,d,f), 5.15 (b,d,f) et 5.16 (b,d,f) nous montrent que la puissance du test atteint son maximum à l'instant exact de rupture c'est-à-dire exactement au point (t_1, t_2) . Donc comme dans le cas d'un seul point de rupture, la puissance théorique du test construit est un bon indicateur de l'endroit où on a des points de rupture pour toute valeur $\alpha \in [0, 1]$. La Table 5.3 obtenue après 500 réplifications, nous montre que notre test

$(t_1, t_2) \backslash \beta$	$\begin{pmatrix} 0 \\ 0.2 \\ -0.2 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.4 \\ 0.4 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.5 \\ 0.5 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.7 \\ -0.7 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.9 \\ -0.9 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1.3 \\ -1.3 \end{pmatrix}$
(60,100)	(59,100)	(70,100)	(60,100)	(65,100)	(60,100)	(60,100)	(60,100)
(40,120)	(38,120)	(40,120)	(40,120)	(40,120)	(40,120)	(40,120)	(40,120)
(50,140)	(40,140)	(50,140)	(50,140)	(50,138)	(50,140)	(50,140)	(50,140)
(80,160)	(80,160)	(80,160)	(80,159)	(80,159)	(80,160)	(80,160)	(80,160)
(100,170)	(99,170)	(100,170)	(100,170)	(100,170)	(100,169)	(100,170)	(100,170)

TABLE 5.3 – $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$

détecte généralement les bons instants de rupture dans les données simulées, mais lorsque les composantes du vecteur β sont de même signe c'est-à-dire lorsque les changements consécutifs ont des amplitudes du même signe, notre méthode, tout en indiquant l'existence de ruptures, ne détecte généralement pas les bons points de rupture, c'est-à-dire les bons couples (t_1, t_2) .

5.2.2.3 Encore des points de rupture dans les données

Nous considérons de nouveau le modèle (5.1) avec $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$ et $\gamma_0 = 0$.

- **Cas où $k = 1, 2$**

Comme à la section 5.2.2.2, nous traçons les courbes de la puissance du test lorsqu'on a un seul point de rupture. Nous obtenons les Figures 5.11 (a) et 5.11 (c) pour $\alpha = 0.05$, les Figures 5.12 (a) et 5.12 (c) pour $\alpha = 0.01$, et les Figures 5.13 (a) et 5.13 (c) pour $\alpha = 0.1$. Dans les 3 cas considérés (les Figures 5.11, 5.12 et 5.13), on constate que la puissance du test est maximale à l'instant de rupture. Toujours, comme à la section 5.1.2.2.2, nous représentons graphiquement la puissance du test considéré lorsque $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1, \theta_2 = 0, \gamma_0 = 0$ et $k = 2$. Les Figures 5.14 (a,c,e), 5.15 (a,c,e) et 5.16 (a,c,e) nous montrent que la puissance du test atteint son maximum à l'instant de rupture c'est-à-dire au point (t_1, t_2) exact. Donc on conclut que dans le cas où $k = 1, 2$, les points de rupture sont obtenus par les indices de temps donnant la plus grande puissance.

La Table 5.4 obtenu après 500 répliquions, nous montre que lorsque $k = 2$, notre test détecte les bons instants de ruptures dans les données simulées. Mais lorsque les composantes du vecteur β sont de même signe c'est-à-dire lorsque les changements consécutifs

(t_1, t_2) \ β	$\begin{pmatrix} 0 \\ -0.25 \\ 0.25 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.3 \\ -0.3 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.4 \\ -0.4 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -0.6 \\ 0.6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.75 \\ -0.75 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.8 \\ -0.8 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -1.2 \\ 1.2 \end{pmatrix}$
(40,100)	(40,101)	(40,101)	(40,100)	(42,102)	(42,100)	(41,100)	(40,100)
(30,120)	(30,121)	(33,120)	(30,120)	(30,120)	(31,121)	(30,120)	(30,121)
(50,150)	(49,150)	(50,150)	(50,150)	(50,150)	(49,150)	(50,151)	(50,150)
(80,160)	(77,160)	(76,161)	(80,160)	(80,160)	(80,160)	(81,161)	(80,160)
(120,170)	(120,172)	(120,170)	(120,170)	(120,169)	(120,170)	(119,170)	(120,170)

TABLE 5.4 – $n = 200$, $\rho_1 = \rho_2 = 0$, $\theta_1 = 1$ et $\theta_2 = 0$

ont des amplitudes du même signe, notre méthode ne détecte généralement pas les bons instants.

5.1.2.4 Modèle ARCH

On considère le modèle (1.1) lorsque $T(x) = 0$ et $V(x) = 0.5x$. Il s'agit d'un modèle ARCH. On étudie le cas d'un seul point de rupture. C'est-à-dire le cas où $k = 1$. On prend f la densité Gaussienne standard. Nous représentons la puissance du test à la Figure 5.17, pour $\alpha = 0.05$ et $\alpha = 0.01$.

On remarque que même dans le cas où on a une volatilité, c'est-à-dire le cas où la fonction V de modèle (1.1) n'est pas constante, notre test détecte bien les instants de rupture qui sont les temps donnant la plus grande puissance.

5.2.2.5 Une rupture connue contre la présence d'une rupture faible supplémentaire

Ici, nous traitons le cas de la présence d'une rupture contre l'hypothèse de la présence d'une rupture faible supplémentaire, c'est-à-dire le cas où le premier instant de rupture t_1 est connu et de deuxième point de rupture t_2 est inconnu. Plus particulièrement, nous considérons un échantillons X_1, X_2, \dots, X_n , issu du modèle

$$X_t = \begin{cases} \mu_1 + \varepsilon_t & \text{pour } 1 \leq t \leq t_1 \\ \mu_2 + \varepsilon_t & \text{pour } t_1 + 1 \leq t \leq t_2 \\ \mu_2 + \frac{\beta_1}{\sqrt{n}} + \varepsilon_t & \text{pour } t_2 + 1 \leq t \leq n, \end{cases}$$

où μ_1, μ_2, β_1 sont des constantes connues et ε_t est un bruit blanc gaussien standard. Cela correspond à notre problème de test pour $V(x) = 1$, $T(x) = \mu_1$ pour $1 \leq t \leq t_1$ et μ_2 pour $t_1 + 1 \leq t \leq n$, $\gamma_0 = (\mu_1, \mu_2, \mu_2)^\top$ et $\beta = (0, 0, \beta_1)^\top$, $\beta_1 \in \mathbb{R}$.

La Figure 5.1 (f) présente le chronogramme des données simulées dans le cas où f est la densité d'une loi normale centrée réduite. Les Figures 5.18 et 5.19 correspondent à la puissance du test pour $\alpha = 0.05$ et 0.01 respectivement. On remarque que la puissance du test est maximale pour la bonne valeur de t_2 quelque soit $\alpha \in [0, 1]$. Donc même dans ce cas, on obtient de bons résultats.

5.2.2.6 Fausse alarme

Une fausse alarme peut être considérée comme un signal annonçant une brève perturbation d'un phénomène qui évolue de manière stable, et qui retrouve sa stabilité après ce signal. Plusieurs travaux sont basés sur la détection de fausse alarme, comme par exemple, Ruggieri et Antonellis (2016) qui proposent une méthode pour détecter une fausse alarme par une approche bayésienne. Dans ce paragraphe, nous appliquons nos résultats à la détection de fausses alarmes. Il s'agit pour nous d'étudier l'aptitude de notre méthode à détecter les fausses alarmes. Pour cela, nous considérons $n = 200$ et deux sous-modèles de (5.1). Le premier avec $\rho_1 = \rho_2 = 0$, $\theta_1 = 1$ et $\theta_2 = 0$ et le deuxième avec $\rho_1 = 0.5$, $\rho_2 = 0$, $\theta_1 = 1$, $\theta_2 = 0$. La puissance dans ces deux cas est représentée dans les Figures 5.20 (a)-5.20 (d) pour $\alpha = 0.05$, et les Figures 5.21 (a)-5.21 (d) pour $\alpha = 0.01$. Par exemple pour la Figures 5.20 (a) et 5.20 (a), nous simulons une série des données présentant une fausse alarme à l'instant $t_1 = 60, 80, 120, 100, 140, 160$ pour $\beta = (0, -0.6)^\top, (0, 0.7)^\top, (0, -0.5)^\top, (0, -0.8)^\top, (0, 0.2)^\top, (0, 0.3)^\top$ respectivement. Nous obtenons les courbes respectivement en bleu, vert, rouge, violet, marron et jaune. On constate que l'instant auquel on a une fausse alarme présente une pente vertigineuse dans la courbe de la puissance du test.

5.3 Application aux données réelles

Dans ce paragraphe nous appliquons les résultats du chapitre 3 aux données d'anomalies de températures annuelles mondiales issues de la base de données de la National Oceanic and Atmospheric Administration (NOAA), et les données du COVID-19, issues du site web Worldometers. Pour les données du Covid-19 nous nous restreignons aux données françaises et nous considérons celles du taux de mortalité, celles du nombre de décès quotidiens et celles du nombre des nouveaux cas quotidiens. Les résultats présentés ici se trouvent dans Ngatchou-Wandji et Ltaifa (2021a).

5.3.1 Modélisation

D'après les graphiques des séries brutes que nous étudions, celles-ci présentent des tendances et ne semblent pas présenter de saisonnalité. Nos résultats ne peuvent pas s'appliquer directement à ces séries. Nous les décomposons comme suit :

$$Y_t = Z_t + X_t,$$

où X_t est une série stationnaire par morceaux de moyenne μ et de variance σ sur chaque "morceau" et Z_t représente la composante tendancielle supposée continue.

Il s'agit pour nous de considérer la série corrigée de la tendance, c'est-à-dire la série résiduelle qui exhibera les mêmes points de ruptures que la série brute.

Il y a plusieurs méthodes d'estimation de la tendance d'une série chronologique. Nous utilisons ici la méthode des moyennes mobiles, en particulier les moyennes mobiles arithmétiques d'ordre 5. Car ce sont elles qui nous ont semblé produire les meilleures estimations. Nous estimons donc Z_t par

$$\hat{Z}_t = \frac{1}{5} \sum_{j=-2}^2 Y_{t+j}.$$

Cette estimation est obtenue en utilisant la commande "ma" du logiciel R.

Ensuite, si on suppose qu'on a ℓ potentiels points de rupture qu'on note t_1, t_2, \dots, t_ℓ , tels que $t_1 < t_2 < \dots < t_\ell$, sur chaque intervalle $[t_{i-1}, t_i)$, où la série résiduelle semble stationnaire, nous ajustons à cette série un modèle de la forme $X_t = \mu + (\beta_i/\sqrt{n}) + \sigma_i \varepsilon_t$, $i \in \{1, \dots, k+1\}$, k le nombre des ruptures, n le nombre des données, $\mu + (\beta_i/\sqrt{n})$ et σ_i sont respectivement la moyenne et la variance de X_t sur l'intervalle $[t_{i-1}, t_i)$, $\beta_1 = 0$ et pour tout $i = 2, \dots, k+1$, $\beta_i \in \mathbb{R}$, et ε_t est un bruit blanc gaussien centré réduit. Cela correspond à notre problème de test pour $T(x) = \mu$, $V(x) = \sigma_i$ sur chaque intervalle $[t_{i-1}, t_i)$, $\gamma_0 = (\mu, \mu, \dots, \mu)^\top$, $\gamma_n = (\mu, \mu + (\beta_2/\sqrt{n}), \mu + (\beta_3/\sqrt{n}), \dots, \mu + (\beta_{k+1}/\sqrt{n}))^\top$ et $\beta = (0, \beta_2, \beta_3, \dots, \beta_{k+1})^\top$.

Nous commençons dans le paragraphe ci-dessous par l'étude des ruptures dans les données sur les anomalies climatiques annuelles.

5.3.2 Anomalie climatique annuelle

Nous prenons les données sur les anomalies climatiques annuelles, de la période de 1880 à 2015. La Figure 5.22 (a) présente le chronogramme des données brutes et semble exhiber 3 points de ruptures correspondants aux années 1910, 1944 et 1975. Nous désignons par points de ruptures ici les années qui semblent indiquer un changement (une brusque augmentation ou une diminution) dans la moyenne des données par rapport aux années précédentes. Nous voulons savoir si les années potentielles ruptures représentent

des points de ruptures ou si ces points sont dans leur voisinage. Nous commençons par estimer la tendance par la méthode de moyenne mobile. Cette estimation est représentée en rouge dans la Figure 5.22 (a). Puis, nous prenons la série résiduelle représentée dans la Figure 5.22 (b). En utilisant le test de normalité Shapiro-Wilk sur la série résiduelle et sur chaque morceau par l'utilisation de la commande `shapiro.test` sous R, nous retenons l'hypothèse de normalité. Nous appliquons les tests de Box-Pierce et Box-Ljung en utilisant la commande `Box.test` sous R sur ces données. Nous trouvons $p\text{-valeur}=0.03858$. Donc pour un seuil critique $\alpha = 0.05$, nous rejetons l'hypothèse iid des résidus. D'où l'hypothèse d'hétéroscédasticité. Ici nous prenons $k = 3$, $t_1 = 1910$, $t_2 = 1944$, $t_3 = 1975$, $\mu + (\beta_i/\sqrt{n})$ et σ_i sont respectivement la moyenne et la variance des données résiduelles sur l'intervalle $[t_{i-1}, t_i)$, pour $i = 1, 2, 3, 4$, où $t_0 = 1880$ et $t_4 = 2015$, $\beta_1 = 0$ et pour tout $i = 2, \dots, k + 1$, $\beta_i \in \mathbb{R}$ et $(\varepsilon_t)_t$ est un bruit blanc gaussien centré réduit.

Ensuite, nous calculons la puissance du test autour de chaque point t_i et nous prenons les années donnant la puissance maximale. Nous obtenons les points suivants : $\hat{t}_1 = 1915$, $\hat{t}_2 = 1944$ et $\hat{t}_3 = 1975$ représentés dans la Figure 5.22 (c). Ces points sont ceux considérés hormis le point \hat{t}_1 qui est proche de t_1 considéré. Ces points semblent coller à la réalité. En effet, dans le premier intervalle la tendance globale est à la baisse. Elle est à la hausse dans l'intervalle suivant, puis constante dans le suivant et à la hausse dans le dernier.

5.3.3 COVID-19 en France

Dans ce paragraphe nous étudions trois séries de données journalières relatives à la première vague du Covid-19 en France en 2020. Nous désignons par les numéros $1, 2, \dots, n$ les jours des données dans l'ordre. C'est-à-dire que : 1 correspond au premier jour, 2 au deuxième jour, \dots , n au dernier jour. Dans la première, la deuxième et la troisième applications, 1 correspond respectivement aux 15/02/2020, 27/02/2020 et 25/02/2020 et n correspond au 10/07/2020. Nous commençons d'abord par donner quelques éléments de la chronologie de cette première vague qui peuvent aider à l'interprétation des ruptures dans les séries étudiées.

- Le 14/02/2020 : Enregistrement du premier décès en France, celui d'un touriste chinois de 80 ans.
- Le 25/02/2020 : Enregistrement du premier décès français et cinq nouveaux cas.
- Le 15/03/2020 : Fermeture des lieux publics.
- Le 17/03/2020 : Mise en place du confinement.
- Le 11/05/2020 : Début de la phase 1 du dé-confinement.

- Le 02/06/2020 : Début de la phase 2 du dé-confinement.
- Le 22/06/2020 : Début de la phase 3 du dé-confinement.

• Taux de mortalité

La première série étudiée est celle du taux de mortalité dans la période du 15/02/2020 au 10/07/2020. La Figure 5.23 (a) correspond au chronogramme des données brutes. Ce graphique, semble présenter 5 potentiels points de ruptures correspondants aux jours $t_1 = 11, t_2 = 31, t_3 = 47, t_4 = 58$ et $t_5 = 110$, à compter à partir du 15/02. C'est-à-dire que les potentielles ruptures correspondent respectivement aux 25/02/2020, 16/03/2020, 01/04/2020, 12/04/2020 et 03/06/2020. Ici t_1, t_2, t_3 et t_4 nous semblent indiquer des instants de changement dans la moyenne de la série.

En utilisant le test de normalité Shapiro-Wilk sur la série résiduelle et sur chaque morceau par l'utilisation de la commande `shapiro.test` sous R, nous retenons l'hypothèse de normalité sur chaque morceau. L'utilisation des tests de Box et Pierce nous conduit à rejeter l'hypothèse iid. Cela nous permet de justifier la présence des points de ruptures dans les données. En t_3 le changement de la moyenne est significatif mais ce point peut-être une fausse alarme. Nous pouvons remarquer très facilement que la série résiduelle représentée dans la Figure 5.23 (b) exhibe les mêmes points de rupture que la série des données brutes. Donc après l'estimation de la tendance, nous ajustons à la série corrigée de la tendance au modèle décrit dans le paragraphe 5.3.1 pour $k = 5$, $t_1 = 11, t_2 = 31, t_3 = 47, t_4 = 58$ et $t_5 = 110$, $\mu + (\beta_i/\sqrt{n})$ et σ_i sont respectivement la moyenne et la variance des données résiduelle sur l'intervalle $[t_{i-1}, t_i)$, pour $i = 1, \dots, n$ avec $t_0 = 1$ et $t_6 = n$, $\beta_1 = 0$ et pour tout $i = 2, \dots, k + 1$, $\beta_i \in \mathbb{R}$ et ε_t est un bruit blanc gaussien centré réduit. Cet ajustement est dû au fait que p-valeur=0.002681 lorsque nous avons appliqué les tests de Box sur la série résiduelle.

Nous calculons la puissance du test autour des points t_i ci-dessus et nous prenons les années donnant la puissance maximale. Nous obtenons les points $\hat{t}_1 = 11, \hat{t}_2 = 32, \hat{t}_3 = 46, \hat{t}_4 = 60, \hat{t}_5 = 110$ à compter à partir du 15/02/2020. Ils sont représentés dans la Figure 5.23 (c). Nous remarquons que notre méthode fournie comme estimateurs des dates des ruptures, les ruptures potentielles. Pour interpréter ces estimations, nous dirons qu'autour du 25/02/2020, le taux de mortalité qui est à la baisse, va brusquement augmenter pour atteindre son pic le 15/03/2020 et chuter drastiquement le lendemain puis rester presque constant jusqu'au 30/03/2020, avant de remonter lentement jusqu'au 11/04/2020, puis redescendre lentement jusqu'au 03/06/2020 avant de se stabiliser jusqu'à la phase 3 du dé-confinement et même longtemps après.

- **Nombre de décès quotidiens**

Nous cherchons dans cette partie les points de rupture dans la série des données correspondant aux nombres de décès quotidiens du COVID-19 en France dans la période du 27/02/2020 au 10/07/2020. La Figure 5.24 (a) correspond au chronogramme des données brutes. Nous voulons savoir si les potentiels points de rupture dans cette série, tracés en vert dans la figure, représentent réellement des points de rupture. Ces points sont $t_1 = 31, t_2 = 38, t_3 = 55, t_4 = 86$ et $t_5 = 116$ correspondants respectivement aux dates suivantes : 28/03/2020, 04/04/2020, 21/04/2020, 22/05/2020 et 21/06/2020. Comme dans les cas précédents, nous estimons la tendance, puis nous ajustons à la série corrigée de la tendance le modèle décrit dans la section précédente, où 1 et n sont respectivement 27/02/2020 et 10/07/2020.

Cette série est représentée dans la Figure 5.24 (b). Nous calculons la puissance du test autour des t_i et nous prenons les dates donnant la plus grande puissance. Nous obtenons ainsi les dates : $\hat{t}_1 = 35, \hat{t}_2 = 40, \hat{t}_3 = 56, \hat{t}_4 = 88$ et $\hat{t}_5 = 117$ à compter à partir du 27/02/2020. Ces dates représentées dans la Figure 5.24 (c) correspondent respectivement aux 02/03/2020, 06/04/2020, 22/04/2020, 21/05/2020 et 25/06/2020. Les dates estimées sont différentes des potentielles dates de rupture, mais elles leur sont très proches, et les résultats obtenus sont plus proches de la réalité. En effet, autour du $\hat{t}_1 = 02/03/2020$, le nombre de décès augmente drastiquement, atteint son pic et redescend autour du $\hat{t}_2 = 06/04/2020$, puis oscille significativement jusqu'aux environs du $\hat{t}_3 = 22/04/2020$, et un peu moins significativement entre les première et deuxième phases du dé-confinement, jusqu'au $\hat{t}_4 = 21/05/2020$, date à partir de laquelle il se stabilise avant de se réduire considérablement à partir du $\hat{t}_5 = 25/06/2020$, peu après la troisième phase du dé-confinement.

- **Nouveaux cas quotidiens**

Nous prenons les données des nouveaux cas quotidiens du COVID-19 en France dans la période du 25/02/2020 au 10/07/2020 présentées dans la Figure 5.25 (a). On dénombre 6 potentiels points de ruptures, représentés en vert. Ces points sont $t_1 = 27, t_2 = 33, t_3 = 37, t_4 = 75, t_5 = 93$ et $t_6 = 100$ qui correspondent respectivement aux dates suivantes : 23/03/2020, 28/03/2020, 01/04/2020, 09/05/2020, 27/05/2020 et 03/06/2020.

Comme précédemment, nous estimons la tendance par la méthode de moyenne mobile d'ordre 5 et nous ajustons aux données résiduelles le modèle décrit dans le paragraphe 5.3.1 pour $k = 7$, $\mu + (\beta_i/\sqrt{n})$ et σ_i sont respectivement la moyenne et la variance des observations sur l'intervalle $[t_{i-1}, t_i)$, pour $i = 1, \dots, k + 1$, avec $t_0 = 1$ et $t_7 = n$, $\beta_1 = 0$ et pour tout $i = 2, \dots, k + 1$, $\beta_i \in \mathbb{R}$ et ε_t est un bruit blanc gaussien centré réduit.

Ensuite, nous calculons la puissance comme dans les cas précédents. Nous obtenons les dates suivantes : $\hat{t}_1 = 26$, $\hat{t}_2 = 34$, $\hat{t}_3 = 40$, $\hat{t}_4 = 80$, $\hat{t}_5 = 93$ et $\hat{t}_6 = 110$ à compter à partir du 25/02/2020. Ces dates sont représentées dans la Figure 5.25 (c). Elles correspondent respectivement aux 22/03/2020, 29/03/2020, 04/04/2020, 4/05/2020, 27/05/2020 et 13/06/2020. Les dates données par notre méthode sont différentes de celles que nous avons soupçonnées mais elles sont très proches.

