



HAL
open science

Testing a class of time-varying coefficients CHARN models with application to change-point study

Youssef Salman

► **To cite this version:**

Youssef Salman. Testing a class of time-varying coefficients CHARN models with application to change-point study. Mathematics [math]. Université de Lorraine; Université libanaise, 2022. English. NNT : 2022LORR0170 . tel-03994442

HAL Id: tel-03994442

<https://hal.univ-lorraine.fr/tel-03994442v1>

Submitted on 17 Feb 2023

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.



**UNIVERSITÉ
DE LORRAINE**

**BIBLIOTHÈQUES
UNIVERSITAIRES**

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 bibliothèque : ddoc-theses-contact@univ-lorraine.fr
(Cette adresse ne permet pas de contacter les auteurs)

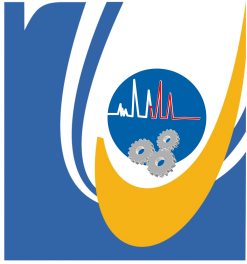
LIENS

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>



Thesis

presented for the degree of Doctor
from **Lorraine University** and **Lebanese University**

Major: Fundamental and Applied Mathematics

by
Youssef SALMAN

Title:

**Testing a class of time-varying coefficients CHARN
models with application to change-point study**

Defence scheduled on the 13rd Decembre in front of the jury made of

Supervisors:	Joseph Ngatchou Wandji	Professor at Lorraine University
	Zaher Khraibani	Professor at Lebanese University
Reviewers:	María Dolores Jiménez Gamero	Professor at Seville University
	Célestin Kokonendji (President of the jury)	Professor at Bourgogne Franche- Comté University
	Elias Ould Said	Professor at Littoral University
Examiners:	Marianne Clausel	Professor at Lorraine University
	Zainab Assaghir	Associate Professor at Lebanese University

Institut Élie Cartan de Lorraine, UMR CNRS 7502, Laboratoire de Mathématiques, BP

70239, 54 506 Vandœuvre-lès-Nancy Cedex, France
LM, Laboratoire des mathématiques, RGHC+4PR, Hadath, Beirut, Lebanon

Acknowledgements

As the time has come to put an end to this manuscript, i would like to thank the people who stand by my side to the last breath in my journey.

First of all, i want to thank God for giving me strength and strong mentality. Next, i would like to thank me for the hard work and effort i did so long.

I would like to thank my supervisor, Mr. Joseph Ngatchou-Wandji who contributed a lot in my thesis to make it in excellent scientific conditions, despite the material difficulties we have faced at the beginning. He has always been available, attentive to my questions, and always interested to follow up with me in every single detail. All the discussions that we had together and his advices made me reach the final results of my work. His analytical skills showed me that the world of research could be an exciting universe. Finally, all the proofreadings and corrections of my thesis were very appreciable. For all this, i say thank you.

I would like to thank Mr. Zaher Khraibani, the other supervisor of this thesis. Thanks for the collaboration with the IECL laboratory that gave me the chance to work with Mr. Joseph Ngatchou-Wandji, who turned out to be a very rewarding experience. In addition to his many occupations, he was always available and a good listener. His encouragement stimulated me a lot.

I would like to thank the Islamic Society of Specialization and Guidance which in turns gave me the PhD scholarship that gave me the chance to enter the world of research.

I would like to thank my family and my wife for their support, encouragement and trust. They believed in my competences and dream during these 4 years.

I am grateful to all the members of the IECL Laboratory, and their former director Xavier Antoine and to the current director Anne Gégot-Petit who ensured that each member of the laboratory would finds his place and succeeds in any mission they have. Particularly, i would like to thank Mrs Élodie Cunat, Laurence Quirot and Mr. Didier Germele for their precious cooperation.

I want to thank someone special who is Al-Sayed Moussa Al-Sader who was credited in keeping us alive in south Lebanon.

Of course, this list is not exhaustive and i thank all those who know me and made me feel like i exist ...

List of Abbreviations

\mathbb{N}	Set of positive numbers.
\mathbb{R}	Set of real numbers.
\mathbb{Z}	Set of relative integers.
\mathbb{P}	Probability.
\mathbb{E}	Expectation.
$\mathbb{1}$	The indicator function.
<i>a.s.</i>	almost surely.
$f = o_p(1)$	g tends to 0 in probability.
$\ v\ _p$	The euclidean norm of $v \in \mathbb{R}^p$
\mathcal{M}_p	Set of $p \times p$ square matrix.
$\ M\ _p$	The norm of the square matrix $M \in \mathcal{M}_p$
$\xrightarrow{\mathcal{D}}$	Converge in distribution.
$\xrightarrow{\mathbb{P}}$	Converge in probability.
$\xrightarrow{a.s.}$	Converge almost surely.
$\mathcal{B}(A)$	A Borelian of A .
\circ	Hadamard product.

Articles/Conferences

- i. Youssef Salman, Joseph Ngatchou Wandji, Zaher Khraibani. Testing a class of time varying CHARN models. JDS 2021 : 52èmes Journées de Statistique de la Société Française de Statistique (SFdS), Jun 2021, Nice (en ligne), France. hal-03329699
- ii. Youssef SALMAN, Joseph Ngatchou Wandji, Zaher Khraibani. "On detecting weak changes in the mean of CHARN models with application to change-point study". This article will be submitted in November 2022.

Talks

- i. **June 2021:** Presentation on "Testing a class of time varying CHARN models" at "52èmes Journées de Statistique de la Société Française de Statistique (SFdS)". <https://hal.archives-ouvertes.fr/hal-03329699>
- ii. **June 2019:** Participation to the "51èmes journées de Statistique organisées par la Société Française de Statistique (SFdS)", from 3 till 7 June 2019 at IECL- Nancy.

Abstract - In this thesis, we study a likelihood ratio test for detecting multiple *weak* changes in the conditional mean of a class of time dependent coefficients CHARN models. We establish the locally asymptotically normal (LAN) structure of the family of likelihoods under study. We prove that the test is asymptotically optimal, and we give an explicit form of its asymptotic local power as a function of candidates change locations and changes magnitudes. We describe some strategies for weak change-points detection and their locations estimates. The estimates are obtained as the time indices maximizing an estimate of the local power. The simulation study we conduct shows the good performance of our methods on the examples considered.

Keywords - CHARN models, Change-points, piece-wise stationarity, Optimal tests, Local Asymptotic Normality, Contiguity.

Résumé - Dans cette thèse, nous étudions un test du rapport de vraisemblance pour détecter les ruptures faibles dans la moyenne conditionnelle d'une classe de modèles CHARN à coefficients dépendants du temps. Nous établissons la structure de normalité asymptotique locale (LAN) de la famille de vraisemblances étudiées. Nous montrons l'optimalité asymptotique du test et donnons une expression explicite de sa puissance locale en fonction des potentiels points de rupture et des amplitudes des ruptures. Nous décrivons des stratégies de détection des ruptures et d'estimation de leurs localisations. Les estimateurs sont obtenus comme indices de temps rendant maximal un estimateur de la puissance locale. Les simulations numériques que nous faisons montrent de bonnes performances de notre méthode sur les exemples considérés.