5.4 Quelques comparaisons

5.4.1 Comparaison avec Horváth *et al.* (2020) et quelques méthodes de segmentation binaire

Le CUSUM test proposé par Horváth *et al.* (2020) est l'une des modifications du CUSUM test standard. C'est une nouvelle classe des statistiques du test de changement qui utilise un schéma de pondération et d'ajustement par le processus des sommes cumulées. Cette construction est inspirée par Rényi (1953) où la statistique considérée est la suivante :

$$D_n = D_n(t_n) = \max_{t_n \leq t \leq n-t_n} \left| \frac{1}{t} \sum_{s=1}^t X_s - \frac{1}{n-t} \sum_{s=t+1}^n X_s \right|,$$

avec X_1, \dots, X_n la série d'observations, n le nombre d'observations et t_n le paramètre d'ajustement.

Horváth *et al.* (2020) généralisent les statistiques de Rényi. Ils trouvent que cette statistique est plus efficace en terme de détection de rupture lorsqu'elle se produit dans les premières ou les dernières observations. Ils observent que le test basé sur cette statistique a la plus grande puissance.

Nous voulons dans ce paragraphe comparer nos méthodes notées NEW-S2 et NEW-S3 avec celles étudiées dans Horváth *et al.* (2020), notées ici SCUSUM (CUSUM test standard) et RCUSUM (CUSUM de Rényi), où le CUSUM test standard est le suivant :

$$CUSUM = \frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \left| \sum_{s=1}^t X_s - \frac{t}{n} \sum_{s=1}^n X_s \right|,$$

où X_1, \dots, X_n est la série d'observations. Pour appliquer la méthode RCUSUM, nous utilisons la commande `CUSUM.test` sous R.

Aussi, nous comparons les résultats donnés par nos méthodes aux résultats donnés par la méthode Wild Binary Segmentation (WBS) proposée par Fryzlewicz *et al.* (2014), la méthode NOT proposée par Baranowski *et al.* (2018) et la méthode Isolate-Detect (ID) proposée par Anastasiou et Fryzlewicz (2019). Pour plus de détails sur ces méthodes, voir

la section 2.3.3.

Pour cela, nous simulons une série d'observations générées par le modèle (5.1) pour $n = 200$, $\rho_1 = 0.5$, $\rho_2 = 0$, $\theta_1 = 1$ et $\theta_0 = 0$. C'est le modèle considéré dans la Figure 5.11 (b) et 5.11 (d). Nous commençons par le cas où les séries simulées ne présentent aucune rupture. Ici, alors que notre stratégie $S1$ n'a jamais trouvé de changement dans toutes les séries simulées, SCUSUM, RCUSUM, WBS et parfois NOT chacune ont trouvé un changement dans presque toutes les séries. ID a même détecté plus d'un changement dans certains d'entre eux.

Pour les données simulées avec un seul point de rupture, les estimations de localisation obtenues avec nos méthodes en utilisant la partie (ii) de la Remarque 2, étaient généralement plus précises (voir la Table 5.5). SCUSUM et RCUSUM n'ont pu fournir des estimations raisonnables que pour les changements de grande amplitude.

Le RCUSUM, proposé par Horváth *et al.* (2020) pour estimer les localisations des points de rupture dans les observations les plus anciennes et les plus récentes, est battu par nos méthodes sur l'exemple étudié. WBS, NOT et ID produisent généralement des séquences de localisations de points de rupture dont certains contiennent des valeurs proches des localisations correctes. Les résultats donnés par la Table 5.5 sont basés sur un seul ensemble de données simulées tandis que ceux de NEW-S2, SCUSUM et RCUSUM sont basés sur 5000 répliques. En effet, comme WBS, NOT et ID donnent parfois plus d'un point, il n'y avait aucun moyen d'évaluer leur moyenne car leur nombre était différent d'une réplique à l'autre.

Comme on peut le voir à partir du Table 5.6, dans le cas d'une seule rupture, pour de petites magnitudes ($|\beta_2| \leq 15$), les estimations de nos méthodes sont globalement plus proches des véritables localisations que les autres, même pour les échantillons aux extrémités. La même chose peut être dite pour les plus grandes amplitudes ($|\beta_2| > 15$). Pour deux ruptures, comme on peut le voir dans la Table 5.7, encore une fois pour les faibles amplitudes, bien qu'elles sous/surestiment parfois le nombre de ruptures, elles sont généralement capables de trouver au moins une, tandis que les autres méthodes sont tombées pour trouver n'importe qui. Pour des grandeurs importantes, nos méthodes sont généralement toujours meilleures.

5.4.2 Comparaison avec Xiong *et al.* (2015)

Xiong *et al.* (2015) étudient les séries hydrologiques multivariées non stationnaires relatives au fleuve Yangzi Jiang en Chine. Ils veulent détecter les ruptures dans les distributions

t exact	30	60	120	170
$(0, \beta_2)$	$(0, -0.5)$	$(0, 0.4)$	$(0, 0.5)$	$(0, -0.75)$
NEW-S2	30	60	120	170
SCUSUM	99	99	100	100
RCUSUM	91	60	143	126
WBS	71;76	17;55;72;124;170	—	25;61;65;134;141
NOT	71;76	17;55;72;114;124	—	61;65;132;141;175
ID	30;58;76;107;110	62;72;114;2124;140	96;101;106;122;129	127;132;141;167;182
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -5)$	$(0, 4)$	$(0, 5)$	$(0, -2)$
NEW-S2	30	60	120	171
SCUSUM	76	81	113	105
RCUSUM	29	63	80	44
WBS	31;44	71	123	152
NOT	31;44;175;182	71	32;84;100;112;158	152
ID	31;47;50;58;63;198	36;38;95;99;113	111;123;142;159	142;152;167;184
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -15)$	$(0, 14)$	$(0, 15)$	$(0, -12)$
NEW-S2	30	60	120	170
SCUSUM	43	64	119	153
RCUSUM	35	77	121	150
WBS	31	68;98;128;149;190	39;122	153
NOT	32;60;63;138;144	52;86;98;128;188	27;39;120	106;115;153
ID	32;37;63;96;144;179	45;52;86;98;128	112;115;123;166	115;153;177
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -25)$	$(0, 24)$	$(0, 25)$	$(0, -22)$
NEW-S2	30	60	120	170
SCUSUM	36	62	120	163
RCUSUM	38	61	121	166
WBS	32	64;129;164	122	150
NOT	14;30;41;122	64;129;164	92;120;126;148;158	150
ID	9;30;52;108;111;198	53;68;124;129;188	108;120;126;148;190	134;150;158;166;184
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -35)$	$(0, 34)$	$(0, 35)$	$(0, -32)$
NEW-S2	30	60	120	170
SCUSUM	34	61	120	166
RCUSUM	30	60	121	170
WBS	17;30	60;80	121;166;173	46;150
NOT	17;30	60;80	65;93;121;172;196	41;150
ID	30;85;128;158;177	21;60;125;156;183	59;67;76;150;188	110;129;150;197

TABLE 5.5 – Les estimations des localisations obtenues par NEW-S2, SCUSUM et RCUSUM sur la base de 5000 réplifications, et par WBS, NOT et ID sur la base d’une réplification pour les échantillons avec des zones de rupture connues.

t exact	30	60	120	170
$(0, \beta_2)$	$(0, -0.5)$	$(0, 0.4)$	$(0, 0.5)$	$(0, -0.75)$
NEW-S2	20,163	54	134,135	171
NEW-S3	26,46	65	116,146	55,65, 175
SCUSUM	67	66	151	75
RCUSUM	67	66	151	78
NOT	25	—	151	—
WBS	25	—	151	—
ID	25	97,104	2,16,151	—
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -5)$	$(0, 4)$	$(0, 5)$	$(0, -2)$
NEW-S2	31	52	121	166
NEW-S3	28,48,186	47,57	116	47,127,171
SCUSUM	42	70	121	62
RCUSUM	40	74	122	62
NOT	—	—	—	—
WBS	—	—	—	—
ID	46,187	—	108	—
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -15)$	$(0, 14)$	$(0, 15)$	$(0, -12)$
NEW-S2	37	61	123	174
NEW-S3	27	62	117	166
SCUSUM	23	64	117	165
RCUSUM	33	64	117	173
NOT	23	61	118	173
WBS	23	61	118	173
ID	23	32,43,61,187	117,191	148,173
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -25)$	$(0, 24)$	$(0, 25)$	$(0, -22)$
NEW-S2	34,94	62	128	170
NEW-S3	30,46,96	56	122	30,168
SCUSUM	40	67	120	167
RCUSUM	30	67	120	170
NOT	30	63	120	170
WBS	30	63	120	170
ID	2,11,30	63	39,40,120,128	171
t exact	30	60	120	170
$(0, \beta_2)$	$(0, -35)$	$(0, 34)$	$(0, 35)$	$(0, -32)$
NEW-S2	31	60	120	166
NEW-S3	30	60	116	166
SCUSUM	31	60	120	160
RCUSUM	31	60	120	168
NOT	31	60	120	170
WBS	31	60	120	170
ID	31	60,120,180,187	120	172

TABLE 5.6 – Les estimations des localisations obtenues par NEW-S2, NEW-S3, SCUSUM, RCUSUM, WBS, NOT et ID sur la base d'un échantillon avec une rupture.

(t_1, t_2) exact	(30,170)	(30,100)	(100,170)	(100,150)
	$(0, \beta_2, \beta_3) = (0, 0.2, 0.5)$			
NEW-S2	20,177	35,103	164	161
NEW-S3	29,166,186	28,98	31,101,166	27,96,150
SCUSUM	103	35	48	121
RCUSUM	103	35	48	121
NOT	—	—	—	—
WBS	—	—	—	—
ID	28	—	48,71	118,121
(t_1, t_2) exact	(30,170)	(30,100)	(100,170)	(100,150)
	$(0, \beta_2, \beta_3) = (0, 0.2, -0.3)$			
NEW-S2	176	21,85	179	114,135
NEW-S3	27,167,187	28,98	31,98,176	101,151
SCUSUM	57	93	88	163
RCUSUM	57	93	88	165
NOT	—	—	—	—
WBS	—	—	—	—
ID	—	—	54,61,81	161,180
(t_1, t_2) exact	(30,170)	(30,100)	(100,170)	(100,150)
	$(0, \beta_2, \beta_3) = (0, 15, -13)$			
NEW-S2	30,167	23,85	170	113,148
NEW-S3	33,170	33,96	102,166	105,164
SCUSUM	170	100	170	151
RCUSUM	170	100	170	151
NOT	33,170	30,100	106,170	99,151
WBS	33,170	30,100	106,170	99,151
ID	33,61,170	27,100,191	100,170	35,90,151
(t_1, t_2) exact	(30,170)	(30,100)	(100,170)	(100,150)
	$(0, \beta_2, \beta_3) = (0, 14, -3)$			
NEW-S2	20,180	35,86	90,180	103,141
NEW-S3	27,166,186	27,96	46,98,166	46,96,149
SCUSUM	169	100	99	103
RCUSUM	169	100	99	103
NOT	36,169	26,98	99,170	103,150
WBS	36,169	26,98	99,170	103,150
ID	36,169,190	26,98	57,99,154,170	45,103,141

TABLE 5.7 – Les estimations des localisations obtenues par NEW-S2, NEW-S3, WBS, NOT et ID sur la base d'un échantillon avec deux ruptures

marginales et les ruptures dans la structure de dépendance entre les variables individuelles. Les auteurs proposent 3 étapes pour détecter ces deux types de ruptures : la détection du point de changement dans les séries univariées individuelles, l'estimation des distributions marginales et la détection du point de rupture dans la structure d'indépendance par l'utilisation d'un test de type Cramer-Von-Mises et un test du rapport de vraisemblance. Ils trouvent à l'aide des méthodes de Monte Carlo que le test basé sur le rapport de vraisemblance est beaucoup plus puissant que la méthode de Cramer-Von-Mises pour détecter une rupture lorsqu'elle a lieu dans la dépendance structurelle. Ils concluent que la dernière méthode n'est pas adaptée à la série multivariée présentant un seul point de rupture caché dans la dépendance structurelle.

Xiong *et al.* (2015) appliquent cette étude à des données réelles. Ils considèrent le plus grand affluent de fleuve Yangzi Jiang, le fleuve Han Jiang, qui a un bassin versant de $159\,000\text{ km}^2$. Au cours de ces dernières décennies, plusieurs grands réservoirs ont été construits dans ce bassin. Le régime d'écoulement de la rivière est influencé par ces réservoirs. Xiong *et al.* (2015) font cette étude en considérant la série des crues tri-variées composée de la AMMDD (annual maxima daily discharge) (Q_1), du volume annuel maximal des crues sur 3 jours (V_3) et du volume maximal annuel des crues sur 15 jours (V_{15}), observés à la station hydrologique d'Ankang pendant la période 1950-2011. Cette station a un bassin versant de $38\,600\text{ km}^2$, où deux grands réservoirs nommés Ankang et Shiquan sont construits comme le montre la Figure 5.26.

La construction de ce dernier réservoir a commencé en 1971 et a été terminée en 1975. La construction du réservoir d'Ankang a commencé en 1978, il a été terminée en 1989. Après un an environ, en décembre 1990, le réservoir a commencé à stoker de l'eau jusqu'en 1992. En 1995, tout le projet est terminé. La capacité de ce réservoir est beaucoup plus grande que celle du réservoir Shiquan. Selon Jiang *et al.* (2015), le régime d'écoulement de la rivière de Haut-Han jiang, observé à la station hydrologique d'Ankang, est principalement influencé par le réservoir d'Ankang. Xiong *et al.* (2015) trouvent que les séries univariées Q_1 , V_3 et V_{15} de cette rivière présentent des points de changement significatifs en raison de la construction de ce dernier réservoir.

Dans ce paragraphe, nous comparons nos résultats aux résultats trouvés par Xiong *et al.* (2015). Pour cela, nous prenons les mêmes données qu'eux. C'est-à-dire les données des crues univariées composées de la AMMDD (Q_1), du volume annuel maximal des crues sur 3 jours (V_3) et du volume maximal annuel des crues sur 15 jours (V_{15}) de la rivière Hanjiang en Chine de la période de 1950 à 2011.

Les données brutes des séries Q_1 , V_3 et V_{15} sont représentées respectivement dans la Figure 5.27 (a) et la Figure 5.28 (a). Nous observons que chacune de ces séries admet une composante tendancielle. Pour cette raison, nous procédons comme dans le paragraphe

5.3.1. Nous estimons la tendance par la méthode de moyenne mobile d'ordre 5. Ensuite, nous prenons la série résiduelle correspondant à chaque série des données représentée dans la Figure 5.26 (b) et la Figure 5.28 (b) respectivement. Nous observons qu'il y a au moins deux points de rupture dans ces séries et non un seul comme trouvé par Xiong *et al.* (2015).

Nous commençons par l'étude de la série Q_1 . Nous ajustons la série corrigée de la tendance au modèle suivant :

$$X_t = \begin{cases} \mu + \frac{\beta_1}{\sqrt{n}} + \sigma_1 \varepsilon_t & \text{pour } 1950 \leq t \leq t_1 \\ \mu + \frac{\beta_2}{\sqrt{n}} + \sigma_2 \varepsilon_t & \text{pour } t_1 + 1 \leq t \leq t_2 \\ \mu + \frac{\beta_3}{\sqrt{n}} + \sigma_3 \varepsilon_t & \text{pour } t_2 + 1 \leq t \leq 2011, \end{cases} \quad (5.2)$$

où $t_1 = 1987$ et $t_2 = 1997$ sont de potentiels points de rupture, $\mu + (\beta_i/\sqrt{n})$ et σ_i sont respectivement la moyenne et la variance des données résiduelles sur l'intervalle $[t_{i-1}, t_i)$, pour $i = 1, 2, 3$ où $t_0 = 1950$ et $t_3 = 2011$, $\beta_1 = 0$ et pour tout $i = 2, 3$, $\beta_i \in \mathbb{R}$, et ε_t est un bruit blanc gaussien centré réduit. Il est à noter que $t_1 = 1987$ est le seul point de rupture estimé par Xiong *et al.* (2015).

Puis, nous calculons la puissance du test autour de chacun de ces t_i , et nous prenons les années donnant la puissance maximale. Notre méthode fournit les points suivants : $\hat{t}_1 = 1986$, $\hat{t}_2 = 1997$ représentés dans la Figure 5.27 (c). Le premier point trouvé est très proche de celui trouvé par Xiong *et al.* (2015). D'après la courbe des observations, à partir de t_2 la tendance est à la hausse. On constate donc qu'il y a un autre point de rupture et notre méthode l'a bien estimé. Ce point peut être dû à la fin du projet de construction des réservoirs en 1995.

Revenons maintenant aux séries V_3 et V_{15} . Nous ajustons chaque série corrigée au modèle (5.5) avec $t_1 = 1985$ et $t_2 = 1997$ les deux potentiels points de rupture, $\mu + (\beta_i/\sqrt{n})$ et σ_i sont respectivement la moyenne et la variance des données résiduelles sur l'intervalle $[t_{i-1}, t_i)$, pour $i = 1, 2, 3$ où $t_0 = 1950$ et $t_3 = 2011$, $\beta_1 = 0$ et pour tout $i = 2, 3$, $\beta_i \in \mathbb{R}$, et $(\varepsilon_t)_t$ est un bruit blanc gaussien centré réduit. Cela peut être justifié en utilisant les commandes Box.test et shapiro.test sous le logiciel R.

Nous calculons la puissance du test autour de chaque point t_i et nous prenons les années donnant la puissance maximale. Notre méthode nous donne les estimations suivantes : $\hat{t}_1 = 1986$ et $\hat{t}_2 = 1997$ représentés dans la Figure 5.28 (c). Comme pour le cas de la série Q_1 , le premier point trouvé est proche de celui trouvé par Xiong *et al.* (2015). D'après la courbe des observations, à partir de t_2 la tendance est à la hausse. On constate donc qu'il y a un autre point de rupture qui est toujours 1997 et notre méthode l'a bien estimé.

Finalement, les résultats trouvés par notre méthode semblent bien estimer les loca-

Méthodes	Q_1	V_3	V_{15}
NEW-S2	1986, 1997	1986, 1997	1986, 1997
CH	1987	1985	1985
NOT	1984	1985	1984
WBS	1984	1985	1984, 2009
ID	1984	1985	1984, 2009

TABLE 5.8 – Table de comparaison

lisations des points de ruptures dans les 3 séries de données, et il nous semble que les points trouvés sont plus proches de la réalité que ceux de Xiong *et al.* (2015). En effet, en utilisant notre méthode, nous pouvons trouver le point trouvé par eux, plus un autre point de rupture dans les données.

Nous essayons aussi d'appliquer les méthodes NOT, WBS et ID citées plus haut à ces données pour faire une comparaison des résultats trouvés. Ces résultats sont affichés dans la Table 5.8, où CH est la méthode proposée par Xiong *et al.* (2015) et NEW-S2 est la nôtre. Nous observons que pour les séries des données Q_1 et V_{15} les méthodes NOT, WBS et ID détectent le même point de rupture 1984. Pour la série V_3 , les 4 méthodes autres que NEW-S2 détectent le même point, 1985. Mais les deux dernières détectent un point de plus qui est 2009. Pour les séries Q_1 et V_3 , seule la méthode NEW-S2 détecte un deuxième point de rupture. Pour la série V_{15} , le premier instant de rupture détecté par WBS et ID est très proche de ceux détectés par CH et de premier instant détecté par NEW-S2. Mais le deuxième point estimé par cette dernière est plus proche de 1995 (l'année dans laquelle le projet est fini) que celui estimé par WBS et ID. À notre avis, la méthode NEW-S2 est la plus proche de la réalité.

5.4.3 Comparaison avec Bardet et Kengne (2014) et Fryzlewicz et Rao (2014)

Dans la section 2.3 du chapitre 2, nous avons cité Bardet et Kengne (2014) parmi les références portant sur les tests d'existence de point de rupture. Dans cet article, les auteurs supposent que la suite d'observations $(X_1, \dots, X_n) \in M_{\mathbb{N}}(M_{\theta_0^*}, f_{\theta_0^*})$. C'est-à-dire $X = (X_1, \dots, X_n)$ satisfait la relation (2.1), avec $\theta = \theta_0^*$ et $T = \mathbb{N}$. Ils ont supposé que le paramètre θ_0^* est inconnu. Ils considèrent le problème de test suivant :

H_0 : θ_0^* est constante sur les observations $X_1, \dots, X_n, X_{n+1}, \dots$, c'est-à-dire $(X_n)_{n \in \mathbb{N}}$
 H_1 : Il existe $K^* > n$, $(\theta_0^*, \theta_1^*) \in \Theta^2$ tel que $\theta_0^* \neq \theta_1^*$

tel que $(X_1, \dots, X_{K^*}) \in M_{\{1, \dots, K^*\}}(M_{\theta_0^*}, f_{\theta_0^*})$ et $(X_{K^*+n})_{n \in \mathbb{N}} \in M_{\{K^*+1, \dots\}}(M_{\theta_1^*}, f_{\theta_1^*})$.

Bardet et Kengne (2014) font cette étude en utilisant le test du rapport de vraisemblance pour détecter la localisation du point de rupture qui est K^* . Puis, ils appliquent ces résultats théoriques à des données réelles.

Dans ce paragraphe, nous prenons les mêmes données. Il s'agit des valeurs de clôture quotidiennes de l'indice boursier FTSE 100, indice boursier des 100 sociétés britanniques les plus capitalisées listées dans London Stock Exchange du 02/01/2004 au 30/04/2013. Elles sont représentées dans la Figure 5.29 (a). Ces données sont étudiées par Bardet et Kengne (2014) et Fryzlewicz et Rao (2014).

Bardet et Kengne (2014) utilisent une approche en-ligne pour détecter les points de ruptures. Ils considèrent la série log-FTSE 100 et ils supposent que les données sont respectivement X_1, X_2, \dots, X_N , où N est le nombre d'observations. Ils procèdent comme suit :

- i. On choisit les observations historiques dans la première partie des données par l'application d'un test rétrospectif (voir par exemple Kengne (2012)) sur toutes les observations, si un changement est détecté, on divise les observations en deux parties et on applique le test rétrospectif sur la première partie. Cette procédure est répétée jusqu'à obtention des données historiques X_1, X_2, \dots, X_n .
- ii. On commence à partir de l'observation numéro $n + 1$, lorsqu'on s'arrête à l'instant \tilde{t}_1 par exemple, on applique le test rétrospectif aux données $X_1, X_2, \dots, X_{\tilde{t}_1}$. Cela détectera et estimera, selon Bardet et Kengne (2014), le premier instant de rupture dans les données (par exemple \hat{t}_1).
- iii. On prend les données historiques (y compris les données $X_{\tilde{t}_1}, X_2, \dots, X_{\tilde{t}_1}$) et on retourne à l'étape précédente. On répète la procédure jusqu'à ce qu'on arrive à la dernière observation X_N .

Les instants de ruptures trouvés par Bardet et Kengne (2014) pour cette série de données sont :

- $\hat{t}_1 \simeq 04/06/2007$ correspondant selon eux à la période juste avant le début de la crise des subprimes aux États-Unis.
- $\hat{t}_2 \simeq 06/09/2008$ correspondant selon eux à la faillite de Lehman Brothers.
- $\hat{t}_3 \simeq 02/01/2009$ correspondant selon eux à l'intervention des gouvernements mondiaux pour résoudre la crise financière.

- $\hat{t}_4 \simeq 30/6/2009$, $\hat{t}_5 \simeq 15/07/2011$ et $\hat{t}_6 \simeq 03/01/2012$ qui indiquent encore selon eux les périodes de turbulences de la Grèce en 2010 et la crise de la dette européenne en 2012.

Fryzlewicz et Rao (2014) proposent une méthode rapide, performante et traitée théoriquement pour détecter plusieurs points de rupture dans la structure d'un modèle ARCH pour les rendements financiers avec des valeurs de paramètres constantes par morceaux. Leur méthode est appelée BASTA (binary segmentation for transformed auto-regressive conditional heteroscedasticity). Cette méthode se déroule en deux étapes. La première est la transformation de processus et la deuxième est la segmentation binaire. La première consiste à décorrélérer le processus original et alléger ses queues. La deuxième estime systématiquement les ruptures. Ils proposent et justifient deux transformations particulières et ils appliquent leur méthode aux données de l'indice Financière Times Stock Exchange (FTSE) 100. Ils utilisent une approche hors-ligne pour estimer les ruptures. Plus précisément, soit $(X_t)_t$ un processus généré par le modèle ARCH(p). L'algorithme BASTA proposé par Fryzlewicz et Rao (2014) est le suivant :

- Étape 1 : On prend $U_t = g(X_t, X_{t-1}, \dots, X_{t-\tau})$, où le choix de g et τ est bien détaillé dans la section 4 de l'article. Le processus U_t est choisi de manière à contenir des informations au cours du temps sur les paramètres changeants de X_t et les points de ruptures correspondants.
- Étape 2 :
 - i. On commence par prendre $(j, l) = (1, 1)$. Soit $s_{j,l} = 0$ et $u_{j,l} = T - 1$, où $T - 1$ est le nombre d'observations.
 - ii. On note $n = u_{j,l} - s_{j,l} + 1$. On calcule

$$\tilde{U}_{s_{j,l}, u_{j,l}}^b = \frac{\sqrt{u_{j,l} - b}}{\sqrt{n(b - s_{j,l} + 1)}} \sum_{t=s_{j,l}}^b U_t - \frac{\sqrt{b - s_{j,l} + 1}}{\sqrt{n(u_{j,l} - b)}} \sum_{t=b+1}^{u_{j,l}} U_t$$

pour tout $b \in (s_{j,l}, u_{j,l})$. On prend $b_{j,l} = \arg \max_b |\tilde{U}_{s_{j,l}, u_{j,l}}^b|$.

- iii. Pour un seuil donné \tilde{b}_T , si $|\tilde{U}_{s_{j,l}, u_{j,l}}^{b_{j,l}}| < \tilde{b}_T$, alors l'algorithme s'arrête dans l'intervalle $(s_{j,l}, u_{j,l})$. Sinon, on ajoute $b_{j,l}$ à l'ensemble des points de rupture estimés et
 - a) On prend $(j_0, l_0) = (j, l)$ et $(s_{j+1, 2l-1}, u_{j+1, 2l-1}) := (s_{j,l}, b_{j,l})$. On met à jour $j := j + 1$, $l := 2l - 1$ et on retourne à l'étape ii.
 - b) On rappelle que $(j, l) = (j_0, l_0)$ stocké à l'étape a). On prend $(s_{j+1, 2l}, u_{j+1, 2l}) := (b_{j,l} + 1, u_{j,l})$. On met à jour $j := j + 1$, $l := 2l$ et on retourne à l'étape ii.

Fryzlewicz et Rao (2014) prennent les mêmes données réelles que Bardet et Kengne (2014) pour appliquer leur méthode, mais du 27/07/2005 au 13/07/2009, soit environ 4 années d'échange. Les ruptures estimées par eux sont :

- $\hat{t}_1 = 467$, correspondant au 5 juin 2007. Selon eux, l'été 2007 est largement considéré comme le début de la crise des hedge funds hypothécaires subprime, la grande banque d'investissement Bear Stearns révélant, en juillet 2007, que leurs deux subprime hedge les fonds avaient perdu presque toute leur valeur.
- $\hat{t}_2 = 773$, correspondant au 18 août 2008. Selon eux, il est probablement prudent d'attribuer ce point de changement estimé à l'effondrement de Lehman Brothers, une grande entreprise de services financiers.
- $\hat{t}_3 = 850$, correspondant au 4 décembre 2008. Selon eux, bien qu'il soit difficile d'attribuer cette date à un événement précis, la fin de l'année 2008 a été l'époque où les gouvernements, les banques nationales et les institutions internationales comme International Monetary Fund ont annoncé et ont commencé à mettre en œuvre une gamme de mesures financières pour aider l'économie mondiale en difficulté.

Pour appliquer notre méthode à ces données, nous prenons la série résiduelle représentée dans la Figure 5.29 (b). Elle présente les huit instants potentiels de ruptures suivants : $t_1 = 605, 810, 910, 1020, 1210, 1310, 1625$ et 1980 , à compter à partir du 02/01/2004, en prenant compte qu'il n'y a pas des données les samedis et les dimanches. C'est-à-dire que les points de ruptures potentiels correspondent respectivement aux 25/05/2006, 15/03/2007, 08/08/2007, 15/01/2008, 14/10/2008, 5/03/2009, 20/05/2010 et 29/09/2011. Les dates estimées par notre méthode sont : 01/06/2006, 19/03/2007, 02/08/2007, 16/01/2008, 9/10/2008, 11/03/2009, 27/05/2010 et 29/09/2011 qui sont très proches des indices potentiels de rupture. Ces dates sont représentées dans la Figure 5.29 (c). Donc notre méthode confirme les localisations de ces ruptures.

Ces localisations des points de ruptures trouvés par notre méthode peuvent être expliquées par les raisons suivantes :

- $\hat{t}_1 \simeq 25/05/2006$ correspondant à la période pendant laquelle le prêteur de subprimes Ameriquest annonce qu'il supprimera 3800 emplois, fermera ses 229 succursales de vente en détail et s'appuiera plutôt sur le web.
- $\hat{t}_2 \simeq 02/04/2007$ correspond à la période où la société New Century Financial, le plus grand prêteur subprimes américain, dépose une demande de faillite au titre du chapitre 11, où le chapitre 11 est une loi sur les faillites des États-Unis. Il est

accessible à toutes les entreprises, qu'elles soient organisées en société ou en entreprise individuelle, et aux particuliers, bien qu'elle soit principalement utilisée par les personnes morales.

- $\hat{t}_3 \simeq 19/07/2007$ tombe dans la période où le plus vieil indice boursier de New York et le plus vieil indice boursier du monde Dow Jones Industrial Average a clôturé plus que 14 000 pour la première fois.
- \hat{t}_4 est proche de 24/01/2008 la période pendant laquelle la National Association of Realtors (NAR) annonce que 2007 a connu la plus forte baisse des ventes de maisons existantes depuis 25 ans. Elle est proche aussi du 27/02/2008 période où le marché des titres à taux d'adjudication se fige et les investisseurs refusent d'enchérir.
- $\hat{t}_5 \simeq 11/10/2008$ la période où le Dow Jones Industrial Average plafonne sa pire semaine de son histoire avec sa journée de volatilité la plus élevée jamais enregistrée en 112 ans d'histoire. Au cours des huit derniers jours de bourse, le DJIA a chuté de 22% au milieu des craintes d'une aggravation de la crise du crédit et de la récession mondiale. Les pertes de papier sur les actions américaines s'élèvent désormais à 8,4 billions de dollars par rapport aux sommets du marché de l'année précédente.
- \hat{t}_6 indique la période dans laquelle l'indice Dow Jones tombe à 6440, une baisse en pourcentage dépassant le rythme de la baisse du marché pendant la grande dépression.
- Les dates $\hat{t}_7 = 27/05/2010$ et $\hat{t}_8 = 29/09/2011$ sont proches des deux dernières ruptures trouvées par Bardet et Kengne (2014). Elles indiquent respectivement la période de turbulences en Grèce en 2010, et la période de la crise de la dette européenne.

Nous n'avons pas trouvé les mêmes points de ruptures que Bardet et Kengne (2014) et Fryzlewicz et Rao (2014), mais nous avons trouvé plus de points et les deux derniers sont proches des deux derniers trouvés par Bardet et Kengne (2014).

5.4.4 Comparaison avec Fotopoulos *et al.* (2009)

Fotopoulos *et al.* (2009) considèrent une série d'observations gaussiennes multivariées et indépendantes présentant un changement dans la moyenne ou dans la matrice de covariance. Ils utilisent l'estimation classique d'un point de rupture dans les données qui est $[n\hat{\rho}_n]$, où n est le nombre d'observations et $\hat{\rho}_n$ est un réel compris entre 0 et 1. Puis, ils

expriment $[n\hat{\rho}_n] - [n\rho_n]$ en utilisant le logarithme du maximum de vraisemblance noté ξ_n dans l'article, où $[n\rho_n]$ est le point de rupture exact inconnu. Ensuite, ils étudient les propriétés asymptotiques de ξ_n . Plus précisément, l'estimateur considéré est le suivant :

$$\xi_n = \hat{\tau}_n - \tau_n = [n\hat{\rho}_n] - [n\rho] = \arg \max_{-[n\rho]+1 \leq j \leq n-[n\rho]-1} \sum_{i=1}^{[n\rho]+j} a(Y_{n,i}),$$

où a est le logarithme du rapport de vraisemblance et $Y_{n,i}$ est l'observation numéro i . Fotopoulos *et al.* (2009) appliquent ces résultats théoriques à des données réelles. Ils considèrent les données journalières des rendements logarithmiques sur deux actions de NYSE, nommées Verizon communication (VZ) et Belle South corporation (BLS) au cours de l'année 2002. Pour cette étude, ils supposent dans un premier temps que les données suivent une série gaussienne bivariée indépendante. La détection des points de ruptures pour ce type de données est développée par Csörgö et Horváth (1997).

Fotopoulos *et al.* (2009) appliquent le Théorème 13.1 de cet article, dans lequel on trouve l'expression de la statistique de test du rapport de vraisemblance, pour confirmer l'existence des points de rupture dans les données considérées.

Dans un second temps, ils appliquent leur méthode à ces données pour estimer les localisations des points de ruptures. Ils trouvent $[n\rho] = 46.84 \simeq 47$ pour un seuil $\alpha = 0.1$ qui correspond au 08/03/2002 et $[n\rho] = 38.92 \simeq 39$ pour un seuil $\alpha = 0.05$ qui correspond au 26/02/2002.

Enfin, ils concluent que le changement dans les paramètres et la volatilité de ces séries au cours de la période du 26/02/2002 à 14/05/2002 est due à une faute professionnelle à partir de 07/03/2002 de l'entreprise Woldcom Inc qui était une importante société de communication à l'époque. Elle avait été incapable de résister après la démission de son PDG Bernard Ebbers le 30/04/2002. Selon Fotopoulos *et al.* (2009), les problèmes de Woldcom ont influencé les actions de communication à NYSE, y compris VZ et BLS.

Dans notre étude, nous prenons les données de VZ au cours de l'année 2002 et nous essayons de lui appliquer notre méthode afin de comparer nos résultats à ceux trouvés par Fotopoulos *et al.* (2009). La Figure 5.30 (a) correspond au chronogramme des données brutes. Elle présente 5 potentiels points de rupture. Comme dans les applications précédentes, nous estimons la tendance par la méthode des moyennes mobile d'ordre 5. Puis, nous considérons la série corrigée de la tendance. Elle est représentée dans la Figure 5.30 (b). Les potentielles dates de rupture sont 05/02/2002, 21/02/2002, 08/05/2002, 19/07/2002 et 02/10/2002 et les dates trouvées par notre méthode, représentées dans la Figure 5.30 (c), sont les suivantes :

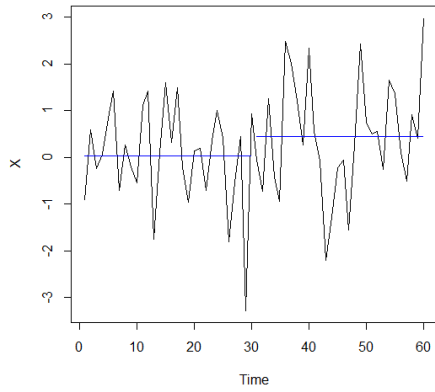
- $\hat{t}_1 = 08/02/2002$ et $\hat{t}_2 = 27/02/2002$ indiquent la période dans laquelle les jeux

olympiques d'hiver se tiennent à Salt Lake City, Utah du 08/02 au 24/02. Dans ces jeux, les États-Unis remportent 10 médailles d'or, 13 d'argent et 11 de bronze.

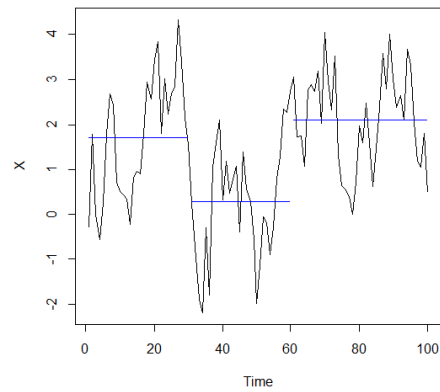
- $\hat{t}_3 = 08/05/2002$: C'est proche de 10/05/2002 la période pendant laquelle l'agent du FBI Robert Hanssen est condamné à la réclusion à perpétuité sans possibilité de libération conditionnelle pour avoir vendu des secrets américains à Mascau pour 1.4 million de dollars en espèces et en diamants.
- $\hat{t}_4 = 17/07/2002$: C'est proche de 21/07/2002 période au cours de laquelle le géant des télécommunications Woldcom dépose la protection contre les faillites du chapitre 11, le plus grand dépôt de ce type dans l'histoire des États-Unis.
- $\hat{t}_5 = 04/10/2002$: C'est proche de 02/10/2002 où les attaques de tireurs d'élite de Betway commencent par 5 fusillades dans le comté de Mantgomery, dans le Maryland et le congrès des États-Unis adopte une résolution commune qui autorise le président à utiliser les forces armées américaines comme il le juge nécessaire et approprié, contre l'Irak.

Tous ces événements influent sur les actions de communications comme VZ et BLS.

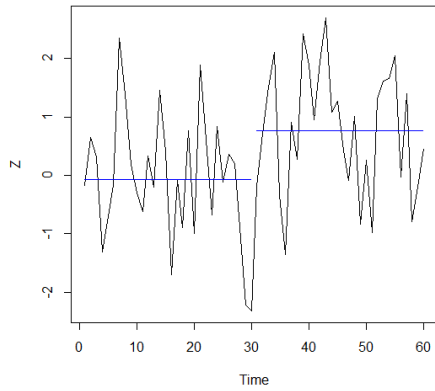
En comparaison de nos résultats à ceux de Fotopoulos *et al.* (2009), nous trouvons leurs conclusions convaincantes. Cependant, il nous semble que notre méthode fait un peu mieux. En effet, elle détecte plusieurs autres ruptures dans les données qui peuvent avoir été causées par des événements bien identifiés.



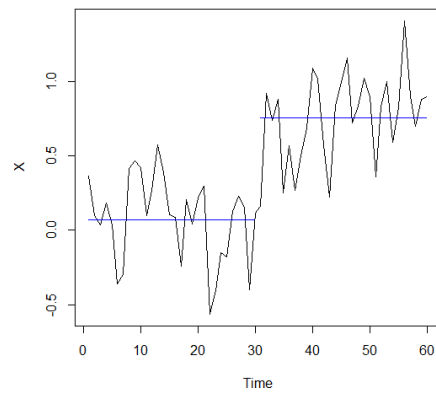
(a) $k = 1, t_1 = 30, \rho_1 = 0.5, \theta_1 = 1, \theta_2 = 0$



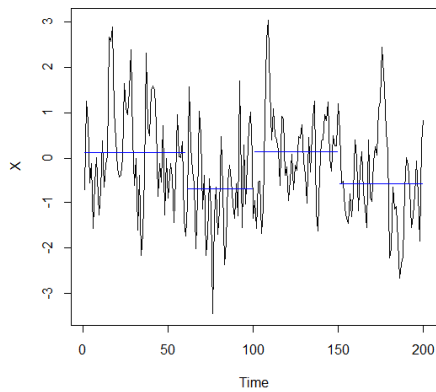
(b) $k = 2, t_1 = 30, t_2 = 60, \rho_1 = 0.75, \theta_1 = 1, \theta_2 = 0$



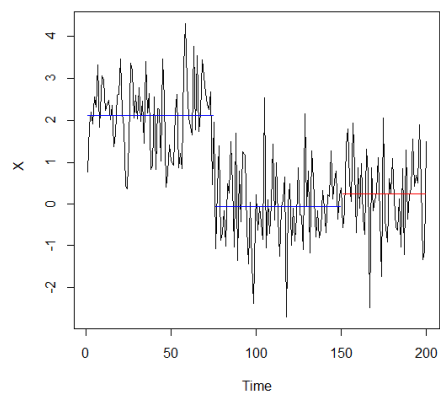
(c) $k = 1, t_1 = 30, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$



(d) $k = 1, t_1 = 30, \rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$



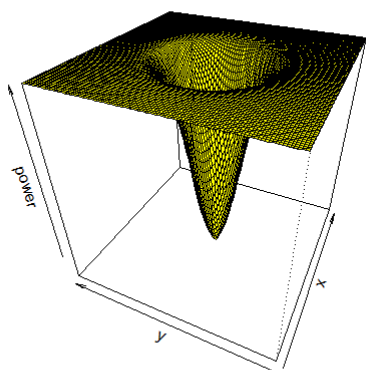
(e) $k = 3, t_1 = 60, t_2 = 100, t_3 = 150, \rho_1 = .5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$



(f) $k = 2, t_1 = 75, t_2 = 150, \mu = .5, \mu_1 = 1.5, \mu_2 = -0.5$

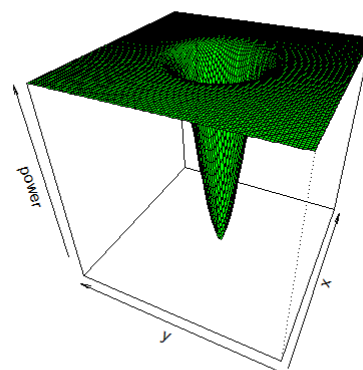
FIGURE 5.1 – Chronogrammes des observations

Gaussian noise



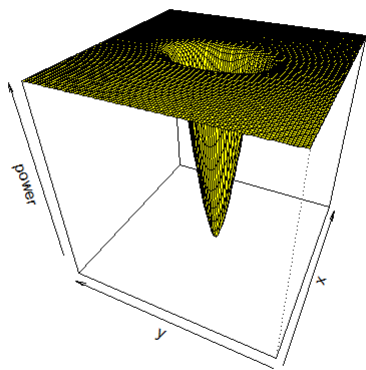
(a) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_0 = 0$

Exponential noise



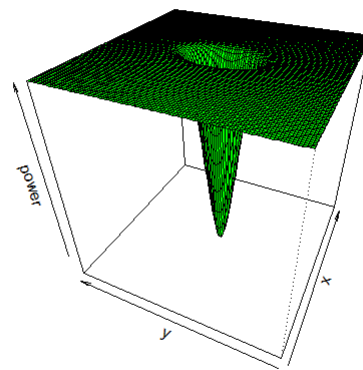
(b) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_0 = 0$

Gaussian noise



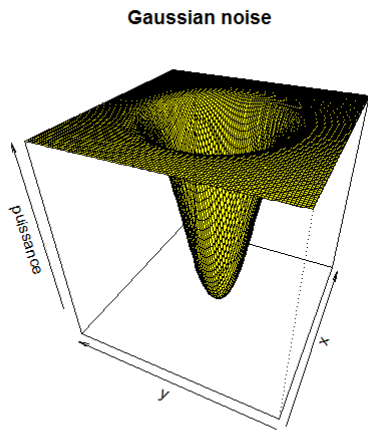
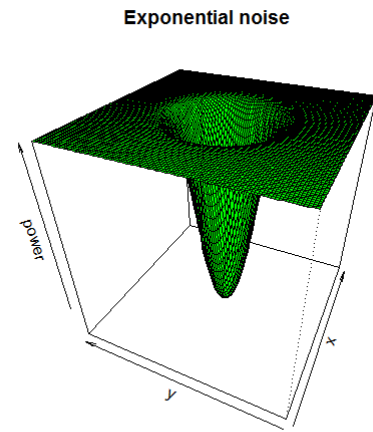
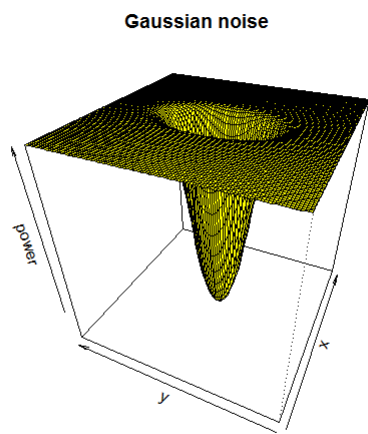
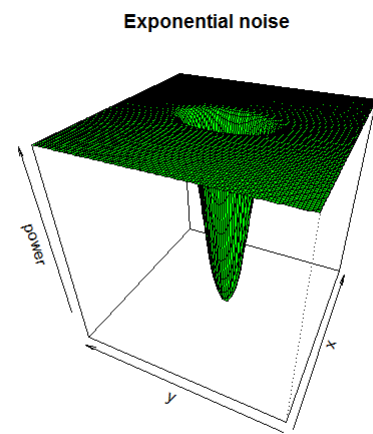
(c) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

Exponential noise

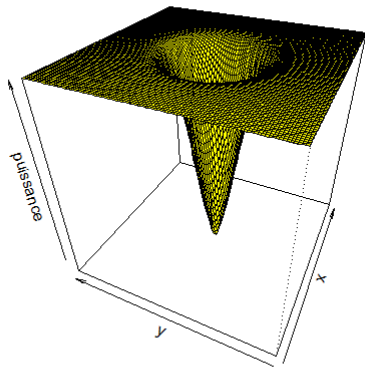


(d) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

FIGURE 5.2 – Puissance locale du test pour $k = 1$ et $\alpha = 0.05$.