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

Résumé du thèse

L'objectif principale de cette thèse est de proposer une méthode pour détecter les petits changements et estimer leurs localisations dans la moyenne conditionnelle dans une classe de modèles CHARN (Conditionnelle - Hétéroscédastique - Autoregressive - Non-linéaire) à coefficients variant dans le temps (voir Härdle et al. (1998)). Nous considérons n observations, X_1, \dots, X_n , générées par le modèle CHARN suivant

$$X_t = T(\boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(t); \mathbf{X}_{t-1}) + V(\mathbf{X}_{t-1})\varepsilon_t, t \in \mathbb{Z}, \quad (1)$$

où $(X_t)_{t \in \mathbb{Z}}$ est un processus stationnaire par morceau et érgodique dans chaque interval $[\tau_{j-1}, \tau_j[$, dont les $\tau_j, j = 0, \dots, k+1$, sont des instants potentiels de ruptures avec $\tau_0 = 1$ et $\tau_{k+1} = n$; $(\varepsilon_t)_{t \in \mathbb{Z}}$ est une bruit blanc de densité f ; $\mathbf{X}_t = (X_t, \dots, X_{t-d+1})^\top$; T et $V > 0$ sont deux fonctions réelles; $\boldsymbol{\rho}_0^\top \in \mathbb{R}^p$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_{k+1}^\top)^\top$, $\boldsymbol{\gamma}_j \in \mathbb{R}^p$ pour tout $j = 1, \dots, k+1$; $\boldsymbol{\omega}(t) = (\mathbb{1}_{[\tau_0, \tau_1)}(t), \mathbb{1}_{[\tau_1, \tau_2)}(t), \dots, \mathbb{1}_{[\tau_{k-1}, \tau_k)}(t), \mathbb{1}_{[\tau_k, \tau_{k+1})}(t))^\top = (\omega_1, \dots, \omega_{k+1}) \in \{0, 1\}^{k+1}$; pour tout $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_{k+1}^\top)^\top \in \mathbb{R}^{p(k+1)}$ et $\boldsymbol{\omega} = (\omega_1, \dots, \omega_{k+1})^\top \in \mathbb{R}^{k+1}$,

$$\boldsymbol{\gamma} \odot \boldsymbol{\omega} = \boldsymbol{\gamma}_1 \omega_1 + \dots + \boldsymbol{\gamma}_{k+1} \omega_{k+1} \in \mathbb{R}^p,$$

où pour $i = 1, \dots, k+1$,

$$\boldsymbol{\gamma}_i \omega_i = (\gamma_{i,1} \omega_i, \dots, \gamma_{i,p} \omega_i) \in \mathbb{R}^p.$$

Nous construisons un test du rapport de vraisemblance pour tester

$$H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \quad \text{contre} \quad H_\beta^{(n)} : \boldsymbol{\gamma} = \boldsymbol{\gamma}_n = \boldsymbol{\gamma}_0 + \frac{\boldsymbol{\beta}}{\sqrt{n}},$$

pour $\boldsymbol{\gamma}_0 = (\boldsymbol{\gamma}_{0,1}^\top, \dots, \boldsymbol{\gamma}_{0,k+1}^\top)^\top$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_{k+1}^\top)^\top \in \mathbb{R}^{p(k+1)}$, avec $\boldsymbol{\gamma}_{0,i}, \boldsymbol{\beta}_i \in \mathbb{R}^p$, $i = 1, \dots, k+1$. Parmi les $\boldsymbol{\gamma}_i$, certaines peuvent être nulles, signifiant la présence des paramètres de nuisance dans le modèle. Ce problème de test prolonge le travail de Ltaifa (2021) et Ngatchou-Wandji and Ltaifa (2021) qui se limite à tester la moyenne. Notre objectif principal dans cette thèse est de généraliser ces travaux.

Premièrement, nous supposons le cas où les paramètres $\boldsymbol{\rho}_0$ et $\boldsymbol{\gamma}_0$ sont connues. Nous calculons la vraisemblance de $\mathbf{X}^{(n)} = (X_n, \dots, X_1)$ conditionnelle à $\mathbf{X}_0 \in \mathbb{R}^d$. Puis, nous trouvons le logarithme du rapport de vraisemblance de H_0 contre $H_\beta^{(n)}$ qui a l'expression suivante

$$\Theta_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) - \Delta_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1),$$

avec

-
- $\Delta_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{1}{2n} \sum_{t=1}^n \left\{ \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right. \\ \left. - \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\}$
 - $\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\}$
 - $M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) N^\top(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})$
 - $N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = \omega(t) \circ D_\gamma [T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]$
 - $\mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = \omega(t) \omega(t)^\top \circ H_\gamma [T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]$
 - $N(\boldsymbol{\gamma}, \mathbf{X}_{t-1}) = \left(\omega_1 \partial_{\gamma_1} [T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})], \dots, \omega_{k+1} \partial_{\gamma_{k+1}} [T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})] \right)^\top \in \mathbb{R}^{p(k+1)}$
 - $\omega_i \in \{0, 1\}$,

avec

$$M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) N^\top(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = \begin{pmatrix} \omega_1^2 \partial_{\gamma_1} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\gamma_1}^\top T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) & \dots & \omega_1 \omega_{k+1} \partial_{\gamma_1} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\gamma_{k+1}}^\top T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \omega_{k+1} \omega_1 \partial_{\gamma_{k+1}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\gamma_1}^\top T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) & \dots & \omega_{k+1}^2 \partial_{\gamma_{k+1}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\gamma_{k+1}}^\top T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \end{pmatrix} \\ \in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

et

$$\mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = \begin{pmatrix} \omega_1^2 \partial_{\gamma_1}^2 T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) & \dots & \omega_1 \omega_{k+1} \partial_{\gamma_1 \gamma_{k+1}}^2 T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \omega_{k+1} \omega_1 \partial_{\gamma_{k+1} \gamma_1}^2 T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) & \dots & \omega_{k+1}^2 \partial_{\gamma_{k+1}}^2 T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

Ensuite, dans le but de trouver la distribution du test construit sous H_0 , nous démontrons que, lorsque $n \rightarrow +\infty$, pour tout $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, sous H_0 ,

- i. $\Delta_n \rightarrow \frac{\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{2}$,
- ii. $\Pi_n \rightarrow \mathcal{N}(0, \eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}))$,

tels que, pour $1 \leq h \leq m \leq p$,

$$\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0), \\ \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) dF_j(x).$$

Ainsi, la propriété LAN (Localement asymptotiquement normale) énoncée dans Dreesbeke and Fine (1996) est vérifiée pour la suite centrale $\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$. D'après la Proposition 4.2 de Dreesbeke and Fine (1996) (troisième lemme de Le Cam), les suites d'hypothèses $\{H_{\boldsymbol{\beta}}^{(n)} : n \geq 1\}$ et $\{H_0^{(n)} = H_0 : n \geq 1\}$ sont contiguës et sous $H_{\boldsymbol{\beta}}^{(n)}$, on a

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} \mathcal{N}(\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}), \eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})).$$

Pour tester H_0 contre $H_{\boldsymbol{\beta}}^{(n)}$, pour $\boldsymbol{\gamma}_0$ connu et pour tout $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, nous considérons la statistique du test suivante

$$\mathcal{T}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}, \quad (2)$$

avec

- $\widehat{\pi}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \sqrt{\widehat{\eta}_n(\boldsymbol{\gamma}_0, \boldsymbol{\beta})}$,
- $\widehat{\eta}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \sum_{j=1}^{k+1} \widehat{\alpha}_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \widehat{\eta}_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)$,
- $\widehat{\eta}_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) d\widehat{F}_j(x)$, où \widehat{F}_j est la fonction de distribution empirique du vecteur \mathbf{X}_t dont les composantes sont dans $[\tau_{j-1}, \tau_j]$.
- Une estimation de $\eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)$ peut être considéré comme

$$\widehat{\eta}_{j,2}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = \frac{I(f)}{n_j(n)} \sum_{\tau_{j-1}}^{\tau_j} \frac{1}{V^2(\mathbf{X}_{t-1})} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}),$$

$$\text{et } \widehat{\alpha}_j = \frac{n_j(n)}{n}.$$

Ensuite, nous démontrons que, sous H_0 ,

$$\mathcal{T}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$

et sous $H_{\boldsymbol{\beta}}^{(n)}$, la puissance asymptotique locale du test basé sur $\mathcal{T}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ est égale à

$$\mathcal{P}_{k,\tau^k} = 1 - \Phi(z_\alpha - \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})),$$

où Φ est la fonction de répartition de la loi normale centrée réduite et z_α est son quantile d'ordre $(1 - \alpha)$. Finalement, nous démontrons que le test basé sur $\mathcal{T}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ est localement asymptotiquement optimal.