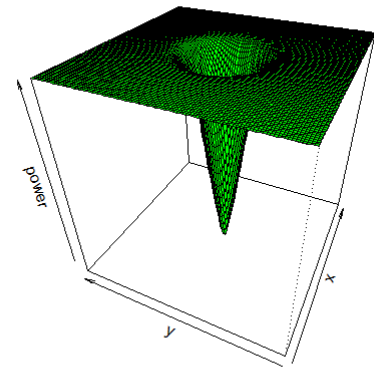
(a) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_0 = 0$ (b) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_0 = 0$ (c) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$ (d) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$ FIGURE 5.3 – Puissance locale du test pour $k = 1$ et $\alpha = 0.01$.

Gaussian noise



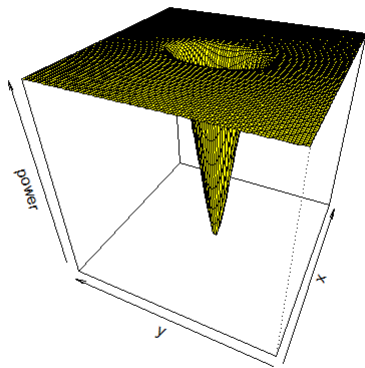
(a) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_0 = 0$

Exponential noise



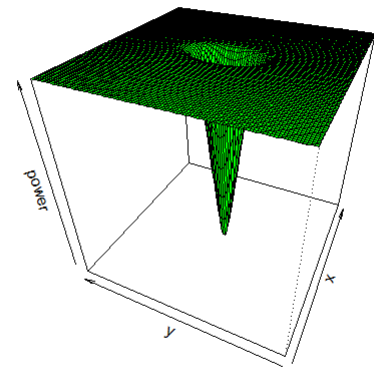
(b) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_0 = 0$

Gaussian noise



(c) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

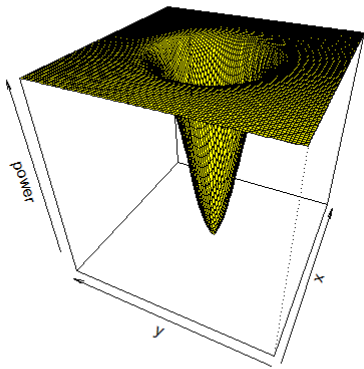
Exponential noise



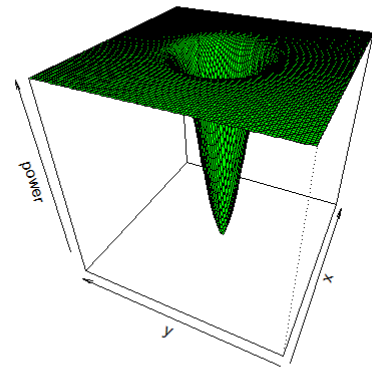
(d) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

FIGURE 5.4 – Puissance locale du test pour $k = 1$ et $\alpha = 0.1$.

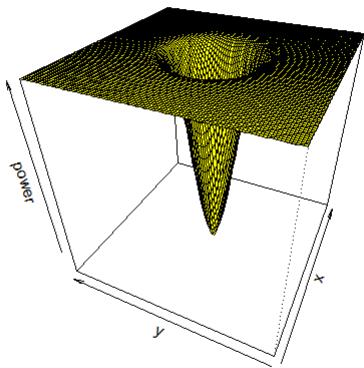
Gaussian noise

(a) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

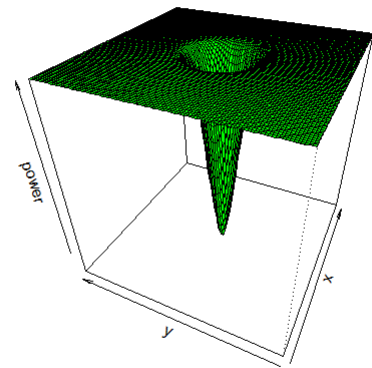
Exponential noise

(b) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

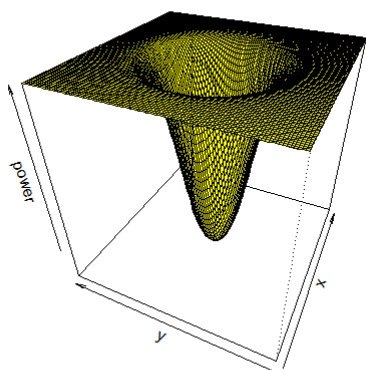
Gaussian noise

(c) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

Exponential noise

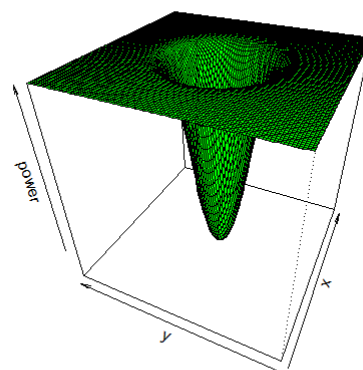
(d) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$ FIGURE 5.5 – Puissance locale du test pour $k = 1$ et $\alpha = 0.05$.

Gaussian noise



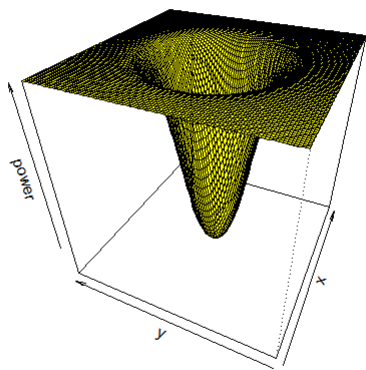
(a) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

Exponential noise



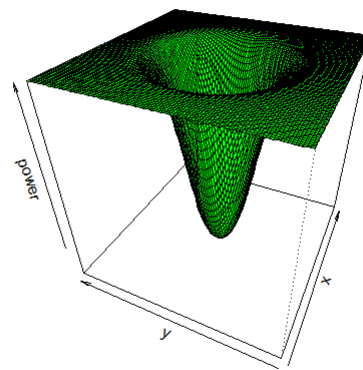
(b) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

Gaussian noise



(c) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

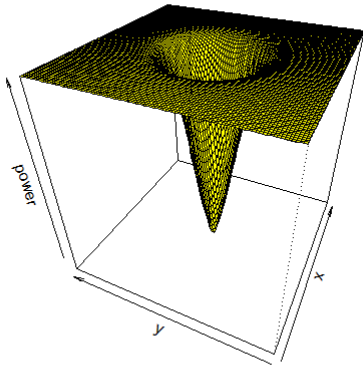
Exponential noise



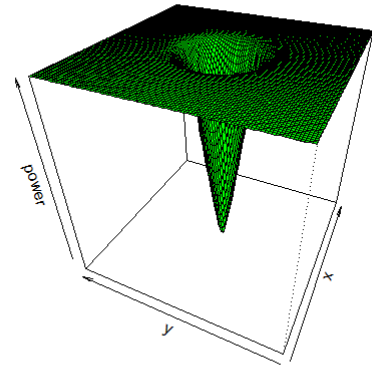
(d) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

FIGURE 5.6 – Puissance locale du test pour $k = 1$ et $\alpha = 0.01$.

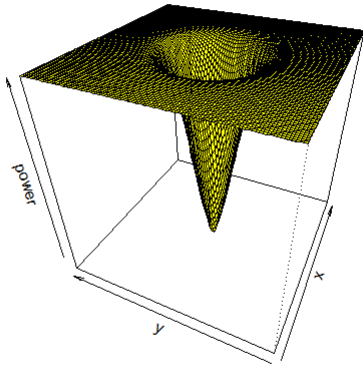
Gaussain noise

(a) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

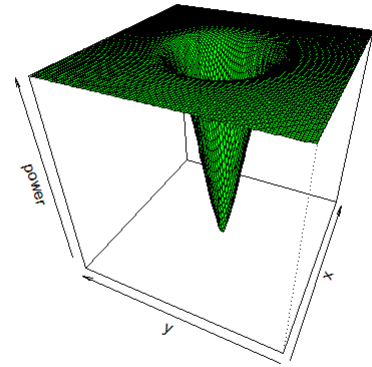
Exponential noise

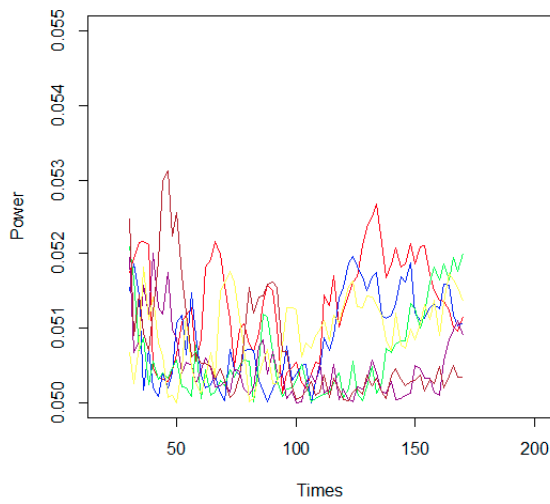
(b) $\rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

Gaussain noise

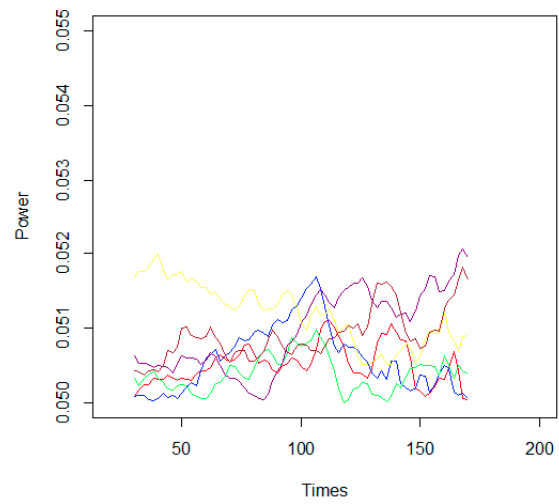
(c) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$

Exponential noise

(d) $\rho_1 = 0.5, \rho_2 = 0.2, \rho_3 = 50, \theta_1 = 0.1, \theta_2 = 0.0025, \theta_3 = 1$ FIGURE 5.7 – Puissance locale du test pour $k = 1$ et $\alpha = 0.1$.

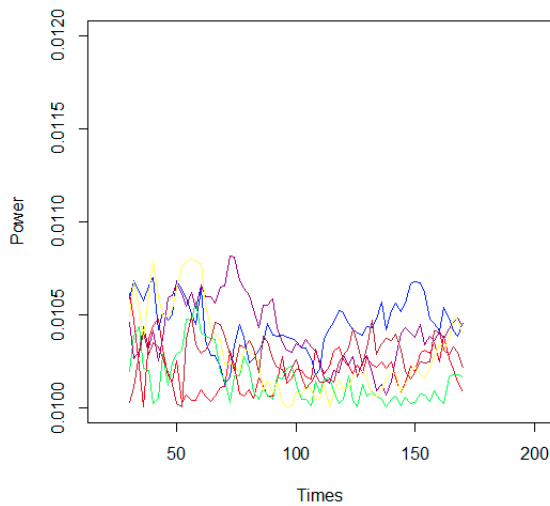


(a) $n = 200, \gamma_0 = 0, \rho_1 = \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

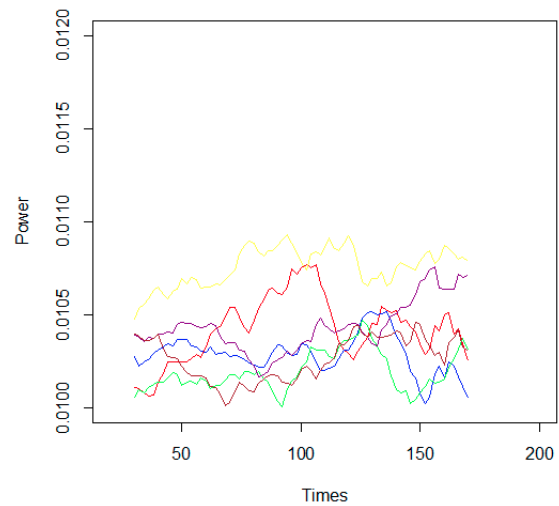


(b) $n = 200, \gamma_0 = 0, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

FIGURE 5.8 – le cas de non-changement pour $\alpha = 0.05$

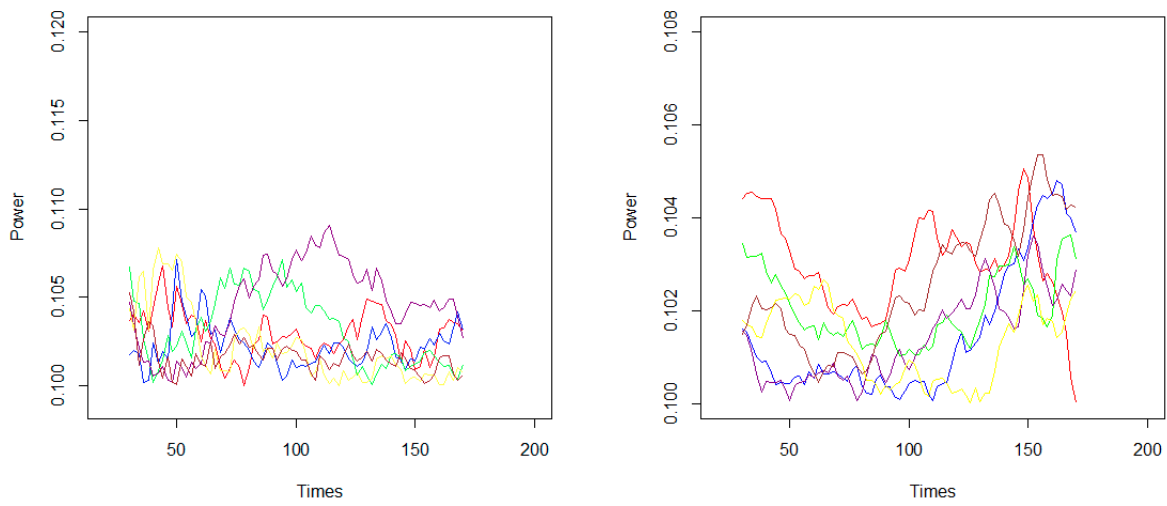


(a) $n = 200, \gamma_0 = 0, \rho_1 = \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$



(b) $n = 200, \gamma_0 = 0, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

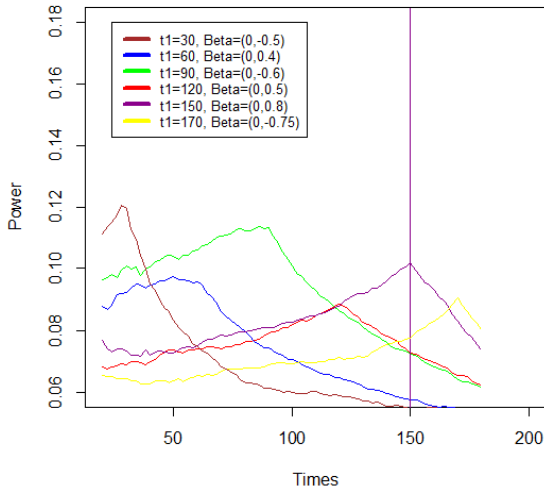
FIGURE 5.9 – Le cas de non-changement pour $\alpha = 0.01$



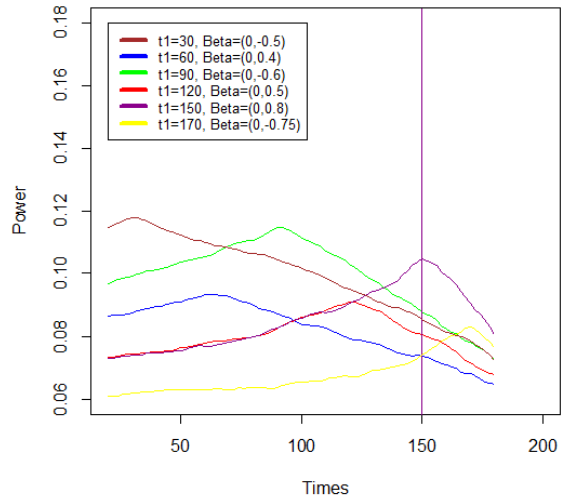
(a) $n = 200, \gamma_0 = 0, \rho_1 = \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

(b) $n = 200, \gamma_0 = 0, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

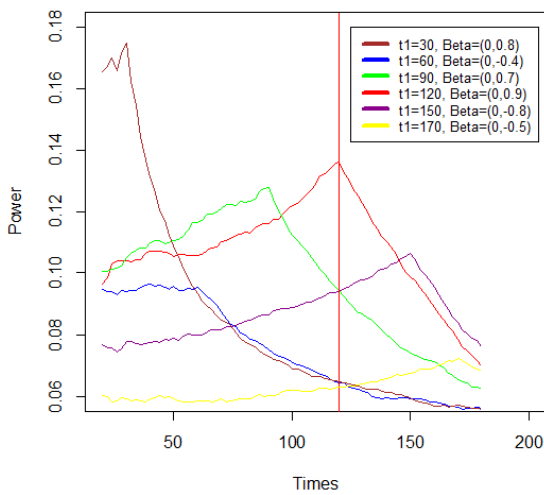
FIGURE 5.10 – le cas de non-changement pour $\alpha = 0.1$



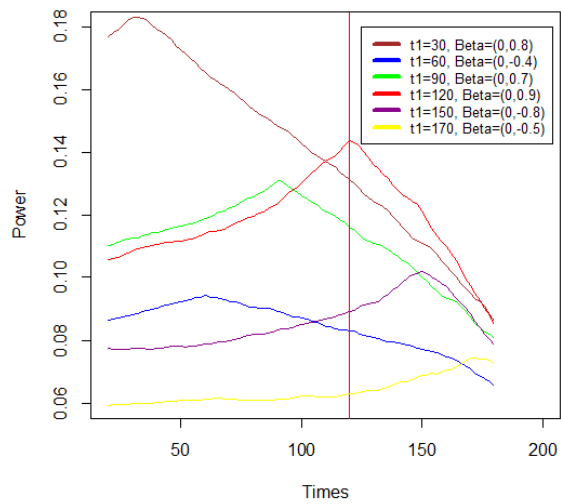
(a) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$



(b) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1,$
 $\theta_2 = 0$

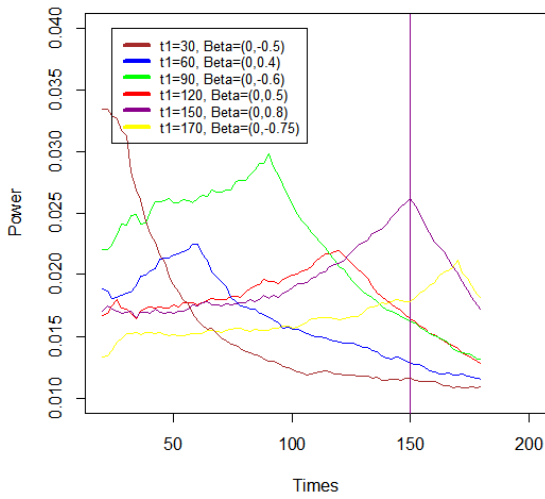


(c) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$

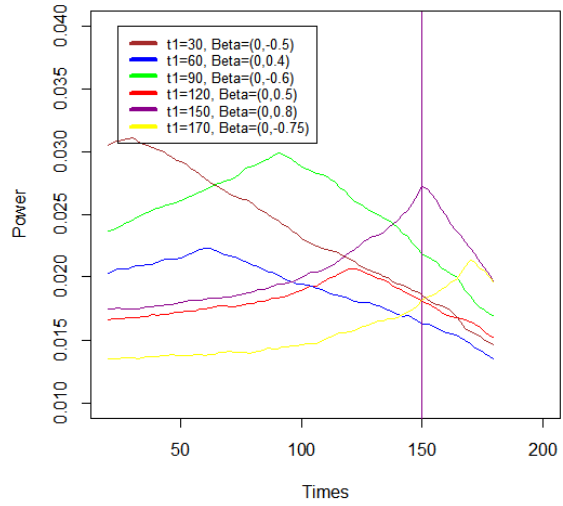


(d) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

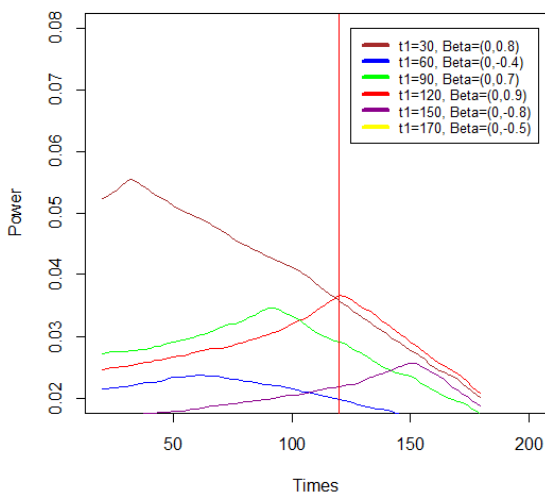
FIGURE 5.11 – Détection des ruptures pour $k = 1$ et $\alpha = 0.05$.



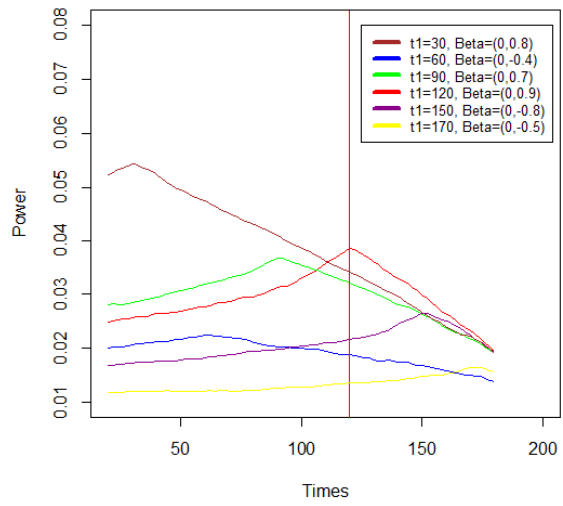
(a) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$



(b) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

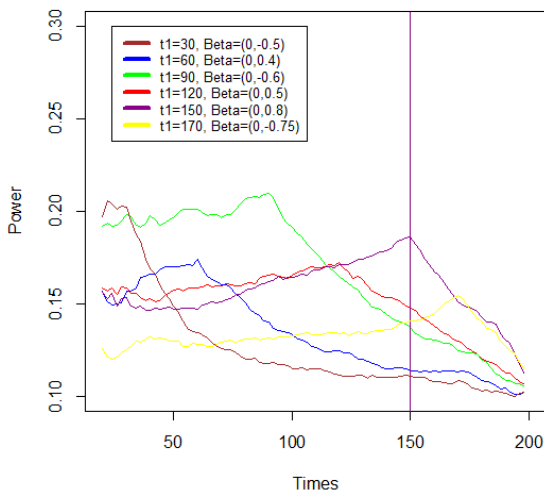


(c) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$

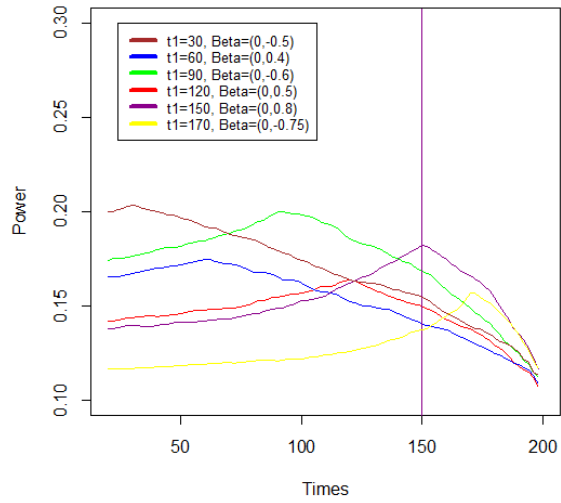


(d) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

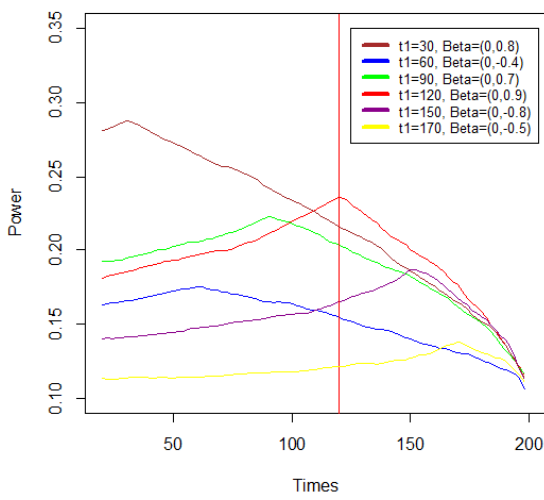
FIGURE 5.12 – Détection des ruptures pour $k = 1$ et $\alpha = 0.01$.



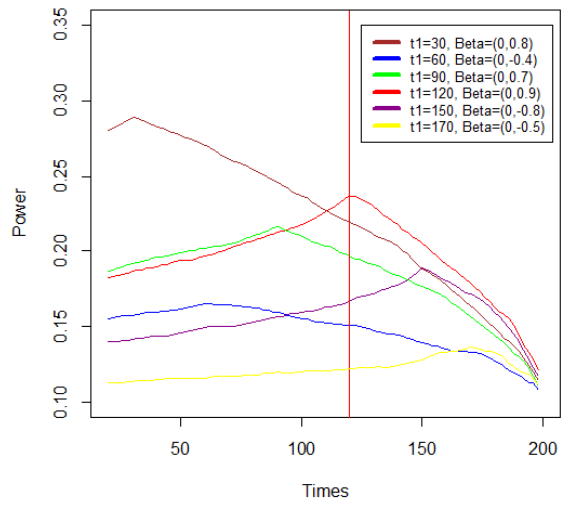
(a) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$



(b) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

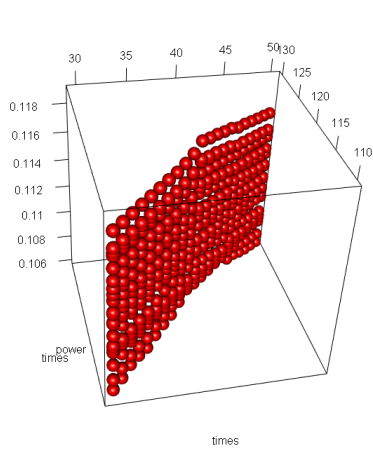


(c) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$

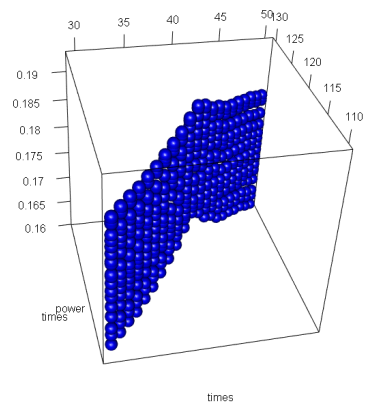


(d) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

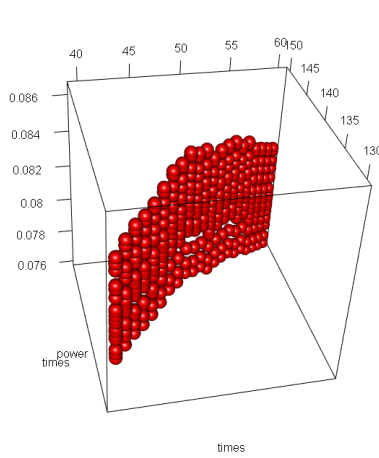
FIGURE 5.13 – Détection des ruptures pour $k = 1$ et $\alpha = 0.1$.



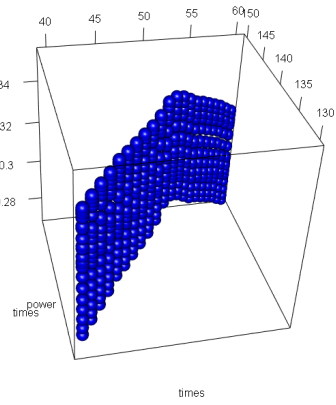
(a) $t_1 = 40, t_2 = 120, \beta = (0, 0.5, -0.5)$



(b) $t_1 = 40, t_2 = 120, \beta = (0, 0.4, -0.4)$

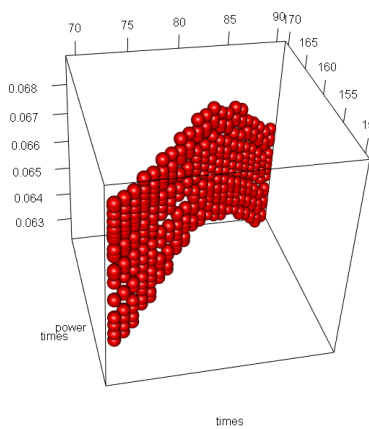


(c) $t_1 = 50, t_2 = 140, \beta = (0, -0.3, 0.3)$

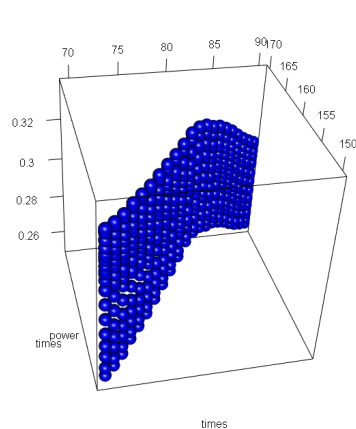


(d)

$t_1 = 50, t_2 = 140, \beta = (0, 0.75, -0.75)$

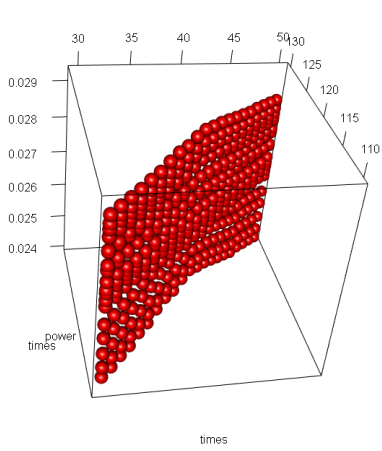


(e) $t_1 = 80, t_2 = 160, \beta = (0, -0.2, 0.2)$

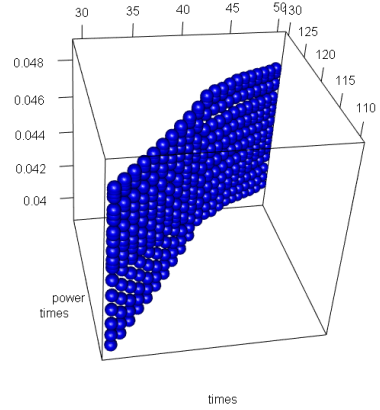


(f) $t_1 = 80, t_2 = 160, \beta = (0, 0.8, -0.8)$

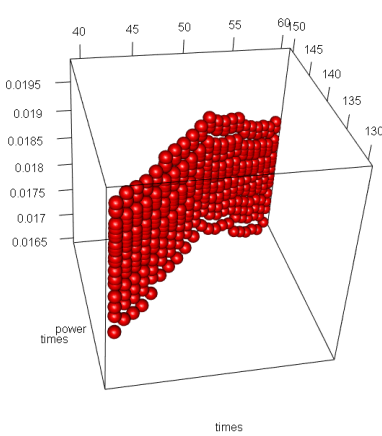
FIGURE 5.14 – Détection des ruptures pour $k = 2$ et $\alpha = 0.05$.



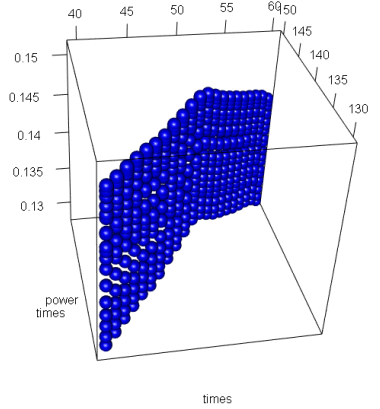
(a) $t_1 = 40, t_2 = 120, \beta = (0, 0.5, -0.5)$



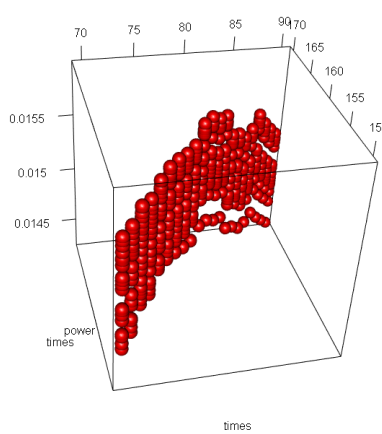
(b) $t_1 = 40, t_2 = 120, \beta = (0, 0.4, -0.4)$



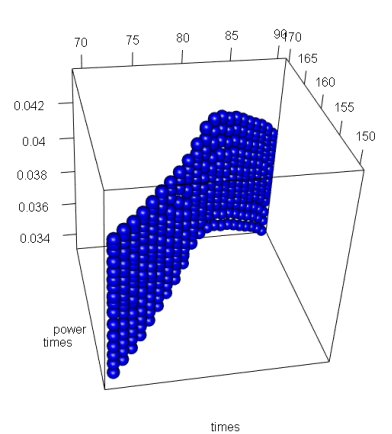
(c) $t_1 = 50, t_2 = 140, \beta = (0, -0.3, 0.3)$



(d) $t_1 = 50, t_2 = 140, \beta = (0, 0.75, -0.75)$

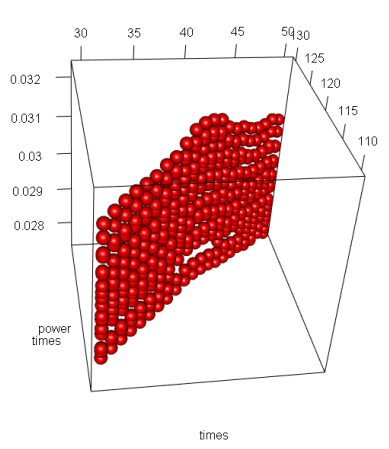


(e) $t_1 = 80, t_2 = 160, \beta = (0, -0.2, 0.2)$

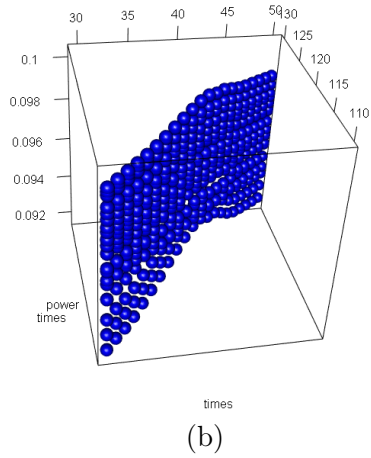


(f) $t_1 = 80, t_2 = 160, \beta = (0, 0.8, -0.8)$

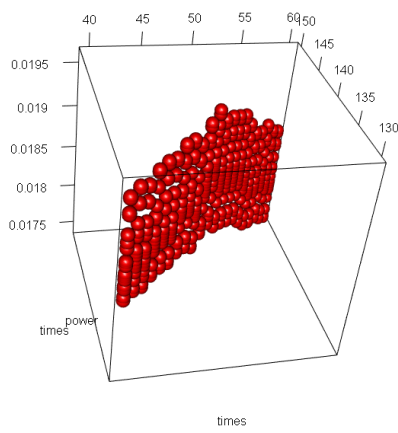
FIGURE 5.15 – Détection des ruptures pour $k = 2$ et $\alpha = 0.01$.



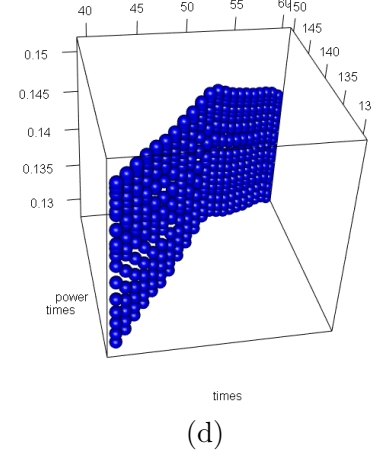
(a) $t_1 = 40, t_2 = 120, \beta = (0, 0.5, -0.5)$



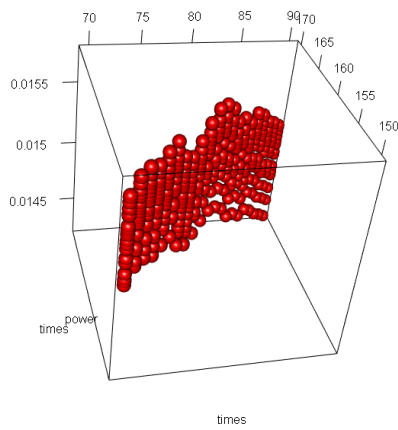
(b) $t_1 = 40, t_2 = 120, \beta = (0, 0.4, -0.4)$



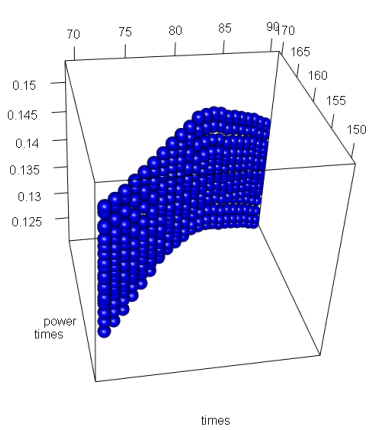
(c) $t_1 = 50, t_2 = 140, \beta = (0, -0.3, 0.3)$



(d) $t_1 = 50, t_2 = 140, \beta = (0, 0.75, -0.75)$

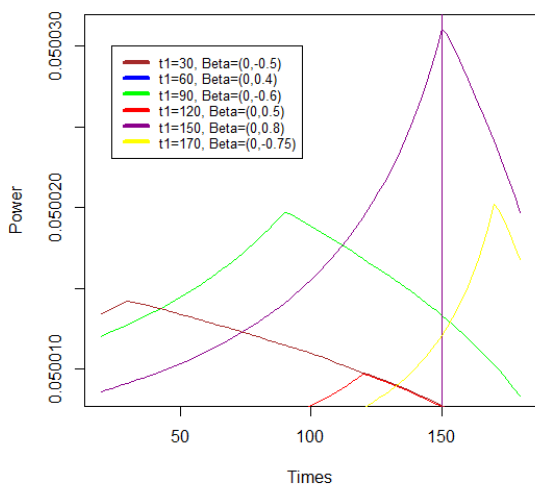


(e) $t_1 = 80, t_2 = 160, \beta = (0, -0.2, 0.2)$

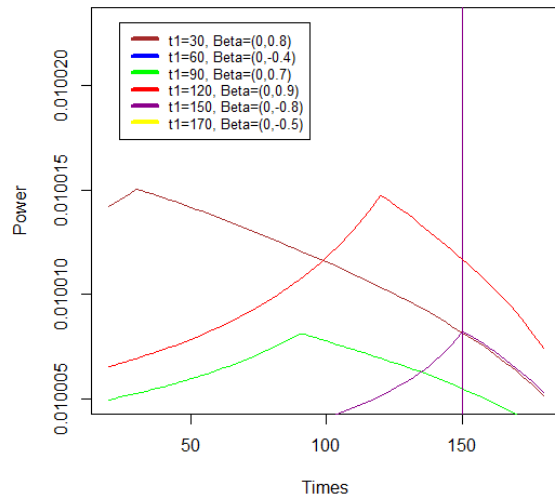


(f) $t_1 = 80, t_2 = 160, \beta = (0, 0.8, -0.8)$

FIGURE 5.16 – Détection des ruptures pour $k = 2$ et $\alpha = 0.1$.

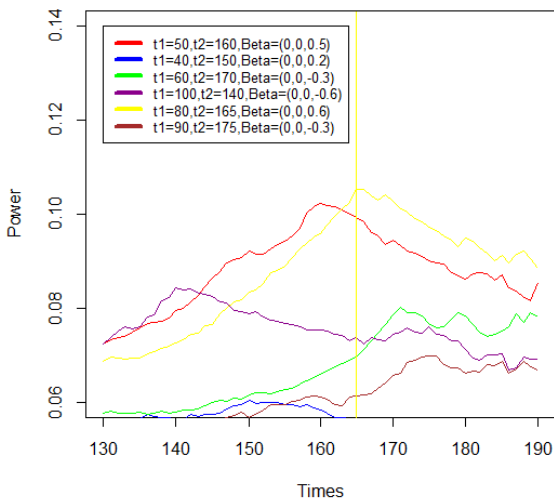


(a) $n = 200$, $T(x) = 0$, $V(x) = 0.5x$ et $\alpha = 0.05$.

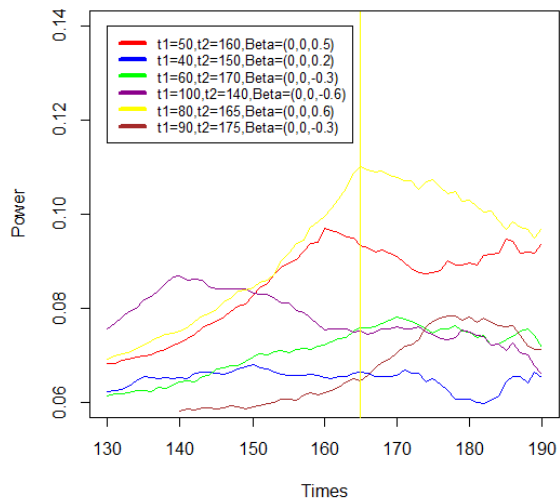


(b) $n = 200$, $T(x) = 0$, $V(x) = 0.5x$ et $\alpha = 0.01$.

FIGURE 5.17 – Détection des ruptures pour $k = 1$.



(a) $\mu_1 = 0.2$, $\mu_2 = -0.8$



(b) $\mu_1 = 0.2$, $\mu_2 = 0.6$

FIGURE 5.18 – Une rupture connue contre la présence d’une rupture faible supplémentaire pour $\alpha = 0.05$.