Maintenant, nous nous plaçons dans le cadre du modèle (1) avec les fonctions T et V de

forme connues mais le paramètre $\boldsymbol{\rho}_0$ est inconnue. Tout d'abord, nous considérons un estimateur consistant de $\boldsymbol{\rho}_0$, noté par $\boldsymbol{\rho}_n$ et qui verifie la représentation de Bahadur et al. (1960) suivante

$$n^{\frac{1}{2}}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) = n^{-\frac{1}{2}} \sum_{t=1}^n \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \aleph(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)) + o_P(1), \quad (3)$$

où

- $\Upsilon(x, \boldsymbol{\rho}_0) = (\Upsilon_1(x, \boldsymbol{\rho}_0), \dots, \Upsilon_p(x, \boldsymbol{\rho}_0))^\top \in \mathbb{R}^p$
- Pour tout $j = 1, \dots, k+1$, $x \in \mathbb{R}^d$, $\exists a \geq 0$ telle que $\int_{\mathbb{R}^d} \|\Upsilon(\boldsymbol{\rho}_0, x)\|^{2+a} dF_j(x) < \infty$.
- $\int_{\mathbb{R}} |\aleph(x)|^{2+a} f(x) dx < \infty$ et $\int_{\mathbb{R}} \aleph(x) f(x) dx = 0$,

Ensuite, nous trouvons que, sous H_0 ,

$$\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma),$$

et sous $H_{\boldsymbol{\beta}}^{(n)}$,

$$\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathcal{C}, \Sigma),$$

où

$$\mathcal{C} = \int_{\mathbb{R}} \aleph(x) \phi_f(x) f(x) dx \sum_{j=1}^n \alpha_j \sum_{h=1}^p \beta_{j,h} \int_{\mathbb{R}^d} \frac{\Upsilon(\boldsymbol{\rho}_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, x) dF_j(x) \in \mathbb{R}^p,$$

et

$$\Sigma = \int_{\mathbb{R}} \aleph^2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \Upsilon(\boldsymbol{\rho}_0, x) \Upsilon^\top(\boldsymbol{\rho}_0, x) dF_j(x) \in \mathcal{M}_p(\mathbb{R}).$$

Ensuite, nous supposons que $\boldsymbol{\gamma}_0$ est connue et nous demontrons la convergence de la suite centrale à version estimé vers la suite centrale à version exacte. En d'autre terme, nous demontrons que, sous certaines conditions, pour tout $n \geq 0$, pour tout estimateur consistant et asymptotiquement normale $\boldsymbol{\rho}_n$ de $\boldsymbol{\rho}_0$, nous avons, pour tout suite des entiers positives $s(n)$ telle que $n/s(n) \rightarrow 0$, lorsque $n \rightarrow +\infty$.

i.

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1),$$

ii.

$$\widehat{\pi}_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \longrightarrow \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}).$$

Dans ce cas, pour tester H_0 contre $H_{\beta}^{(n)}$, pour tout $\beta \in \mathbb{R}^{p(k+1)}$, nous considérons la statistique du test suivante

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}.$$

Ensuite, nous démontrons que, sous H_0 , lorsque $n \rightarrow +\infty$,

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1),$$

et sous $H_{\beta}^{(n)}$, pour un seuil de signification $\alpha \in]0, 1[$, la puissance locale asymptotique du test basé sur $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ est

$$\mathcal{P}_{k, \tau^k} = 1 - \Phi(z_{\alpha} - \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})),$$

où Φ est la fonction de répartition de la loi normale centrée réduite et z_{α} est son quantile d'ordre $(1 - \alpha)$. Finalement, nous démontrons que le test basé sur la statistique $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ est localement asymptotiquement optimale.

En suivant les mêmes techniques, nous démontrons que, en remplaçant $\boldsymbol{\gamma}_0$ par un estimateur consistant $\widehat{\boldsymbol{\gamma}}_{0,s(n)}$, la statistique du test $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})$ définit par

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})},$$

converge sous H_0 vers une variable aléatoire qui suit une loi normale centrée réduite. Sous $H_{\beta}^{(n)}$, pour un seuil de signification $\alpha \in]0, 1[$, la puissance asymptotique du test basé sur $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ est

$$\mathcal{P}_{k, \tau^k} = 1 - \Phi(z_{\alpha} - \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})),$$

où Φ est la fonction de répartition de la loi normale centrée réduite et z_{α} est son quantile d'ordre $(1 - \alpha)$. Finalement, nous démontrons que le test basé sur la statistique $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})$ est localement asymptotiquement optimale.

Pour utiliser les résultats théoriques achevés, nous construisons un nouveau algorithme, motivé par le calcul de la puissance locale, pouvant détecter les ruptures d'une manière séquentielle. En utilisant des données simulées par le logiciel R, nous démontrons que notre méthode est efficace et capable à détecter les petites ruptures dans une classe de modèle linéaire, soit AR(1), et non linéaire, soit AR(1)-ARCH(1).

Enfin, nous comparons la performance de notre méthode avec la méthode de Horváth et al. (2020). Les résultats montrent que notre méthode est plus puissante et plus précise.

Contents

1	General introduction	1
1.1	Introduction	1
1.2	A review on change-point study	2
1.3	Our main contributions	7
1.4	Outline of the thesis	7
2	Mathematical tools	9
2.1	Introduction	9
2.2	On testing and estimating time series models	10
2.2.1	Linear process	10
2.2.2	Non-linear process	11
2.2.3	Parameters estimation methods	12
2.2.4	Tests for breaks detection	14
2.3	Some important notions	17
2.3.1	Identifiability	17
2.3.2	Stationarity	18
2.3.3	Ergodicity	19
2.3.4	Local asymptotic normality	21
2.3.5	Contiguity	22
2.3.6	Le Cam's Third Lemma	23
2.3.7	Central limit theorem of martingales	24
3	Testing the conditional mean of a time-dependent coefficients CHARN model: Part I	25
3.1	Introduction	25
3.2	Notation and technical assumptions	26
3.2.1	Notation	26
3.3	General assumptions	28

3.4	The functions involved in the model and the parameters are known.	29
3.4.1	Likelihood ratio test	29
3.4.2	Asymptotic behavior of the test statistic	42
3.5	Conclusion	50
4	Testing the conditional mean of a time-dependent coefficients CHARN model: Part II	51
4.1	Introduction	51
4.2	The functions involved in the model and the nuisance parameters are unknown	51
4.2.1	Parameters estimation	52
4.2.2	Asymptotic behavior of the new test statistic	61
4.3	The functions involved in the model and the parameters are unknown. . .	81
4.4	Conclusion	103
5	Application to change-point study	105
5.1	Introduction	105
5.2	Some algorithms for change-point detection and the location estimation . .	105
5.2.1	Automatic procedure based on parametric autoregressive model (<i>Auto-PARM</i>)	106
5.2.2	Smooth localized complex exponentials (<i>Auto-SLEX</i>) functions approach	106
5.2.3	Informational approaches	107
5.2.4	Binary segmentation	108
5.3	Some recent methods	109
5.3.1	Horváth et al. (2020)	109
5.3.2	Ngatchou-Wandji and Ltaifa (2021)	110
5.4	Algorithm for break detection and location estimation	110
5.4.1	The algorithm	110
5.4.2	Comments	112
6	Practical considerations	113
6.1	Introduction	113
6.2	Numerical study	113
6.2.1	Power study	114
6.2.2	Breaks detection	119
6.3	Comparison with Horváth et al. (2020)	129
6.4	Conclusion	130

7	General conclusion and outlook	131
7.1	Conclusion	131
7.2	Outlook	132

List of Figures

6.1	Power of test against a shifted white noise.	115
6.2	Power of test in case of change in autocorrelation.	115
6.3	Power against an AR(1).	117
6.4	Power against an AR(1).	118
6.5	Level of significance 5%.	118
6.6	Level of significance 1%.	118
6.7	No break ($\alpha = 5\%$).	119
6.8	No break in the data.	119
6.9	Estimation of the break location for $\tau_1 = 100$ in a class of AR(1) models (first case).	121
6.10	Estimation of the break location for $\tau_1 = 100$ in a class of AR(1) models (second case).	122
6.11	Estimation of break location for different τ_1 in a class of AR(1) model (third case).	123
6.12	Breaks locations estimation in the case of two breaks.	125
6.13	Behavior of the local power of the test in case of three breaks in a class of shifted models.	126
6.14	Detection of multiple change-points in a class of ARCH models	129

Chapter 1

General introduction

1.1 Introduction

The main purpose of our work is the study of *weak* changes detection and the location estimation of the changes in the conditional mean of a class of time-dependent coefficients Conditional Heteroscedastic AutoRegressive Nonlinear (CHARN) models (see Härdle et al. (1998)). We mean by *weak* change that whose magnitude is too small. Such a change can happen for example in a phenomenon which changes gradually.