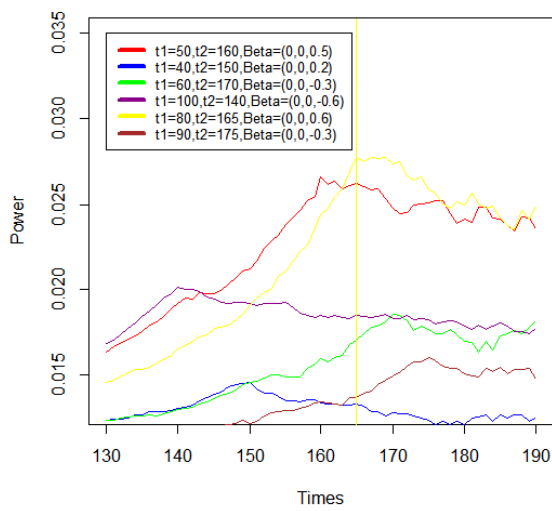
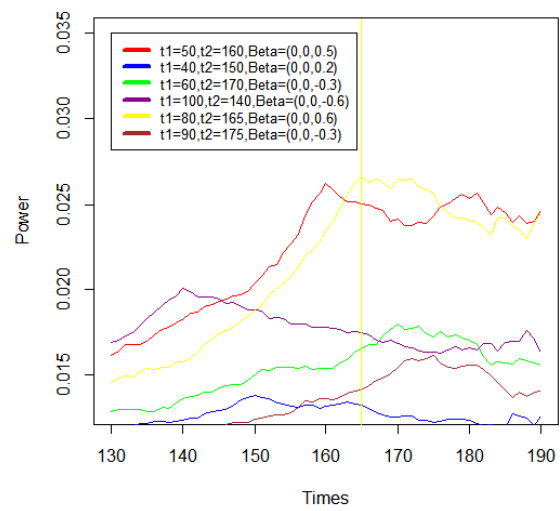
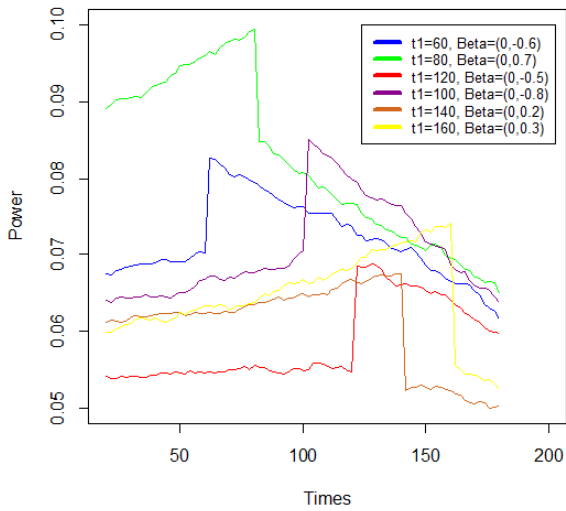
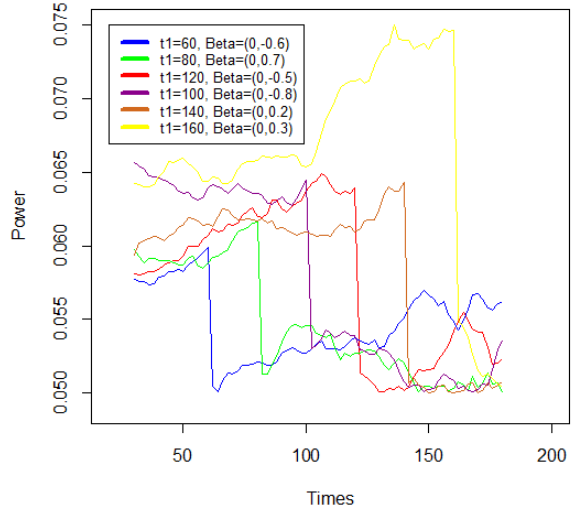
(a) $\mu_1 = 0.2, \mu_2 = -0.8$ (b) $\mu_1 = 0.2, \mu_2 = 0.6$

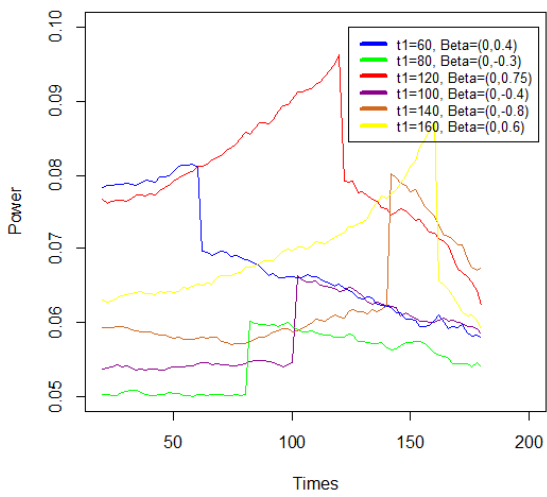
FIGURE 5.19 – Une rupture connue contre la présence d’une rupture faible supplémentaire pour $\alpha = 0.01$.



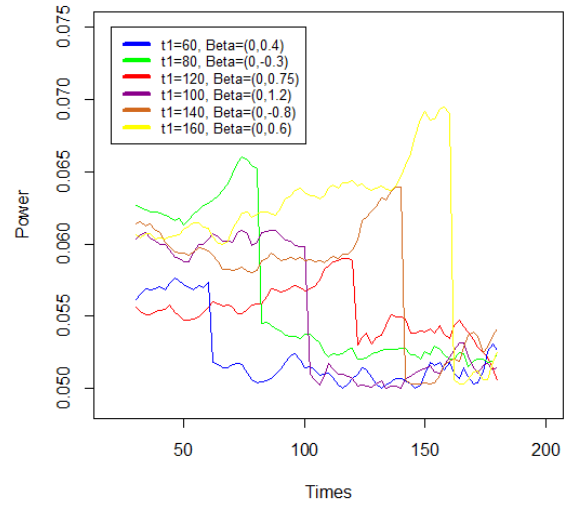
(a) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$



(b) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

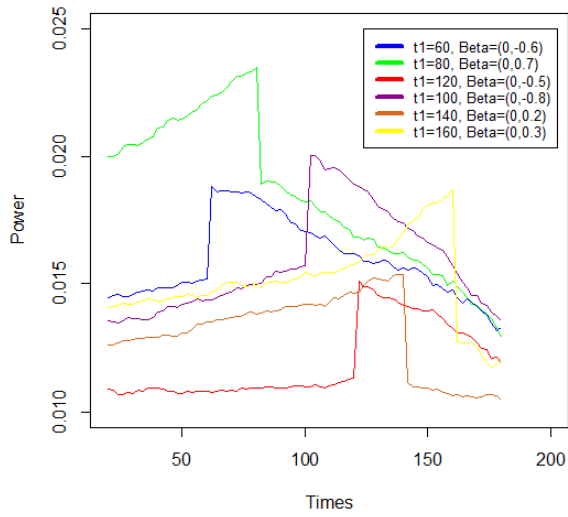


(c) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$

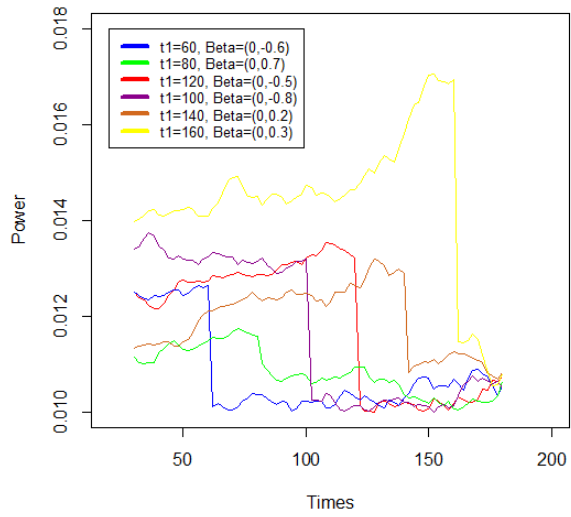


(d) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

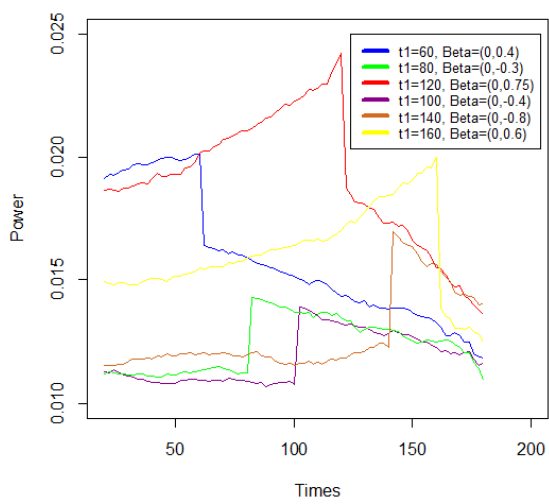
FIGURE 5.20 – Détection des ruptures pour $k = 1$ et $\alpha = 0.05$.



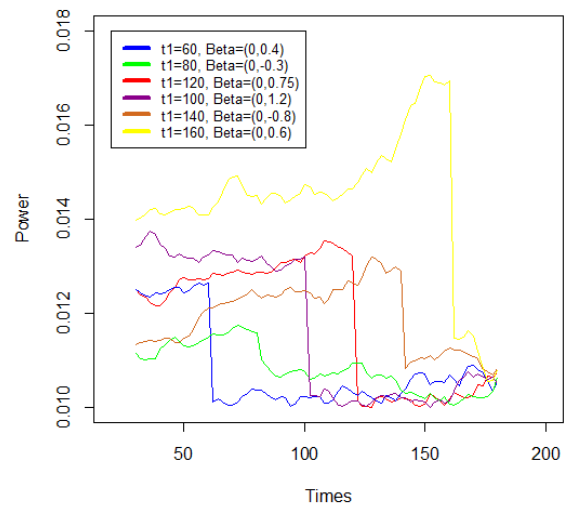
(a) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$



(b) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

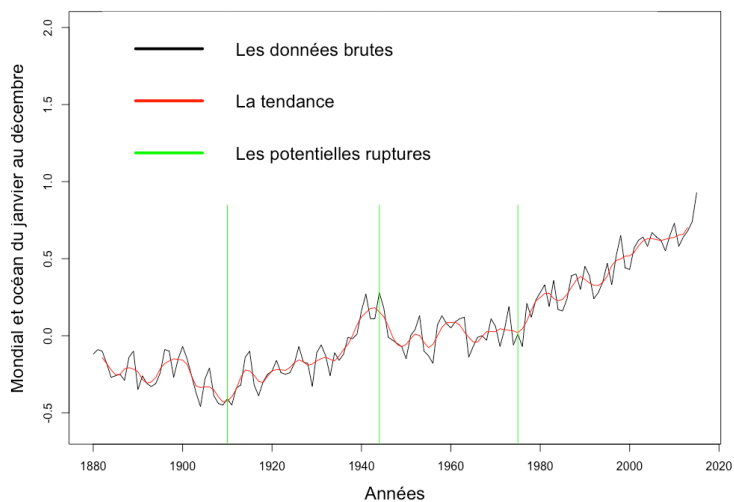


(c) $n = 200, \rho_1 = \rho_2 = 0, \theta_1 = 1$ et $\theta_2 = 0$

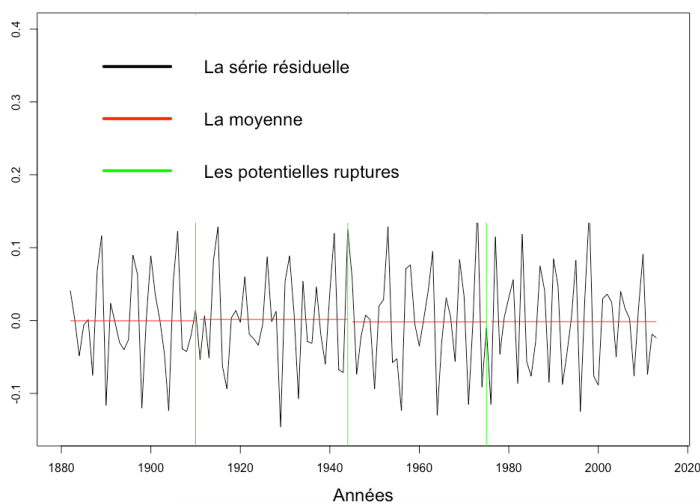


(d) $n = 200, \rho_1 = 0.5, \rho_2 = 0, \theta_1 = 1, \theta_2 = 0$

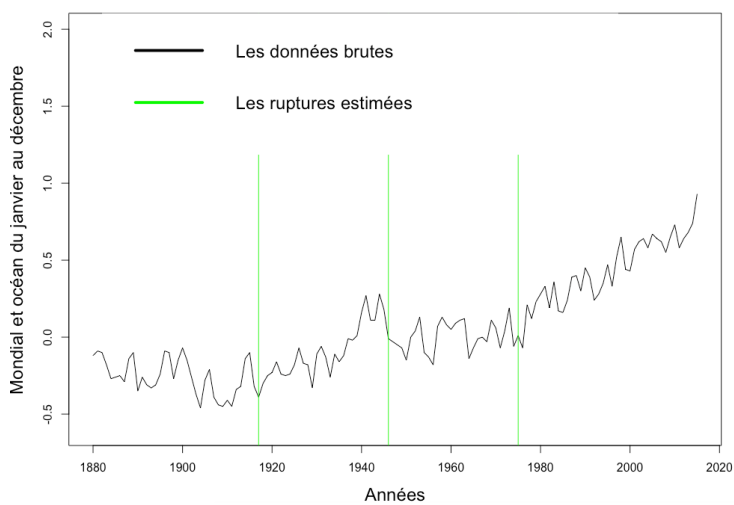
FIGURE 5.21 – Détection des ruptures pour $k = 1$ et $\alpha = 0.01$.



(a) Série brute

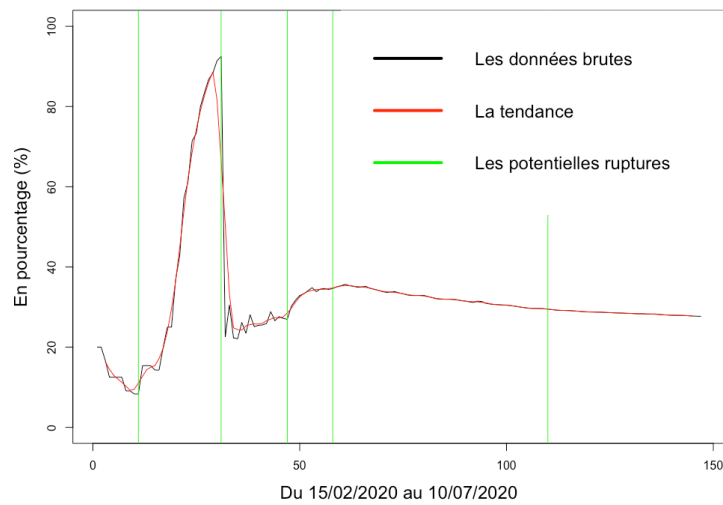


(b) Série résiduelle

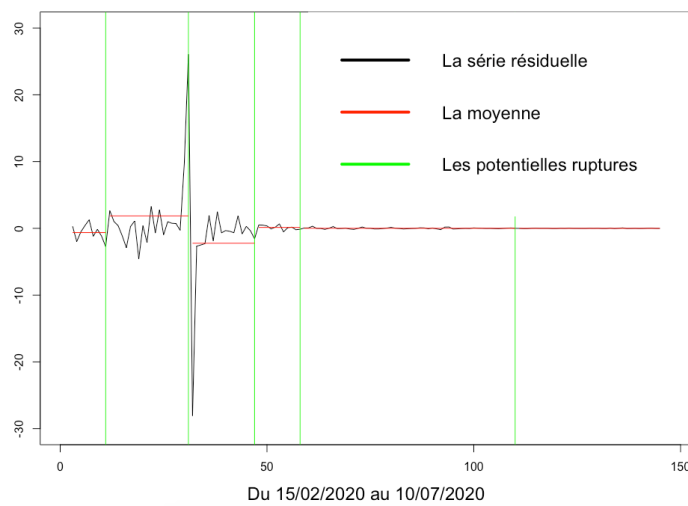


(c) Dates de rupture estimées

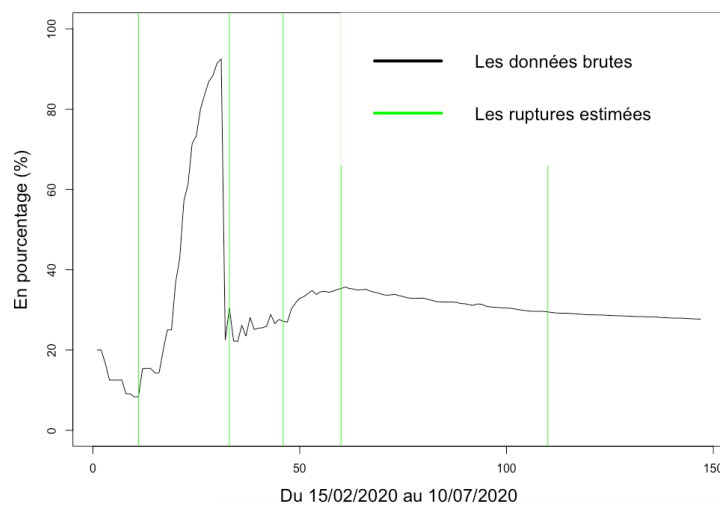
FIGURE 5.22 – Anomalies climatiques annuelles mondiales de 1880 à 2015.



(a) Série brute

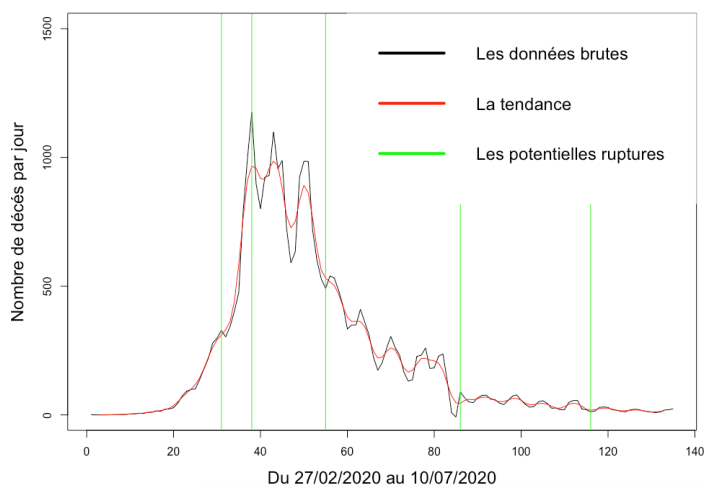


(b) Série résiduelle

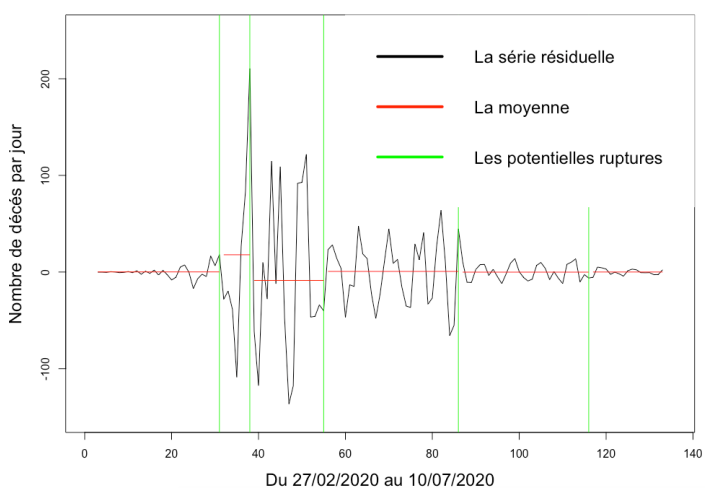


(c) Dates de rupture estimées

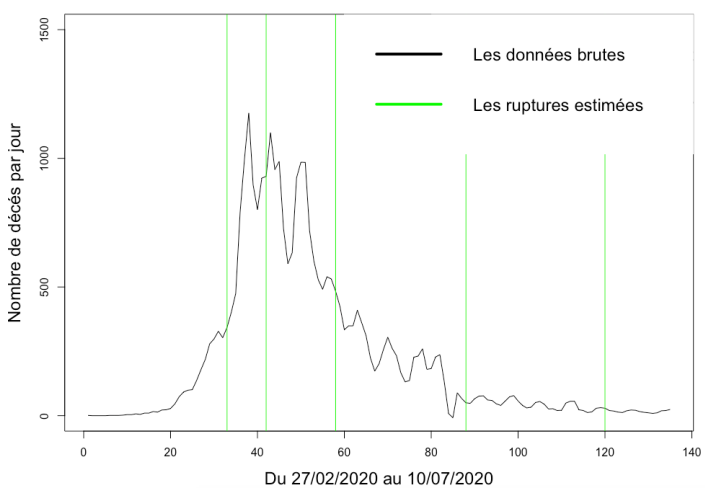
FIGURE 5.23 – Taux de mortalité du COVID-19 en France du 15/02/2020 au 10/07/2020.



(a) Série brute

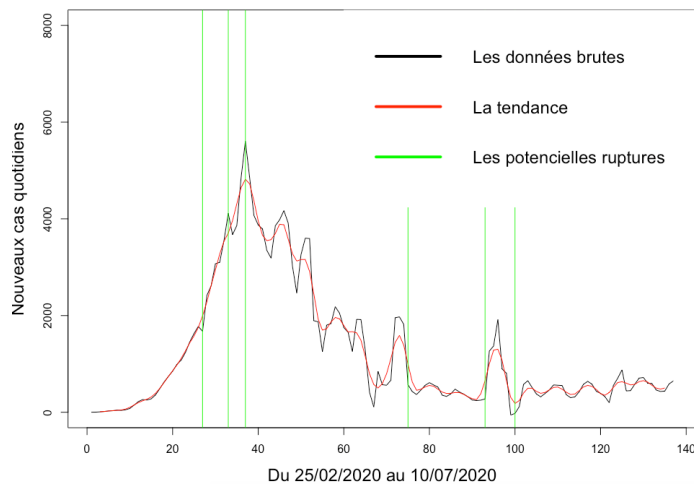


(b) Série résiduelle

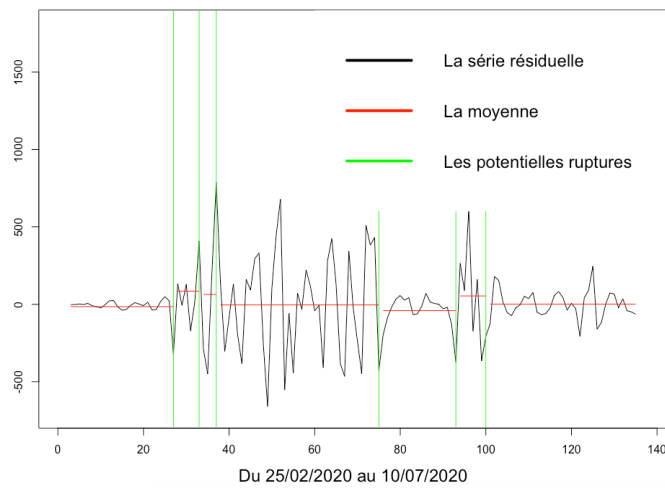


(c) Dates de rupture estimées

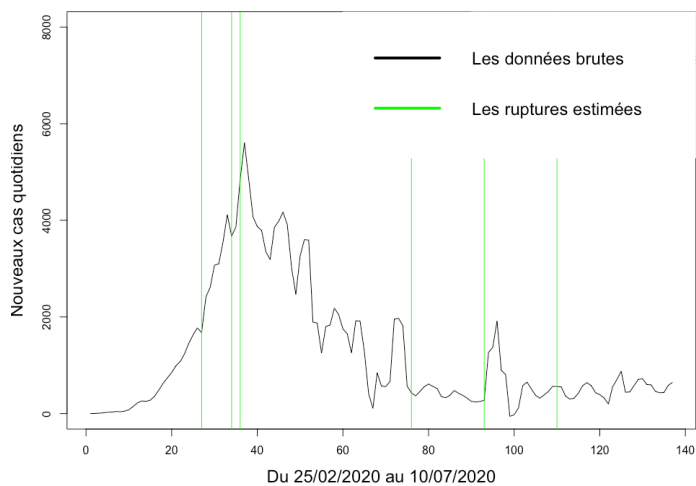
FIGURE 5.24 – Nombre de décès quotidiens du COVID-19 en France du 27/02/2020 au 10/07/2020.



(a) Série brute



(b) Série résiduelle



(c) Dates de rupture estimées

FIGURE 5.25 – Nouveaux cas quotidiens du COVID-19 en France du 25/02/2020 au 10/07/2020.

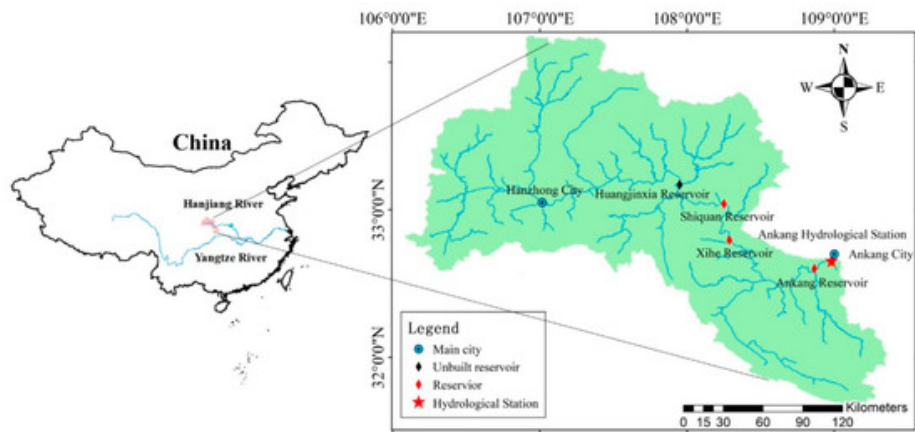
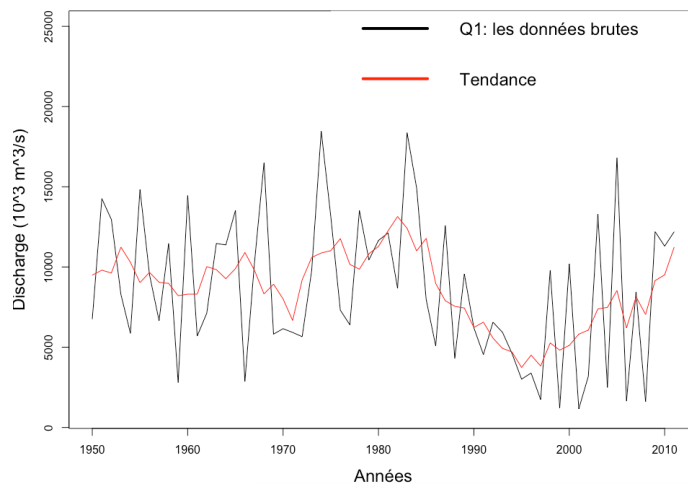
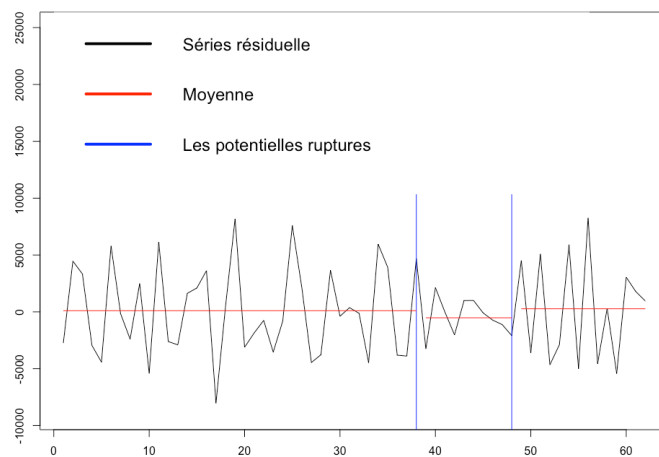


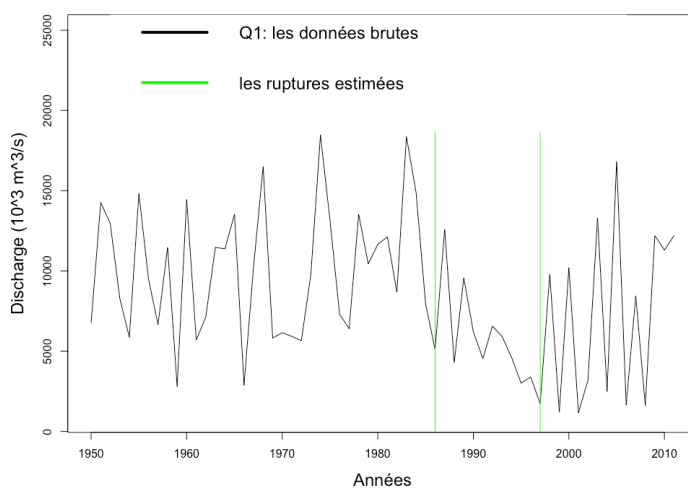
FIGURE 5.26 – La rivière Han jiang de Chine.



(a) Série brute

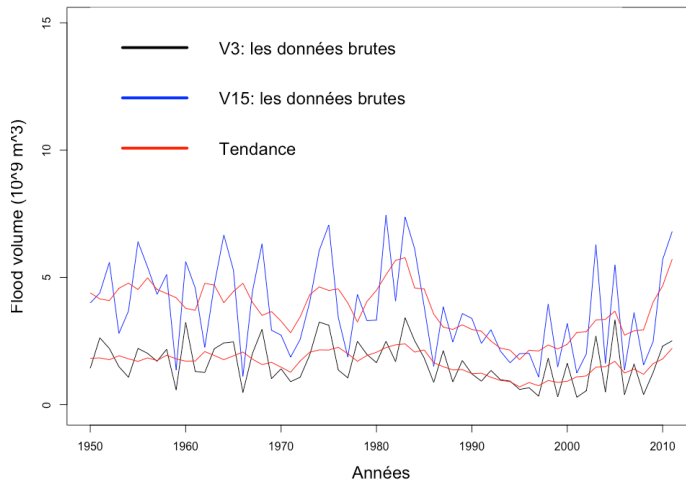


(b) Série résiduelle

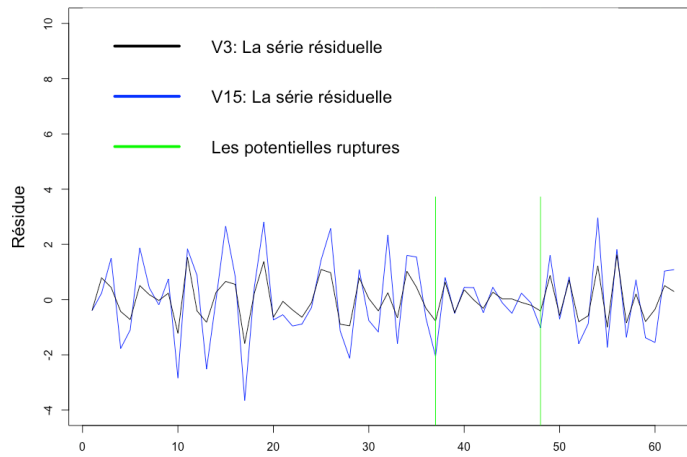


(c) Dates de rupture estimées

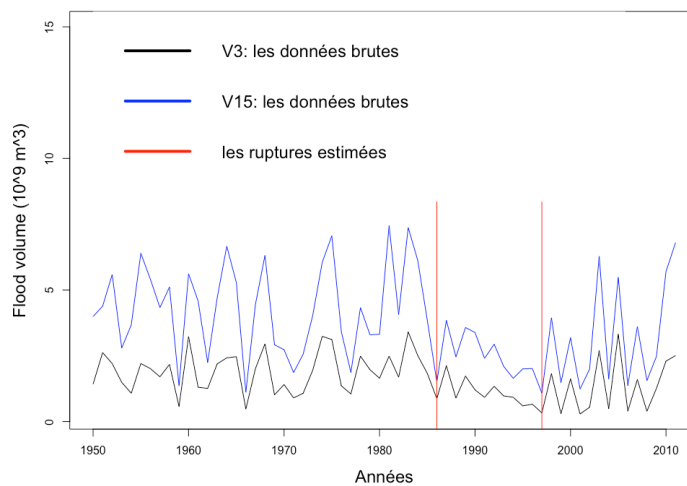
FIGURE 5.27 – Les données des crues univariées composées de la AMMDD (Q_1) de 1950 à 2011.



(a) Série brute

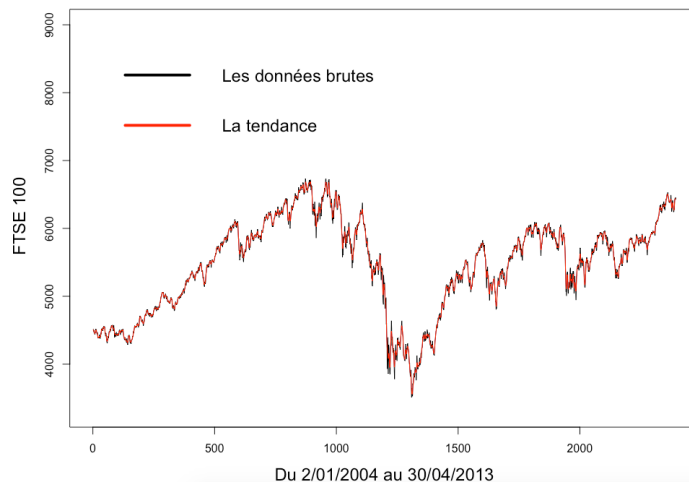


(b) Série résiduelle

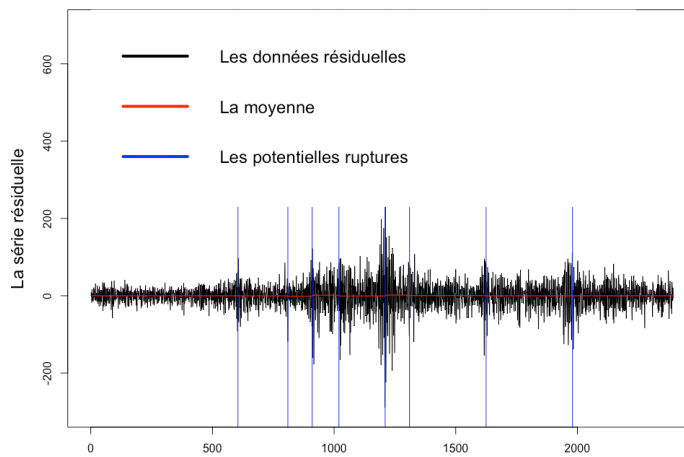


(c) Dates de ruptures estimées

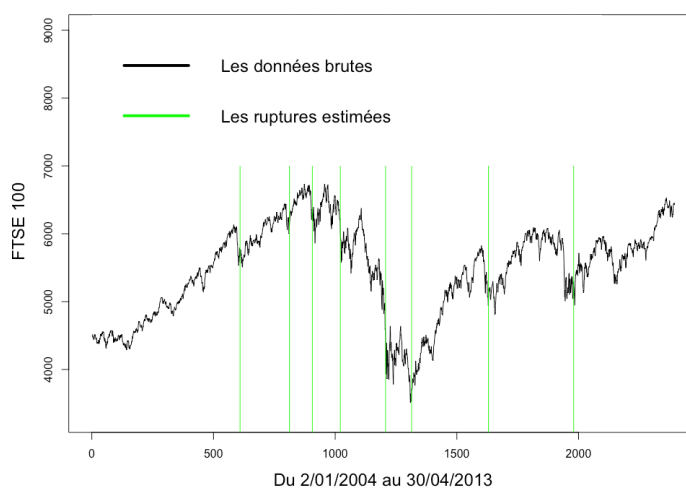
FIGURE 5.28 – Les données du volume annuel maximal des crues sur 3 jours (V_3) et du volume maximal annuel des crues sur 15 jours (V_{15}) de 1950 à 2011.



(a) Série brute

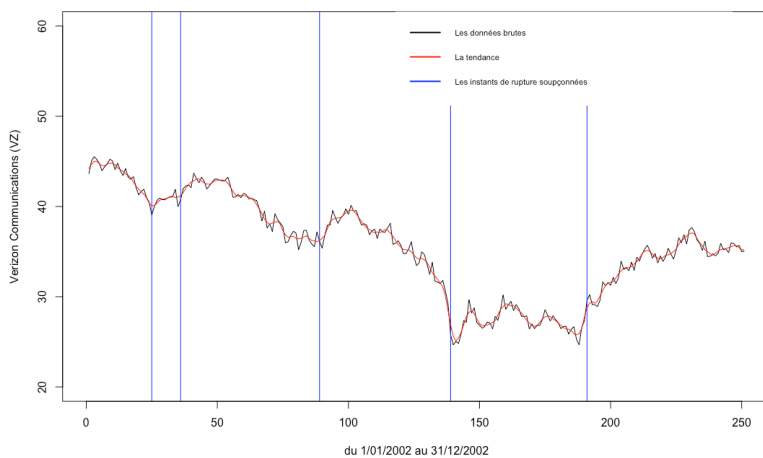


(b) Série résiduelle

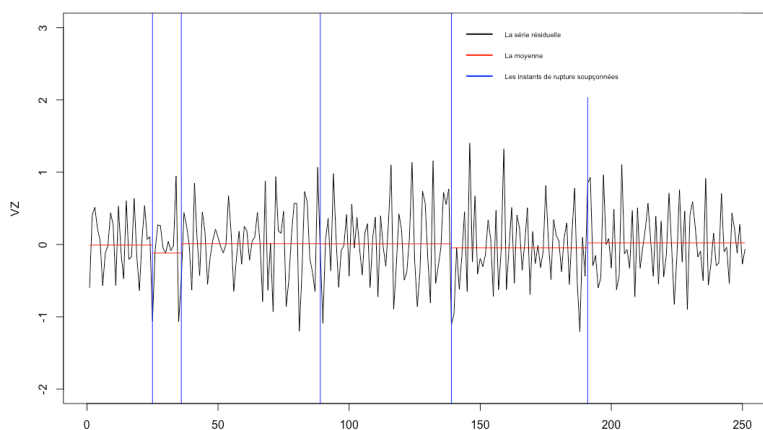


(c) Dates de rupture estimées

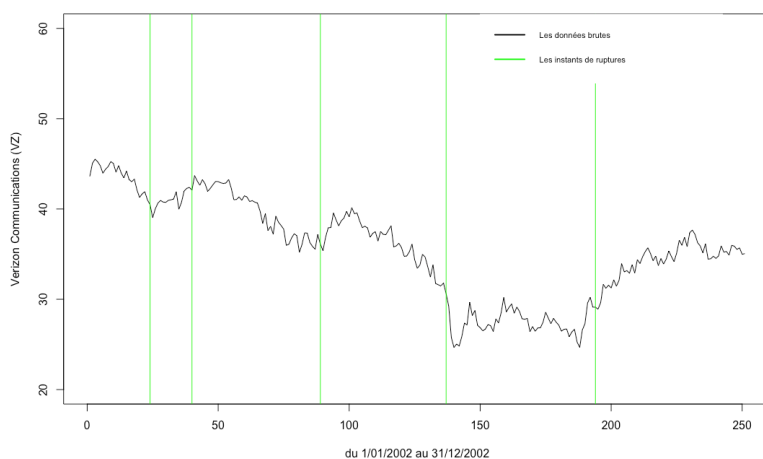
FIGURE 5.29 – Clôture quotidienne de l'indice boursier FTSE 100 du 02/01/2004 au 30/04/2013.



(a) Série brute



(b) Série résiduelle



(c) Dates de rupture estimées

FIGURE 5.30 – Les données journalières des rendements logarithmiques sur l’action Verizon communication (VZ) en 2002.

Chapitre 6

Conclusion et perspectives

6.1 Conclusion générale

Le problème que nous avons étudié dans cette thèse est celui de la détection des points de rupture, sujet qui suscite beaucoup d'intérêt depuis plusieurs années, et qui a fait l'objet de nombreux travaux. Notre étude a été faite à travers des modèles CHARN et l'utilisation d'un test du rapport de vraisemblance.

Nous avons supposé dans un premier temps que les deux fonctions T et V du modèle sont connues et que γ_0 , le paramètre sur lequel est basé le test, est connu. Nous avons établi la propriété LAN et nous avons calculé la puissance locale du test. Dans un second temps, nous avons supposé les fonctions des modèles T et V de formes connues, mais dépendant des paramètres inconnus i.e $T(x) = m(x; \rho)$; $V(x) = \sigma(x; \theta)$; $\rho \in \Theta \subset \mathbb{R}^l$ et $\theta \in \tilde{\Theta} \subset \mathbb{R}^q$. Dans ce dernier cas, les paramètres ont été estimés, et l'expression explicite de la puissance locale du test construit a été donnée lorsque γ_0 est connue et lorsqu'elle est estimée par l'estimateur du maximum de vraisemblance. Ensuite, nous avons étudié la convergence faible du processus de statistique centrale indexé par β pour étudier le cas du test d'un modèle CHARN contre un ensemble de modèles CHARN. Cette étude a été faite dans l'espace de Fréchet $\mathbb{C}(\mathbb{R}^{k+1}, \mathbb{R})$ des fonctions réelles définies dans \mathbb{R}^{k+1} . Dans ce dernier cas, nous avons utilisé un test du type Cramer-Von-Mises.

Les simulations numériques considérées dans le chapitre 5 nous ont montré que les tests construits dans cette thèse pour détecter les points de ruptures dans le cas d'une série d'observations stationnaires par morceaux sont consistants, et les applications aux données réelles ont montré que nos travaux contribuent à la détection des faibles ruptures dans la moyenne des modèles CHARN. Nous avons observé que dans presque tous les exemples considérés, nos stratégies sont meilleures que certaines méthodes auxquelles nous les avons comparées. Cependant, chaque méthode présente des avantages et des inconvénients. Il

n'y a pas une méthode qui pourrait être considérée la meilleure dans tous les cas et en toutes circonstances. Voici quelques inconvénients de nos méthodes : pour que la stratégie S2 donne des bonnes estimations des localisations des points de rupture, il faut que la distance minimale h entre les ruptures soit assez grande pour avoir de bonnes estimations de paramètre β . De plus, cette stratégie exige que la zone qui contient la localisation du premier point de rupture soit connue. Lorsque la série présente plusieurs ruptures, il est préférable d'utiliser la stratégie S3 plutôt que la stratégie S2 car elle marche d'une façon séquentielle et dans ce cas elle sera plus rapide.

Nos méthodes sont basées sur le graphique, il est tout à fait possible d'oublier des localisations de ruptures. Pour l'éviter, on peut considérer le plus de localisations possibles.

6.2 Perspectives

- i. Dans la partie théorique de cette thèse, le travail a été effectué sous la condition où la loi du processus des innovations est supposée connue. Nous n'avons traité le cas où elle est inconnue que par des simulations numériques dans la section 5.2.1.2 du chapitre 5. Il serait intéressant d'étudier ce problème théoriquement.
- ii. Nous nous sommes intéressés à la détection des points de rupture dans la moyenne des modèles CHARN. Il serait intéressant de traiter le cas où le changement a lieu dans la variance ou à la fois dans la moyenne et la variance de ces modèles. Ceci est le sujet d'une thèse en cours.
- iii. Il serait intéressant de créer un package sous R pour implémenter nos stratégies à la détection des faibles ruptures dans la moyenne des modèles CHARN.
- iv. Dans le chapitre 4, nous avons trouvé la loi limite de la statistique de test considérée sous l'hypothèse alternative mais nous n'avons pas réussi à trouver l'expression explicite de sa puissance comme le cas du test traité au chapitre 3. Il serait intéressant de l'étudier par des simulations numériques pour évaluer la performance des résultats trouvés, et les appliquer à des données réelles.
- v. Il serait intéressant aussi de généraliser ce travail dans le cas des modèles CHARN, multidimensionnels, c'est-à-dire on prend $\mathbf{X}_t = (X_t^1, \dots, X_t^p)$, T et V deux fonctions de \mathbb{R}^q à valeurs dans \mathbb{R}^p , $\boldsymbol{\varepsilon}_t = (\varepsilon_t^1, \dots, \varepsilon_t^p)$ et on cherche à détecter les ruptures dans la moyenne et/ou la variance du modèle CHARN multidimensionnel suivant :

$$\mathbf{X}_t = T(\mathbf{X}_{t-1}) + V(\mathbf{X}_{t-1})\boldsymbol{\varepsilon}_t.$$

Nos travaux futurs porteront essentiellement sur les points iii et iv, et nous sommes disposés à collaborer aux recherches en cours sur le point ii.

Bibliographie

- T. AMANO : Asymptotic optimality of estimating function estimator for charn model. *Advances in Decision Sciences*, 2012, 2012.
- A. ANASTASIOU et P. FRYZLEWICZ : Detecting multiple generalized change-points by isolating single ones. *arXiv preprint arXiv :1901.10852*, 2019.
- E. ANDREOU : Restoring monotone power in the cusum test. *Economics Letters*, 98 (1):48–58, 2008.
- D. W. ANDREWS : Tests for parameter instability and structural change with unknown change point. *Econometrica : Journal of the Econometric Society*, p. 821–856, 1993.
- D. W. ANDREWS et J. C. MONAHAN : An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica : Journal of the Econometric Society*, p. 953–966, 1992.
- D. W. ANDREWS et W. PLOBERGER : Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica : Journal of the Econometric Society*, p. 1383–1414, 1994.
- A. L. BADAGIÁN, R. KAISER et D. PEÑA : Time series segmentation procedures to detect, locate and estimate change-points. *In Empirical Economic and Financial Research*, p. 45–59. Springer, 2015.
- R. BARANOWSKI, Y. CHEN et P. FRYZLEWICZ : not : Narrowest-over-threshold change-point detection. *R Package Version*, 1, 2016.
- R. BARANOWSKI, Y. CHEN et P. FRYZLEWICZ : Narrowest-over-threshold change-point detection. *arXiv preprint arXiv :1609.00293*, 2018.
- J.-M. BARDET et W. KENGNE : Monitoring procedure for parameter change in causal time series. *Journal of Multivariate Analysis*, 125:204–221, 2014.