We consider n observations, X_1, \dots, X_n , generated by the following CHARN model

$$X_t = T(\boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(t); \mathbf{X}_{t-1}) + V(\mathbf{X}_{t-1})\varepsilon_t, t \in \mathbb{Z} \quad (1.1)$$

where $(X_t)_{t \in \mathbb{Z}}$ is a piece-wise stationary and ergodic process in each segment $[\tau_{j-1}, \tau_j]$, where the τ_j , $j = 0, \dots, k+1$, are potential instants of change with $\tau_0 = 1$ and $\tau_{k+1} = n$ and n denotes the sample size; $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a standard white noise with density f ; $\mathbf{X}_t = (X_t, \dots, X_{t-d+1})^\top$; T and $V > 0$ are real smooth functions; $\boldsymbol{\rho}_0^\top \in \mathbb{R}^p$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_{k+1}^\top)^\top$, $\boldsymbol{\gamma}_j \in \mathbb{R}^p$ for any $j = 1, \dots, k+1$; $\boldsymbol{\omega}(t) = (\mathbb{1}_{[\tau_0, \tau_1)}(t), \mathbb{1}_{[\tau_1, \tau_2)}(t), \dots, \mathbb{1}_{[\tau_{k-1}, \tau_k)}(t), \mathbb{1}_{[\tau_k, \tau_{k+1})}(t))^\top = (\omega_1, \dots, \omega_{k+1}) \in \{0, 1\}^{k+1}$; for any $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_{k+1}^\top)^\top \in \mathbb{R}^{p(k+1)}$ and $\boldsymbol{\omega} = (\omega_1, \dots, \omega_{k+1})^\top \in \mathbb{R}^{k+1}$,

$$\boldsymbol{\gamma} \odot \boldsymbol{\omega} = \boldsymbol{\gamma}_1 \omega_1 + \dots + \boldsymbol{\gamma}_{k+1} \omega_{k+1} \in \mathbb{R}^p,$$

where for $i = 1, \dots, k+1$,

$$\boldsymbol{\gamma}_i \omega_i = (\gamma_{i,1} \omega_i, \dots, \gamma_{i,p} \omega_i) \in \mathbb{R}^p.$$

We construct a log-likelihood ratio test for testing

$$H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \quad \text{against} \quad H_{\boldsymbol{\beta}}^{(n)} : \boldsymbol{\gamma} = \boldsymbol{\gamma}_n = \boldsymbol{\gamma}_0 + \frac{\boldsymbol{\beta}}{\sqrt{n}}, \quad (1.2)$$

for some $\boldsymbol{\gamma}_0 = (\boldsymbol{\gamma}_{0,1}^\top, \dots, \boldsymbol{\gamma}_{0,k+1}^\top)^\top$, $\boldsymbol{\beta} = (\boldsymbol{\beta}_1^\top, \dots, \boldsymbol{\beta}_{k+1}^\top)^\top \in \mathbb{R}^{p(k+1)}$, with $\boldsymbol{\gamma}_{0,i}$, $\boldsymbol{\beta}_i \in \mathbb{R}^p$, $i = 1, \dots, k+1$. Some of $\boldsymbol{\gamma}_i$'s may be nil, meaning that there may be nuisance parameters in the model. This testing problem extends the work of Ltaifa (2021) and Ngatchou-Wandji and Ltaifa (2021) who restrict to testing the mean. Our main purpose in this thesis is to generalize these works.

1.2 A review on change-point study

The first work on change-points study started with Page (1954), in the context of iid data. The monograph of Csörgő et al. (1997) presents notions and theory of change-points. The literature on change-points is large and various. In this section, we review some change-point papers, focusing on time series.

MacNeill (1974) proposes a test for parameter changes at unknown points using a Brownian Motion test. The observations are assumed to follow an exponential distribution. He presents a derivation, using the method of Chernoff and Zacks (1964). He uses the likelihood ratio statistic for testing "no change" against the existence of change. He studies the large sample distribution theory of the test statistic, and obtains its asymptotic distribution. Finally, he studies the distributions of certain functionals of Brownian Motion and illustrates the relation between the means obtained by a simulation experiment.

Davis et al. (1995) study the problem of testing the change in parameter values when it has not yet occurred. The authors show that if the white noise is weakly stationary with finite fourth moments, then under the null hypothesis (no change-point), the normalized Gaussian likelihood ratio test statistic converges in distribution to the Gumbel extreme value distribution. They investigate the null hypothesis behavior of the likelihood ratio statistic when the orders of the **AR** model are permitted to be different before and after the change-point. They study the change in the variance of the white noise before and after the change-point. Also, they check that the likelihood ratio test converges in probability under two conditions. They illustrate their theoretical results with a simulation experiment.

Vogelsang (1997) proposes tests statistics for detecting a break at an unknown date in the trend function of a dynamic univariate time series. The tests are based on the mean and exponential statistics of Andrews and Ploberger (1994) and the supremum statistic of Andrews (1993). Their results are extended to allow trending and unit root regressors. Asymptotic results are obtained. When the errors are highly persistent and it is not known which asymptotic theory provides a better approximation, the authors provide a conservative approach based on nearly integrated asymptotics. They show the power of

the mean statistic to be nonmonotonic with respect to the break magnitude and is dominated by the exponential and supremum statistics. Also, they propose different versions of the tests applicable to first differences of the data. Finally, they applied the tests to some macroeconomic time series, and the null hypothesis of a stable trend function is rejected in many cases.

Another method for detecting change-points is introduced in Lavielle and Lebarbier (2001). The authors present a multiple change-point analysis for which Markov Chain Monte Carlo (MCMC) sampler plays a fundamental role. They propose an attractive methodology for the change-points problem in a Bayesian context. It is used for estimating the posterior distribution of the unknown sequence of change-points instants, and also for estimating the hyperparameters of the model. The reversible jump algorithm is based on the fact that the dimension of the model can change according to the number of changes. Unfortunately, this algorithm converges slowly, and many iterations are needed for estimating correctly the posterior distribution. The authors start their article by describing the model of change-points in the mean and detail the prior modeling. They next use the Hastings-Metropolis algorithm to estimate the posterior distribution of the change-points instants. They explain the problem of recovering the sequence of means. The reversible jump algorithm is presented and finally, they explain the SAEM algorithm (a stochastic approximation of expectation maximization) for the estimation of the hyperparameters of the model. They use the Bayesian approach to estimate the number of change-points. They determine the posterior distribution of the non-observed sequences conditionally to the set of observations. After finding this distribution using the Markov chain Monte-Carlo method, they start the Hastings-Metropolis algorithm. They explain the basic algorithm, and to illustrate it, they simulate a sequence of 500 observations. Finally, they conclude that the MCMC is faster than the Reversible Jump algorithm for the determination of the posterior distribution of change-points.

Lebarbier (2005) also studies the problem of detecting change-points in the mean of a signal corrupted by additive noise. The number of changes and their position are unknown. These are estimated by a method based on a penalized least-square criterion. The problem in this paper is to detect and locate the change-point instants and to estimate the jumps in the mean. The author shows that when the variance of the noise is known, the two penalty constants are calibrated optimally. Next, the heuristic method to estimate the noise variance is presented and a calibration of this method is proposed. Finally, different simulation studies are done to assess the performance of the calibrated method proposed.

Davis et al. (2008) consider the problem of detecting break points for a nonstationary time series. Specifically, the time series is assumed to follow a parametric nonlinear time-

series model in which the parameters may change at fixed times. The authors assume that the number and locations of the break-points are unknown. They use the minimum description length (MDL) as a criterion for estimating the number of break-points, the locations of the break-points and the parametric model in each segment. They find the best segmentation by minimizing MDL using a genetic algorithm. The authors illustrate their results using generalized autoregressive conditionally heteroscedastic (GARCH) models, stochastic volatility models and generalized state-space models as the parametric model for the segments. An empirical study shows the good performance of their methods.

Fotopoulos et al. (2009) study the problem of detecting the instant of change. The use of the maximum likelihood estimates the instant of the change $\hat{\tau}_n$, but instead of working with a discrete value, the authors estimate a continuous number between 0 and 1 after assuming that the instant of change is the product of the sample size with this number. They study the asymptotic distribution of the test used by contiguity. After the proof of the convergence of the test, the authors apply the theoretical part to simulated data.

To monitor the parameters (mean, variance, and parameter of the model), the Brownian Motion method is used by Gombay and Serban (2009) to detect changes. In an autoregressive model, the authors introduce a new change parameter and estimate it in case of change. Finally, they find a relation between the score and the Gaussian process, and they prove that the difference between them (in addition to constant) tends to zero almost surely. In other words, the authors prove the effectiveness of the test. They end their work by a simulation experiment.

Huh (2010) investigates the regression function or its ν th derivative in generalized linear models which may have a change (discontinuity) point at an unknown location. The estimation of this change is based on one-sided kernel weighted local-likelihood functions. The asymptotic distribution is also studied for the estimators of location and jump size. The theoretical part is illustrated by a simulation and beetle mortality examples.

Shao and Zhang (2010) considers the CUSUM-based (cumulative sum) test for detecting a change in a time series. The authors propose a self-normalization (SN) based Kolmogorov–Smirnov test, where the formation of the self-normalizer takes the change-point alternative into account. They prove that the resulting test statistic is asymptotically distribution free and its power is monotonic. They extend the SN-based test to test for a change in other parameters associated with a time series, such as marginal median, autocorrelation at lag one, and spectrum at certain frequency bands. Monte-Carlo simulations are conducted to compare the finite sample performance of the new SN-based test with the traditional Kolmogorov–Smirnov test. Illustrations using real data examples are presented.

Jarušková (2010) studies the change-point in the mean of a sequence of independent nor-

mally distributed random vectors. The asymptotic distribution of the test statistic is studied by using results from Pit (1994). Also, on the same subject, Chen et al. (2011) study this problem for independent normal means as a multiple testing problem. The authors consider two stepwise methods, the binary segmentation method of Vostrikova (1981), and the maximum residual down method of Cohen et al. (2009). They prove the consistency of these two methods. More precisely, they proved that the probability of making errors in the acceptance or the rejection of the hypothesis tends to zero when the sample size tends to infinity.

Ciuperca (2011) studies the number of change-points using an M -estimator procedure. BIC with and without changes are defined. The important aspect of this paper is that the number of observations between two instants of changes is unknown. Maximizing the M -criterion leads to the estimation of the number of change-points. The consistency is proved. The results are illustrated with a simulation study.

Döring (2011) studies the asymptotic distribution for a class of multiple change-point estimators. The author considers a nonparametric model and uses the maximizing point of a weighted multivariate U -statistic process to estimate the change in the parameters. The most important idea presented in this article is the transition from the maximization of the parameter to the maximization of another process while maintaining the same goal. The detection of change is studied in Prášková and Chochola (2014). The authors study the existence of the change in the regression parameters in a linear model where the regressors and the errors are weakly dependent. They use the M -estimators and the Weighted M -residuals to construct a test statistic. They also study the asymptotic distribution under the null hypothesis and contiguous alternatives. The results are illustrated by a simulation experiment.

The problem of changes can be studied not just in time series and observations, but also in the effect of the change in the mean on the density function (tail, peak, *etc.*). In Dupuis et al. (2015) this problem is studied. The authors develop a new method of detecting and estimating the change-points in the tail of multiple time series data. They explain the change in the mean, in the variance, and in the tails. They also discuss the effect of the mean and variance's change on the tails. They focus on the detection of change-points in the upper-tail of the distribution of the variable of interest basing on multiple cross-sectional time series. They start their work by identifying the change-points. Next, they consider a tail model and a likelihood ratio test to study the change in the parameters of this model. Their null hypothesis is "no change in the parameters" and the alternative is the "existence of the change". This method is based on approximating the distribution of the exceedances over a high threshold by the GP (generalized Pareto) distribution. Next, the authors compare the performance of the methods used. Finally, they apply these

methods to a set of real data that represent the temperatures in New York City between 1956 and 2005.

In Badagián et al. (2015), the problem of detecting and estimating the change-points location in a time series is studied. The authors consider the problem of modeling a non-stationary time series by segmenting it into blocks that are fitted by stationary processes. This paper proposes a procedure based on the *BIC* joint with the binary segmentation algorithm to look for changes in the mean, the autoregressive coefficients, and the variance of the perturbation in piece-wise autoregressive processes, by using a procedure and comparing this procedure with several others available in the literature. The authors present the change-points problem and explain briefly the Auto-PARM and Auto-SLEX methods. They present different algorithms that are useful to the search for multiple change-points, and they compute and compare the size and the power of the presented approaches. Finally, they apply them to real data from speech recognition.

Yau and Zhao (2016) propose a likelihood ratio scan method for estimating multiple change points in piecewise stationary processes. The authors use scan statistics which reduces the computationally infeasible global multiple-change-point estimation problem to a number of single-change-point detection problems in various local windows. They prove the consistency for the estimated numbers and locations of the change points. Moreover, they develop a procedure for constructing confidence intervals for each of the change points. Finally, the authors illustrate their results and the efficiency of the likelihood ratio scan method by simulation experiments and real data analysis.

Ruggieri and Antonellis (2016) aim to estimate the instant of change in a regression model. They use a sequential Bayesian change-point algorithm that provides uncertainty bounds on both the number and location of the change. The most important thing is that this algorithm can quickly update itself in linear time as each new data point is recorded and uses the exact posterior distribution to infer whether or not a change-point has been observed. The authors illustrate the theoretical part with simulation data and real data about the temperature anomalies over the last 130 years (the real data is taken from the National Oceanic and Atmospheric Administration NOAA).

A new class of change-point test statistics is proposed in Horváth et al. (2020) that utilizes a weighting and trimming scheme for the cumulative sum (CUSUM) process inspired by Rényi. Using an asymptotic analysis and simulations, the authors demonstrate that this new class of statistics possess superior power compared to traditional change-point statistics based on the CUSUM process when the change-point is near the beginning or end of the sample. The authors develop a generalization of these “Rényi” statistics in order to testing for changes in the parameters in linear and nonlinear regression models, and in generalized method of moments estimation. In these contexts, the authors apply

the proposed statistics, as well as several others, to test for changes in the coefficients of Fama–French factor models. They observe that the Renyi statistic is the most effective in terms of retrospectively detecting change-points that occur near the endpoints of the sample.

1.3 Our main contributions

Very few of the papers reviewed in the preceding section study the weak change problems. To contribute to filling this gap, Ltaifa (2021) and Ngatchou-Wandji and Ltaifa (2021) propose a new approach for weak change-points study. They consider the testing problem presented in Section 1.1, for the particular case where $T(\boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \omega(t), \mathbf{X}_{t-1}) = \boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \omega(t) + U(\mathbf{X}_{t-1})$, for some function U . Then they prove a LAN property and the contiguity of the hypotheses tested. From these results, they compute the theoretical local power of their likelihood-ratio test, and use it in an algorithm for detecting changes and estimating their locations. In our work, we proceed as them, but with a more general model and a new change detection algorithm. Our main contributions are:

- A locally asymptotically normal (LAN) result, from which we deduce the contiguity of the hypotheses tested.
- From the contiguity result, using Le Cam third Lemma, we give an explicit expression for the local power of the test studied. This power is a function of the potential change-points and the magnitudes of the changes.
- Considering an estimated version of the power with the magnitudes of the changes replaced by their estimators, we propose an algorithm for change-points detection and for estimating the changes locations. This algorithm finds change-points at blind, sequentially.
- We conduct a simulation experiment which shows that our method performs well on the examples tested and compares favorably with other methods.

1.4 Outline of the thesis

This thesis is structured as follows.