- J.-M. BARDET, W. KENGNE, O. WINTENBERGER *et al.* : Multiple breaks detection in general causal time series using penalized quasi-likelihood. *Electronic Journal of Statistics*, 6:435–477, 2012.
- J.-M. BARDET, W. C. KENGNE et O. WINTENBERGER : Detecting multiple change-points in general causal time series using penalized quasi-likelihood. *arXiv preprint arXiv :1008.0054*, 2010.
- J.-M. BARDET, O. WINTENBERGER *et al.* : Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *The Annals of Statistics*, 37 (5B):2730–2759, 2009.
- I. BASAWA et S. Y. HWANG : Estimation for nonlinear autoregressive models generated by beta-arch processes. *Sankhyā : The Indian Journal of Statistics*, p. 744–762, 2003.
- Y. BENGHABRIT et M. HALLIN : Locally asymptotically optimal tests for autoregressive against bilinear serial dependence. *Statistica sinica*, p. 147–169, 1996.
- Y. BENGHABRIT et M. HALLIN : Locally asymptotically optimal tests for ar (p) against diagonal bilinear dependence. *Journal of statistical planning and inference*, 68(1):47–63, 1998.
- I. BERKES, L. HORVÁTH et P. KOKOSZKA : Testing for parameter constancy in garch (p, q) models. *Statistics & probability letters*, 70(4):263–273, 2004.
- P. BHATTACHARYA et H. ZHOU : Nonparametric stopping rules for detecting small changes in location and scale families. In *From Statistics to Mathematical Finance*, p. 251–271. Springer, 2017.
- G. K. BHATTACHARYYA et R. A. JOHNSON : Nonparametric tests for shift at an unknown time point. *The Annals of Mathematical Statistics*, p. 1731–1743, 1968.
- P. BILLINGSLEY : *Convergence of probability measures*. Wiley-Blackwell, 1968.
- L. BIRGÉ et P. MASSART : Gaussian model selection. *Journal of the European Mathematical Society*, 3(3):203–268, 2001.
- P. J. BROCKWELL, R. A. DAVIS et M. V. CALDER : *Introduction to time series and forecasting*, vol. 2. Springer, 2002.
- R. L. BROWN, J. DURBIN et J. M. EVANS : Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society : Series B (Methodological)*, 37(2):149–163, 1975.

- S. A. CHANDRA et M. TANIGUCHI : Estimating functions for nonlinear time series models. *Annals of the Institute of Statistical Mathematics*, 53(1):125–141, 2001.
- K.-m. CHEN, A. COHEN et H. SACKROWITZ : Consistent multiple testing for change points. *Journal of multivariate analysis*, 102(10):1339–1343, 2011.
- H. CHO et P. FRYZLEWICZ : Multiple change point detection under serial dependence : Wild energy maximisation and gappy schwarz criterion. *arXiv preprint arXiv :1609.00293v2*, 2020.
- G. CIUPERCA : A general criterion to determine the number of change-points. *Statistics & probability letters*, 81(8):1267–1275, 2011.
- A. COHEN, H. B. SACKROWITZ, M. XU *et al.* : A new multiple testing method in the dependent case. *The Annals of Statistics*, 37(3):1518–1544, 2009.
- M. CSÖRGÖ et L. HORVÁTH : *Limit theorems in change-point analysis*, vol. 18. John Wiley & Sons Inc, 1997.
- S. CSÖRGÖ : Kernel-transformed empirical processes. *Journal of multivariate analysis*, 13(4):517–533, 1983.
- R. DAHLHAUS *et al.* : Fitting time series models to nonstationary processes. *The annals of Statistics*, 25(1):1–37, 1997.
- D. A. DARLING, P. ERDÖS *et al.* : A limit theorem for the maximum of normalized sums of independent random variables. *Duke Math. J.*, 23(1):143–155, 1956.
- R. A. DAVIS, D. HUANG et Y.-C. YAO : Testing for a change in the parameter values and order of an autoregressive model. *The Annals of Statistics*, p. 282–304, 1995.
- P. DEHEUVELS et G. V. MARTYNOV : Cramér-von mises-type tests with applications to tests of independence for multivariate extreme-value distributions. *Communications in Statistics-Theory and Methods*, 25(4):871–908, 1996.
- H. DEHLING, A. ROOCH et M. S. TAQQU : Non-parametric change-point tests for long-range dependent data. *Scandinavian Journal of Statistics*, 40(1):153–173, 2013.
- H. DEHLING, A. ROOCH, M. S. TAQQU *et al.* : Power of change-point tests for long-range dependent data. *Electronic Journal of Statistics*, 11(1):2168–2198, 2017.
- A. DENG et P. PERRON : A non-local perspective on the power properties of the cusum and cusum of squares tests for structural change. *Journal of Econometrics*, 142(1):212–240, 2008.

- M. DÖRING : Multiple change-point estimation with u-statistics. *Journal of Statistical Planning and Inference*, 140(7):2003–2017, 2010.
- M. DÖRING : Convergence in distribution of multiple change point estimators. *Journal of statistical planning and inference*, 141(7):2238–2248, 2011.
- J.-J. DROESBEKE et J. FINE : *Inférence non paramétrique : Les statistiques de rangs*. Ed. de l'Université de Bruxelles ; Ed. Ellipses, 1996.
- E. EBERLEIN : On strong invariance principles under dependence assumptions. *The Annals of Probability*, p. 260–270, 1986.
- R. F. ENGLE : Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica : Journal of the econometric society*, p. 987–1007, 1982.
- F. ENIKEEVA, A. MUNK, F. WERNER *et al.* : Bump detection in heterogeneous gaussian regression. *Bernoulli*, 24(2):1266–1306, 2018.
- S. B. FOTOPOULOS, V. K. JANDHYALA et L. TAN : Asymptotic study of the change-point mle in multivariate gaussian families under contiguous alternatives. *Journal of statistical planning and inference*, 139(3):1190–1202, 2009.
- C. FRANCO et J.-M. ZAKOÏAN : Strict stationarity testing and estimation of explosive and stationary generalized autoregressive conditional heteroscedasticity models. *Econometrica*, 80(2):821–861, 2012.
- K. FRICK, A. MUNK et H. SIELING : Multiscale change point inference. *Journal of the Royal Statistical Society : Series B : Statistical Methodology*, p. 495–580, 2014.
- P. FRYZLEWICZ : Detecting possibly frequent change-points : Wild binary segmentation 2 and steepest-drop model selection. *Journal of the Korean Statistical Society*, 49(4):1027–1070, 2020.
- P. FRYZLEWICZ *et al.* : Wild binary segmentation for multiple change-point detection. *Annals of Statistics*, 42(6):2243–2281, 2014.
- P. FRYZLEWICZ et S. S. RAO : Multiple-change-point detection for auto-regressive conditional heteroscedastic processes. *Journal of the Royal Statistical Society : Series B : Statistical Methodology*, p. 903–924, 2014.
- L. GARDNER : On detecting changes in the mean of normal variates. *The Annals of Mathematical Statistics*, 40(1):116–126, 1969.

- E. GOMBAY et L. HORVÁTH : An application of the maximum likelihood test to the change-point problem. *Stochastic Processes and their Applications*, 50(1):161–171, 1994.
- E. GOMBAY et D. SERBAN : Monitoring parameter change in ar (p) time series models. *Journal of Multivariate Analysis*, 100(4):715–725, 2009.
- C. GRANGER et A. ANDERSEN : An introduction to bilinear time series models. vanden-hoek & ruprecht, göttingen. *Recent Advances in ARCH Modelling*, 35:36, 1978.
- D. GUÉGAN *et al.* : Séries chronologiques non linéaires à temps discret. Rap. tech., 1994.
- J. GURLAND : Quadratic forms in normally distributed random variables. *Sankhyā : The Indian Journal of Statistics (1933-1960)*, 17(1):37–50, 1956.
- P. HALL et C. C. HEYDE : *Martingale limit theory and its application*. Academic press, 1980.
- W. HALL et D. J. MATHIASON : On large-sample estimation and testing in parametric models. *International Statistical Review/Revue Internationale de Statistique*, p. 77–97, 1990.
- W. HÄRDLE et A. TSYBAKOV : Local polynomial estimators of the volatility function in nonparametric autoregression. *Journal of econometrics*, 81(1):223–242, 1997.
- W. HÄRDLE, A. TSYBAKOV et L. YANG : Nonparametric vector autoregression. *Journal of Statistical Planning and Inference*, 68(2):221–245, 1998.
- D. M. HAWKINS : Fitting multiple change-point models to data. *Computational Statistics & Data Analysis*, 37(3):323–341, 2001.
- N. HENZE et T. WAGNER : A new approach to the bhep tests for multivariate normality. *Journal of Multivariate Analysis*, 62(1):1–23, 1997.
- D. V. HINKLEY : Inference about the change-point in a sequence of random variables. 1970.
- L. HORVÁTH : The maximum likelihood method for testing changes in the parameters of normal observations. *The Annals of statistics*, p. 671–680, 1993.
- L. HORVÁTH : Change-point detection in long-memory processes. *Journal of Multivariate Analysis*, 78(2):218–234, 2001.
- L. HORVÁTH et M. HUŠKOVÁ : Testing for changes using permutations of u-statistics. *Journal of Statistical Planning and Inference*, 128(2):351–371, 2005.

- L. HORVÁTH, M. HUŠKOVÁ, P. KOKOSZKA et J. STEINEBACH : Monitoring changes in linear models. *Journal of Statistical Planning and Inference*, 126(1):225–251, 2004.
- L. HORVÁTH, M. HUSKOVÁ, G. RICE et J. WANG : Asymptotic properties of the cusum estimator for the time of change in linear panel data models. *Econometric Theory*, 33(2):366, 2017.
- L. HORVÁTH et P. KOKOSZKA : The effect of long-range dependence on change-point estimators. *Journal of Statistical Planning and Inference*, 64(1):57–81, 1997.
- L. HORVÁTH, P. KOKOSZKA et A. ZHANG : Monitoring constancy of variance in conditionally heteroskedastic time series. *Econometric Theory*, p. 373–402, 2006.
- L. HORVÁTH, C. MILLER et G. RICE : A new class of change point test statistics of rényi type. *Journal of Business & Economic Statistics*, 38(3):570–579, 2020.
- L. HORVÁTH et J. STEINEBACH : Testing for changes in the mean or variance of a stochastic process under weak invariance. *Journal of statistical planning and inference*, 91(2):365–376, 2000.
- J. HUH : Detection of a change point based on local-likelihood. *Journal of multivariate analysis*, 101(7):1681–1700, 2010.
- M. HUŠKOVÁ, C. KIRCH, Z. PRÁŠKOVÁ et J. STEINEBACH : On the detection of changes in autoregressive time series, ii. resampling procedures. *Journal of Statistical Planning and Inference*, 138(6):1697–1721, 2008.
- M. HUŠKOVÁ et A. KOUBKOVÁ : Monitoring jump changes in linear models. *J. Statist. Res*, 39(2):51–70, 2005.
- M. HUŠKOVÁ, Z. PRÁŠKOVÁ et J. STEINEBACH : On the detection of changes in autoregressive time series i. asymptotics. *Journal of Statistical Planning and Inference*, 137(4):1243–1259, 2007.
- S. Y. HWANG et I. BASAWA : Nonlinear time series contiguous to ar (1) processes and a related efficient test for linearity. *Statistics & probability letters*, 52(4):381–390, 2001.
- J.-P. IMHOF : Computing the distribution of quadratic forms in normal variables. *Biometrika*, 48(3/4):419–426, 1961.
- C. INCLAN et G. C. TIAO : Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89(427):913–923, 1994.

- D. JARUŠKOVÁ : Asymptotic behaviour of a test statistic for detection of change in mean of vectors. *Journal of statistical planning and inference*, 140(3):616–625, 2010.
- C. JIANG, L. XIONG, C.-Y. XU et S. GUO : Bivariate frequency analysis of nonstationary low-flow series based on the time-varying copula. *Hydrological Processes*, 29(6):1521–1534, 2015.
- P. JIANG et E. KUROZUMI : Power properties of the modified cusum tests. *Communications in Statistics-Theory and Methods*, 48(12):2962–2981, 2019.
- H. KANAI, H. OGATA et M. TANIGUCHI : Estimating function approach for charn models. *Metron*, 68(1):1–21, 2010.
- W. C. KENGNE : Testing for parameter constancy in general causal time-series models. *Journal of Time Series Analysis*, 33(3):503–518, 2012.
- T. G. KHAKHUBIA : A limit theorem for a maximum-likelihood estimate of the disorder time. *Theory of Probability & Its Applications*, 31(1):141–144, 1987.
- S. KIM, S. CHO et S. LEE : On the cusum test for parameter changes in garch (1, 1) models. *Communications in Statistics-Theory and Methods*, 29(2):445–462, 2000.
- S. KOTZ, N. L. JOHNSON et D. BOYD : Series representations of distributions of quadratic forms in normal variables. i. central case. *The Annals of Mathematical Statistics*, 38(3):823–837, 1967a.
- S. KOTZ, N. L. JOHNSON et D. BOYD : Series representations of distributions of quadratic forms in normal variables ii. non-central case. *The Annals of Mathematical Statistics*, 38(3):838–848, 1967b.
- O. KOUAMO, E. MOULINES et F. ROUEFF : Testing for homogeneity of variance in the wavelet domain. *In Dependence in probability and statistics*, p. 175–205. Springer, 2010.
- W. KRÄMER, W. PLOBERGER et R. ALT : Testing for structural change in dynamic models. *Econometrica : Journal of the Econometric Society*, p. 1355–1369, 1988.
- W. KRÄMER et P. SCHOTMAN : Range vs. maximum in the ols-based version of the cusum test. *Economics letters*, 40(4):379–381, 1992.
- M. LAVIELLE : Optimal segmentation of random processes. *IEEE Transactions on Signal Processing*, 46(5):1365–1373, 1998.

- M. LAVIELLE et E. LEBARBIER : An application of mcmc methods for the multiple change-points problem. *Signal processing*, 81(1):39–53, 2001.
- L. LE CAM : *Asymptotic methods in statistical decision theory*. Springer-Verlag, 1986.
- É. LEBARBIER : Detecting multiple change-points in the mean of gaussian process by model selection. *Signal processing*, 85(4):717–736, 2005.
- S. LEE et J. SONG : Test for parameter change in arma models with garch innovations. *Statistics & Probability Letters*, 78(13):1990–1998, 2008.
- S. LEE, Y. TOKUTSU et K. MAEKAWA : The residual cusum test for the constancy of parameters in garch (1, 1) models. *Submitted for publication*, 2003.
- E. L. LEHMANN et J. P. ROMANO : *Testing statistical hypotheses*. Springer Science & Business Media, 2006.
- H. LI, A. MUNK, H. SIELING *et al.* : Fdr-control in multiscale change-point segmentation. *Electronic Journal of Statistics*, 10(1):918–959, 2016.
- T. LOUNIS : *Inférences dans les modèles ARCH : tests localement asymptotiquement optimaux*. Thèse de doctorat, Université de Lorraine, 2015.
- I. B. MACNEILL : Tests for change of parameter at unknown times and distributions of some related functionals on brownian motion. *The Annals of Statistics*, p. 950–962, 1974.
- S. G. MEINTANIS, J. NGATCHOU-WANDJI et E. TAUFER : Goodness-of-fit tests for multivariate stable distributions based on the empirical characteristic function. *Journal of Multivariate Analysis*, 140:171–192, 2015.
- M. J. MENNE et C. N. WILLIAMS JR : Detection of undocumented changepoints using multiple test statistics and composite reference series. *Journal of Climate*, 18(20):4271–4286, 2005.
- A. MONFORT : *Cours de statistique mathématique*. Economica, 1996.
- J. NGATCHOU-WANDJI : Checking nonlinear heteroscedastic time series models. *Journal of statistical planning and inference*, 133(1):33–68, 2005.
- J. NGATCHOU-WANDJI : Testing for symmetry in multivariate distributions. *Statistical methodology*, 6(3):230–250, 2009.

- J. NGATCHOU-WANDJI et M. HAREL : A cramér-von mises test for symmetry of the error distribution in asymptotically stationary stochastic models. *Statistical Inference for Stochastic Processes*, 16(3):207–236, 2013.
- J. NGATCHOU-WANDJI et N. LAÏB : Local power of a cramér-von mises type test for parametric autoregressive models of order one. *Computers & Mathematics with Applications*, 56(4):918–929, 2008.
- J. NGATCHOU-WANDJI et M. LTAIFA : Détection de rupture dans des données climatiques et dans des données de santé. *Biostatistiques et sciences de la santé*, 2(1), 2021a. ISSN 2632-8291. URL <http://www.openscience.fr/Detection-de-rupture-dans-des-donnees-climatiques-et-dans-des-donnees-de-sante>.
- J. NGATCHOU-WANDJI et M. LTAIFA : On detecting weak changes in the mean of charn models. *ArXiv : 2101.08597*, 2021b.
- J. NGATCHOU-WANDJI, M. L. PURI, M. HAREL et E. ELHARFAOUI : Testing nonstationary and absolutely regular nonlinear time series models. *Statistical Inference for Stochastic Processes*, 22(3):557–593, 2019.
- K. NOSEK : Schwarz information criterion based tests for a change-point in regression models. *Statistical Papers*, 51(4):915–929, 2010.
- A. B. OLSHEN, E. VENKATRAMAN, R. LUCITO et M. WIGLER : Circular binary segmentation for the analysis of array-based dna copy number data. *Biostatistics*, 5(4):557–572, 2004.
- E. PAGE : A test for a change in a parameter occurring at an unknown point. *Biometrika*, 42(3/4):523–527, 1955.
- E. PARZEN : On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076, 1962.
- W. PLOBERGER et W. KRÄMER : The local power of the cusum and cusum of squares tests. *Econometric Theory*, p. 335–347, 1990.
- W. PLOBERGER, W. KRAMER *et al.* : On studentizing a test for structural change. *Economics Letters*, 20(4):341–344, 1986.
- Z. PRÁŠKOVÁ et O. CHOCHOLA : M-procedures for detection of a change under weak dependence. *Journal of Statistical Planning and Inference*, 149:60–76, 2014.

- A. RÉNYI : On the theory of order statistics. *Acta Mathematica Academiae Scientiarum Hungarica*, 4(3-4):191–231, 1953.
- F. RIESZ et B. SZÖKEFALVI-NAGY : Leçons d'analyse fonctionnelle. *laf*, 1972.
- H. RUBEN : Probability content of regions under spherical normal distributions, iv : The distribution of homogeneous and non-homogeneous quadratic functions of normal variables. *The Annals of Mathematical Statistics*, 33(2):542–570, 1962.
- E. RUGGIERI et M. ANTONELLIS : An exact approach to bayesian sequential change point detection. *Computational Statistics & Data Analysis*, 97:71–86, 2016.
- M. SANKARAN : On the non-central chi-square distribution. *Biometrika*, 46(1/2):235–237, 1959.
- M. SANKARAN : Approximations to the non-central chi-square distribution. *Biometrika*, 50(1/2):199–204, 1963.
- G. SCHWARZ *et al.* : Estimating the dimension of a model. *Annals of statistics*, 6(2):461–464, 1978.
- A. SEN et M. S. SRIVASTAVA : On tests for detecting change in mean. *The Annals of statistics*, p. 98–108, 1975.
- P. K. SEN : Sequential nonparametrics : Invariance principles and statistical inference. *John Wiley & Sons, New York*, 1981.
- J. SIMON : *Espaces de Banach, Fréchet, Hilbert et Neumann*, vol. 1. ISTE Group, 2017.
- D. SUL, P. C. PHILLIPS et C.-Y. CHOI : Prewhitening bias in hac estimation. *Oxford Bulletin of Economics and Statistics*, 67(4):517–546, 2005.
- A. R. SWENSEN : The asymptotic distribution of the likelihood ratio for autoregressive time series with a regression trend. *Journal of Multivariate Analysis*, 16(1):54–70, 1985.
- H. TONG : *Non-linear time series : a dynamical system approach*. Oxford University Press, 1990.
- A. W. Van der VAART : *Asymptotic statistics*, vol. 3. Cambridge university press, 2000.
- T. J. VOGELSANG : Wald-type tests for detecting breaks in the trend function of a dynamic time series. *Econometric Theory*, p. 818–849, 1997.

- T. J. VOGELSANG : Sources of nonmonotonic power when testing for a shift in mean of a dynamic time series. *Journal of Econometrics*, 88(2):283–299, 1999.
- L. Y. VOSTRIKOVA : Detecting “disorder” in multidimensional random processes. *In Doklady Akademii Nauk*, vol. 259, p. 270–274. Russian Academy of Sciences, 1981.
- Q. WANG, P. C. PHILLIPS *et al.* : A specification test for nonlinear nonstationary models. *The Annals of Statistics*, 40(2):727–758, 2012.
- P. WHITTLE : Estimation and information in stationary time series. *Arkiv för matematik*, 2(5):423–434, 1953.
- H. WONG, W. IP *et al.* Y. LI : Detection of jumps by wavelets in a heteroscedastic autoregressive model. *Statistics & probability letters*, 52(4):365–372, 2001.
- L. XIONG, C. JIANG, C.-Y. XU, K.-x. YU *et al.* S. GUO : A framework of change-point detection for multivariate hydrological series. *Water Resources Research*, 51(10):8198–8217, 2015.
- Y. YANG *et al.* Q. SONG : Jump detection in time series nonparametric regression models : a polynomial spline approach. *Annals of the Institute of Statistical Mathematics*, 66(2):325–344, 2014.
- A. ZEILEIS : p values and alternative boundaries for cusum tests. Rap. tech., Technical Report, 2001.
- A. ZEILEIS : Alternative boundaries for cusum tests. *Statistical Papers*, 45(1):123–131, 2004.
- Z. ZHOU *et al.* : Nonparametric specification for non-stationary time series regression. *Bernoulli*, 20(1):78–108, 2014.
- C. ZOU, G. YIN, L. FENG, Z. WANG *et al.* : Nonparametric maximum likelihood approach to multiple change-point problems. *Annals of Statistics*, 42(3):970–1002, 2014.