In Chapter 2, we present the main tools that we use in our work. We present some usual time series models and recall some main methods for parameter estimation in these

models. We also describe the main methods used for testing change-points in time series models. Next, we recall some important notions as identifiability, stationarity and ergodicity. We recall some notions of the Le Cam theory as LAN and contiguity. We finally recall a central limit theorem from Hall and Heyde (2014).

In Chapter 3, we study the problem presented in the general introduction in the case where all the parameters are known. We first precise the notations and make the general assumptions. Our likelihood-ratio test is derived and its asymptotic properties are studied under the null and under the local alternatives considered. For the study under the local alternatives, we use the contiguity of the hypotheses under study that we previously established. An explicit expression of the local power is provided.

In Chapter 4, all the parameters involved in the problem are no more known. We show that replacing them by their estimators, the results of Chapter 3 remained valid.

In Chapter 5, we apply the results of Chapter 4 to change-point detection and to estimating their locations. We describe an algorithm for multiple change-points detection and for estimating their locations. We previously give a brief review of some existing change-point detection algorithms.

In Chapter 6, our algorithm is applied to simulated data from a class of CHARN(1,1) model for detecting change-points and estimating their locations.

The last chapter concludes of our work and outlines some perspectives.

Chapter 2

Mathematical tools

2.1 Introduction

In this chapter, we present some definitions of linear and nonlinear models and introduce some methods for their parameters estimation. We also review some tests for breaks detection. Finally, we present the fundamental notions, notations and definitions that we use. One of these notions is the **LAN** property from the theory of Le Cam (1960). It is of a primary importance in our work.

The **LAN** property used is the modified version defined in Hall and Mathiason (1990) and used in Hwang and Basawa (2001) to establish the asymptotic normality of the likelihood ratio test under the null hypothesis.

We use other tools like central limit theorem of martingals presented in Hall and Heyde (2014) and the ergodic theorem (see Birkhoff (1931)). These tools are also those used in Ngatchou-Wandji and Ltaifa (2021), Lounis (2012), Lounis (2011), and can be found in Le Cam (1960), Hall and Heyde (2014) and Dreesbeke and Fine (1996). The outline of this chapter is as follow:

We present some time series models, some methods for estimating their parameters, and some tests. We also present some tests for detecting breaks and some methods for estimating the break-locations. Finally, we outline and present the main mathematical tools used in our work.

2.2 On testing and estimating time series models

In this section, we recall the definition of linear and nonlinear models and we give some examples. Next, we present some methods used in literature to estimate the models parameters. Furthermore, we enumerate some tests used in literature for breaks detection.

2.2.1 Linear process

A stochastic process $\{X_t\}_{t \in \mathbb{Z}}$ is said to be linear if, for any $t \in \mathbb{Z}$, X_t can be written as a weighted infinite sum

$$X_t = \rho + \sum_{i=0}^{+\infty} \alpha_i \varepsilon_{t-i}, \quad t \in \mathbb{Z},$$

where ρ is a constant, the α_i are real coefficients with $\alpha_0 = 1$ and $(\varepsilon_t)_t$ is a sequence of iid zero-mean random variables with variance $\sigma_\varepsilon^2 < +\infty$ such that $\mathbb{E}(\varepsilon_t | X_{t-1}, X_{t-2}, \dots) = 0$. Some examples of linear process are ARMA, ARIMA, SARIMA.

Let d, D, p, P, q, Q and s be non negative integers. Let B is the lag operator. Define the following operators.

- $\nabla^d = (1 - B)^d$
- $\nabla_s^D = (1 - B^s)^D$
- $\Phi(B) = 1 - \phi_1 B - \dots - \phi_p B^p$, where $\phi = (\phi_1, \dots, \phi_p) \in \mathbb{R}^p$ and $\phi_p \neq 0$
- $\Phi'(B) = 1 - \phi'_1 - \dots - \phi'_P B^P$, where $\phi' = (\phi'_1 - \dots - \phi'_P) \in \mathbb{R}^P$ and $\phi'_P \neq 0$
- $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, where $\Theta = (\theta_1, \dots, \theta_q) \in \mathbb{R}^q$ and $\theta_q \neq 0$
- $\Theta'(B) = 1 - \theta'_1 B - \dots - \theta'_Q B^Q$, where $\Theta' = (\theta'_1, \dots, \theta'_Q) \in \mathbb{R}^Q$ and $\theta'_Q \neq 0$

Definition 2.2.1. We say that $(X_t)_{t \in \mathbb{Z}}$ is an ARIMA process of order (p, d, q) , denoted by ARIMA(p, d, q), if:

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

Definition 2.2.2. We say that $(X_t)_{t \in \mathbb{Z}}$ is a SARIMA process of order $(p, d, q)(P, D, Q)$, denoted by SARIMA($p, d, q)(P, D, Q)$, if:

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

2.2.2 Non-linear process

Some phenomena essentially display nonlinear behavior, such as time irreversibility, asymmetry, and self-sustained stochastic cyclical behavior, which cannot be explained well by linear time series models. Realizing this, researchers have proposed numerous nonlinear time series models since the late 1970s.

Let $(X_t)_{t \in \mathbb{Z}}$ be the process defined, for any $t \in \mathbb{Z}$, by

$$X_t = g(\varepsilon_t, \varepsilon_{t-1}, \dots), \quad (2.1)$$

where (ε_t) is a white noise and g is a function. If g is linear, then (X_t) is a linear process and (2.1) is a linear model. If g is non-linear, we obtain a non-linear model. Some examples of non-linear models are ARCH, TARARCH and EXPAR models studied in Engle (1982), Tong (1990) and Guégan et al. (1994), and bilinear models proposed by Granger and Andersen (1978).

ExpAR models

ExpAR model was first used to model the ship rolling data. There is a lot of work on ExpAR models, (see, e.g., Ozaki and Oda (1977), Ghosh et al. (2015)).

Definition 2.2.3. *A time series $(X_t)_{t \in \mathbb{Z}}$ is an ExpAR model of order p , if for any $t \in \mathbb{Z}$,*

$$X_t = \sum_{i=1}^p \left(c_i + \pi_i e^{-\gamma X_{t-1}^2} \right) X_{t-i} + \varepsilon_t,$$

where $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a white noise such that ε_t independent of X_t , and c_i , π_i and γ are the unknown parameters which need to be estimated.

Later, many versions of ExpAR models were introduced (see, e.g., Ozaki (1985), LeBaron (1992), Koutmos (1997)).

Definition 2.2.4. *An ARCH(q) models can for example be defined as follows:*

$$\varepsilon_t = \sigma_t z_t,$$

where ε_t represents the innovations of the series, and z_t is a standard white noise and $\sigma_t^2 = a_0 + \sum_{i=1}^q a_i \varepsilon_{t-i}^2$ with a_i , $i = 0, \dots, q$, are reals such that $a_0 > 0$, $a_1 \geq 0$ and $a_q \geq 0$ guaranteeing the strict positivity of the conditional variance.

It is easy to see that ExpAR and ARCH models are within our CHARN models (1.1).

Time-varying coefficients models.

In real life, it may be necessary to adjust time-varying coefficients models to some data. Some examples of such data are those from climate change and a machine in a production line. The model (1.1) that we study in this thesis and the models presented in Section 2.2.1 are examples of this type of models, studied by Davis et al. (1995), Badagián et al. (2015) and Dehling et al. (2013).

2.2.3 Parameters estimation methods

In literature, many methods are used to estimate the parameters of time series models. We present those which are mainly used, namely, the least-square and maximum likelihood methods.

Least-squares methods

From Marquardt (1963), most algorithms for the least-squares estimation of nonlinear parameters have centered on two approaches. On the one hand, the model may be expanded as a Taylor series and corrections to the several parameters calculated at each iteration on the assumption of local linearity. On the other hand, various modifications of the method of steepest-descent have been used. Both methods frequently ground to a stop, the Taylor series method because of divergence of the successive iterates, the steepest-descent (or gradient) methods because of agonizingly slow convergence after the first few iterations. The least-squares method can be explained as follow:

Let x_1, \dots, x_m be independent variables, β_1, \dots, β_k are the population values of k parameters. Let the model to be fitted to the data be

$$\begin{aligned}\mathbb{E}(y) &= f(x_1, \dots, x_m; \beta_1, \dots, \beta_k) \\ &= f(\mathbf{X}, \beta)\end{aligned}\tag{2.2}$$

where $\mathbb{E}(y)$ is the expected value of the dependent variable y . Let the data points be denoted by

$$(Y_i, X_{1i}, \dots, X_{mi}), i = 1, \dots, n.$$

The problem is to compute those estimates of the parameters which will minimize

$$\begin{aligned}\Phi &= \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 \\ &= \|Y - \hat{Y}\|^2\end{aligned}$$

where \widehat{Y}_i is the value of y predicted by (2.2) at the i th data point.

Maximum likelihood method

In statistics, maximum likelihood estimation (MLE) is a method for estimating the parameters of a given probability distribution on the basis of observed data. This is achieved by maximizing a likelihood function so that, under the assumed statistical model, the observed data are the most probable. The point in the parameter space that maximizes the likelihood function is called the maximum likelihood estimate. The logic of maximum likelihood is both intuitive and flexible, and as such, the method has become a dominant means of statistical inference.

We write the parameters governing the joint distribution as a vector $\theta = (\theta_1, \theta_2, \dots, \theta_k)^\top$ so that this distribution falls within a parametric family $\{f(\cdot; \theta) \mid \theta \in \Theta\}$, where Θ is called the parameter space, a finite-dimensional subset of Euclidean space. Evaluating the joint density at the observed data sample $\mathbf{x} = (x_1, x_2, \dots, x_n)$ gives a real-valued function,

$$\mathcal{L}_n(\theta, \mathbf{x}) = f_n(\theta, \mathbf{x}),$$

which is called the likelihood function. If the random variables are independent and identically distributed, $f_n(\theta, \mathbf{x})$ will be the product of univariate density functions,

$$f_n(\theta, \mathbf{x}) = \prod_{k=1}^n f_k^{\text{univar}}(\theta, x_k),$$

where f_k^{univar} stands for the univariate density function of x_k . Now, we can estimate θ by

$$\widehat{\theta} = \arg \max_{\theta \in \Theta} [\mathcal{L}_n(\theta, \mathbf{x})].$$

By using the logarithm function \log , the function will be called the log-likelihood function expressed as follow

$$L_n(\theta, \mathbf{X}) = \log(\mathcal{L}_n(\theta, \mathbf{x})).$$

Since the logarithm function is a monotonic function, the maximum of $L_n(\theta, \mathbf{x})$ occurs at the same value of θ as does the maximum of $\mathcal{L}_n(\theta, \mathbf{x})$.

If $L_n(\theta, \mathbf{x})$ is differentiable in Θ , the necessary conditions for the occurrence of a maximum, or a minimum, are

$$\frac{\partial L_n}{\partial \theta_1} = 0, \frac{\partial L_n}{\partial \theta_2} = 0, \dots, \frac{\partial L_n}{\partial \theta_k} = 0,$$

known as the likelihood equations. Sometimes, these equations are sufficient to find an estimator of θ . In general, we use a second order partial derivatives, and we obtain a matrix called by Hessian matrix

$$H(\hat{\theta}) = \begin{pmatrix} \frac{\partial^2 L_n}{\partial \theta_1^2} \Big|_{\theta=\hat{\theta}} & \frac{\partial^2 L_n}{\partial \theta_1 \partial \theta_2} \Big|_{\theta=\hat{\theta}} & \cdots & \frac{\partial^2 L_n}{\partial \theta_1 \partial \theta_k} \Big|_{\theta=\hat{\theta}} \\ \frac{\partial^2 L_n}{\partial \theta_2 \partial \theta_1} \Big|_{\theta=\hat{\theta}} & \frac{\partial^2 L_n}{\partial \theta_2^2} \Big|_{\theta=\hat{\theta}} & \cdots & \frac{\partial^2 L_n}{\partial \theta_2 \partial \theta_k} \Big|_{\theta=\hat{\theta}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 L_n}{\partial \theta_k \partial \theta_1} \Big|_{\theta=\hat{\theta}} & \frac{\partial^2 L_n}{\partial \theta_k \partial \theta_2} \Big|_{\theta=\hat{\theta}} & \cdots & \frac{\partial^2 L_n}{\partial \theta_k^2} \Big|_{\theta=\hat{\theta}} \end{pmatrix},$$

which should be negative semi-definite at $\hat{\theta}$, as this indicates local concavity.

Conveniently, most common probability distributions - in particular the exponential family - are logarithmically concave.

2.2.4 Tests for breaks detection

Most of the break or segmentation tests are based on likelihood or on CUSUM tests.

Likelihood ratio test

The likelihood-ratio test assesses the goodness of fit of two competing statistical models based on the ratio of their likelihoods. The likelihood ratio test is a test for hypothesis in which two different maximum likelihood estimates of a parameter are compared to decide whether to reject or not a restriction on the parameter. Denote the parameter of interest by $\theta = (\theta_1, \dots, \theta_k)$ for some k . A change occurs if θ changes from parameter space Γ_0 to another parameter space Γ_1 such that $\Gamma = \Gamma_1 \cup \Gamma_0$, $\Gamma_1 \neq \Gamma_0$ and $\Gamma_1 \cap \Gamma = \emptyset$.

Consider these two hypotheses:

$$\begin{cases} H_0 : \theta \in \Gamma_0 \\ H_1 : \theta \in \Gamma_1 \end{cases}$$

The likelihood function $L(\theta|x) = f(x|\theta)$ (where f is the probability function) viewed as a function of the parameter θ with x held fixed at the value that was actually observed.

The likelihood ratio statistic denoted by λ , is defined as:

$$\lambda(x) = \frac{\max_{\theta \in \Gamma_1} L(\theta|x)}{\max_{\theta \in \Gamma_0} L(\theta|x)}$$

If λ is larger than a critical value at a level of significance α , the null hypothesis is rejected, otherwise, it is accepted. For more details, we can see Monfort (1982) or Lehmann et al. (2005).

Optimality

Definition 2.2.5. *We say that a test T is optimal, or the most powerful test, with level of significance $\alpha \in]0, 1[$, if it is the most powerful test among any other tests at the same level of significance.*

Definition 2.2.6. *(Neyman-Pearson lemma)*

The likelihood ratio test is the most powerful test for any level of significant $\alpha \in]0, 1[$.

CUSUM test

One of the statistics most often used for the segmentation of the time series is the cumulative sum or the **CUSUM** (introduced by Page (1954)). To explain it, we consider x_1, \dots, x_T as a time series process with m change-points. The test statistic is defined as:

$$IT = \sqrt{\frac{T}{2}} \max_k D_k,$$

where

$$D_k = \frac{\sum_{t=1}^k x_t^2}{\sum_{t=1}^T x_t^2} - \frac{k}{T},$$

with $0 < k < T$. The asymptotic distribution of the statistic IT is the maximum of a Brownian bridge ($B(s)$):

$$IT \longrightarrow_{D[0,1]} \max\{B(s) : s \in [0, 1]\},$$

where B is a Brownian bridge, $D[0, 1]$ is the Skorohod space. This establishes a Kolmogorov-Smirnov type asymptotic distribution. The null hypothesis is rejected when the maximum value of the function IT is greater than the critical value and the change-point is located at a period $k = \hat{k}$ where the maximum is achieved:

$$\hat{k} = \{k : IT > c.v\},$$

where $c.v$ is the corresponding critical value.

Brown et al. (1975) introduce another version of CUSUM test based on the least-squares residuals, denoted by $\text{CUSUM}^{\text{ols}}$. Zeileis (2001), Zeileis (2004) use the CUSUM test

in order to estimate the p -value in a class of the following standard linear regression models

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

where y_t designates a dependent variable observation, $x_t = (1, x_{t2}, \dots, x_{tk})^\top$ is a $k \times 1$, $(u_t)_t$ is a sequence of iid centered random variables with variance σ^2 and β is a $k \times 1$ regression coefficient vector.

The CUSUM^{ols} test is defined using the least-squares residual $\hat{u}_t = y_t - x_t^\top \hat{\beta}$.

The statistic of CUSUM test based on the residuals cumulative sum can be expressed as follow

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

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

$\hat{\beta}^{(i-1)}$ is the ordinary least squares estimation of the regression coefficients $\beta^{(i-1)}$ based on the observations up to $i-1$, k is the number of component of the vector x_t and $t \in [0, 1]$ is the break's instant.

For $c(t) = \lambda(1+2t)$, $\lambda \in]0, 1[$, H_0 will be rejected when $w_n(t)$ tends to $c(t)$ or $-c(t)$. That means, H_0 will be rejected when the test statistic $S > \lambda$, where

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

Krämer et al. (1988) proved that asymptotically,

$$w_n(t) \xrightarrow{\mathcal{D}} B(t),$$

where $B(t)$ is a standard Brownian motion.

For $t \in [0, 1]$, the trajectory of CUSUM^{ols} test can be defined as follows

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

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

Krämer and Schotman (1992) proved that, asymptotically,

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

where $B^0(t)$ is a standard Brownian bridge.

H_0 is rejected when the test statistic $S^0 > \lambda$, $\lambda \in]0, 1[$, where

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

For $\alpha \in]0, 1[$ designate the significant threshold, p -value is defined as follows

$$p\text{-value} = \mathbb{P}_{H_0}(S > \alpha).$$

That means, the p -value of a standard CUSUM test is

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

The asymptotic p -value considered in Zeileis (2004) and Zeileis (2001) is

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

and the asymptotic p -value of CUSUM^{ols} is

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

Zeileis (2001) and Zeileis (2004) bounded the p -value with two approximate bounds for the standard CUSUM test and CUSUM^{ols} test.

One of the major drawbacks is that they have low power, when the changes occur at the beginning or when they occur at the end (late changes). For this reason, they update the bounds, see Zeileis (2004) and Zeileis (2001). Later, many version of CUSUM test was introduced, see Horváth et al. (2020).

We can see a detailed discussion for all the methods based on CUSUM test in Ltaifa (2021).

2.3 Some important notions

This section is devoted to the statement of the mathematical tools used. The main notions are identifiability, stationarity, ergodicity, LAN and central limit theorem.

2.3.1 Identifiability

To simplify the work and to make this thesis more clear, we first define the notation.

In all what follows, d, k, p, q, m, n are non negative integers and Θ is a non-empty open subset of \mathbb{R}^m . We observe the random variables X_1, \dots, X_n in a measurable space $(\mathfrak{X}, \mathfrak{A}, \mathbf{P}_\theta)$, where $\theta \in \Theta$.

Definition 2.3.1. *The parametric family $\{\mathbf{P}_\theta, \theta \in \Theta\}$ admits the injective property (identifiability), if for θ and $\theta' \in \Theta$, we have:*

$$\theta \neq \theta' \Rightarrow \mathbf{P}_\theta \neq \mathbf{P}_{\theta'}.$$

2.3.2 Stationarity

The notion of stationarity is one of the most important in time series analysis and their theoretical studies.

Definition 2.3.2. *A stochastic process (or a time series) $(X_t)_{t \in \mathbb{Z}}$ is said to be strongly or strictly stationary if the joint distribution of $(X_{t_1}, \dots, X_{t_k})$ is identical to that of $(X_{t_1+t}, \dots, X_{t_k+t})$ for all t , where $k \in \mathbb{N}$ and (t_1, \dots, t_k) is a collection of k positive integers.*

In real-life, often of the theoretical study for a real-time series doesn't have the stationarity property. Sometimes, they can admit a weak version for the stationarity, known as weak or wide-sense stationarity. In a mathematical view, a stochastic process is said to be weakly stationary if its mean is independent of the time and its covariance depends on the lag between two observations.

Definition 2.3.3. *A stochastic process $(X_t)_{t \in \mathbb{Z}}$ is said to be weakly stationary if:*

- $\mathbb{E}(X_t) = \mu$, constant and independent of t ,
- $\text{Cov}(X_t, X_{t-d})$ depends only on d .
- $\mathbb{E}(X_t^2) < \infty$.

Non-stationarity

Most economic series are non-stationary, i.e. the process which describes them does not satisfy at least one of the conditions of the definition of a second-order stationary process. This led to define several types of non-stationarity.

- **Deterministic non-stationarity.**

Definition 2.3.4. *We say that the process $(X_t)_t$ is a deterministic non-stationary process, if we can express it as*

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

where f is a time dependent deterministic function and $\{Z_t\}$ is a stationary process. There are particular cases where $\{Z_t\}$ is a white noise.

For example, $f(t) = a + bt$, $(a, b) \in \mathbb{R}^2$, it is the commonly used case. Also, we can consider a quadratic trend $f(t) = a + bt + ct^2$, $(a, b, c) \in \mathbb{R}^3$. Lavielle and Lebarbier (2001), Lebarbier (2005) and Ciuperca (2011) considered this type of non-stationarity.

- **Piece-wise stationarity**

In our work, we are interested in another type of non-stationarity, called piece-wise. It is different from that introduced by Dahlhaus (1997). A process $(X_t)_t$ is said to be piece-wise stationary if over its observation interval it is stationary on sub-intervals. Note that such a process is non-stationary, if it has different laws on at least two sub-intervals of its observation interval. Dahlhaus (1997) is a generalization of Whittle (1953). When the number of observation tends to $+\infty$, Dahlhaus (1997) consider a process that is asymptotically stationary and he defines the non-stationary process as a process having a spectral representation.

Definition 2.3.5. *Non-stationarity in the sense of Dahlhaus (1997).*

We say that $(X_{t,T})_{t \in \mathbb{Z}}$ is non-stationary in the sense of Dahlhaus, if $(X_{t,T})_{t \in \mathbb{Z}}$ has the following representation

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

where $X_{t,T}$ ($t = 1, \dots, T$) represents the series of observations, T is the number of observations, ξ represents a stochastic process on $[-\pi, \pi]$ and $A_{t,T}^0(\lambda) = A(t/T, \lambda)$ such that $A : [0, 1] \times \mathbb{R} \rightarrow \mathbb{C}$ is a 2π -periodic function with $A(u, -\lambda) = \overline{A(u, \lambda)}$ and for any T , $A(u, \lambda)$ and $\mu(u)$ are continuous in u , where \bar{x} represents the conjugate of $x \in \mathbb{C}$.

Remark 2.3.1. *The above non-stationarity can be seen as a local non-stationarity. Ngatchou-Wandji and Harel (2013) and Ngatchou-Wandji et al. (2019) used another type of local non-stationary process.*

2.3.3 Ergodicity

There are two main versions for the law of large numbers: weak law of large numbers and strong law of large numbers.

Definition 2.3.6. (*Weak law of large numbers*).

Let (X_n) be a sequence of integrable variables, independent two by two, and identically distributed. Let m be their common expectation ($\mathbb{E}(X_1) = m$). Let

$$S_n = \frac{X_1 + \cdots + X_n}{n}.$$

Then, when $n \rightarrow +\infty$, the sequence $S_n \xrightarrow{P} m$.

Definition 2.3.7. (*Strong law of large numbers*).

Let (X_n) be a sequence of mutually independent and identically distributed integrable variables with finite mean m . Let

$$S_n = \frac{X_1 + \cdots + X_n}{n}.$$

Then, the sequence $S_n \xrightarrow{a.s.} m$.

The ergodicity is a generalisation of the law of large numbers, it can be defined as:

Definition 2.3.8. A stationary process $(X_t)_{t \in \mathbb{Z}}$ is said to be ergodic if for any Borel function $f : \mathbb{R}^p \rightarrow \mathbb{R}$, we have

$$\frac{1}{n} \sum_{i=1}^n f(X_{t_1+i}, X_{t_2+i}, \dots, X_{t_p+i}) \rightarrow \mathbb{E}[f(X_{t_1}, \dots, X_{t_p})]$$

Remark 2.3.2. The class of CHARN models (1.1) is very large. It contains models such as ARMA, ARCH, GARCH, ExpAR and GExpAR models. The invertibility of these models is readily obtained for example for $|V(x)| > 0$. Ngatchou-Wandji (2005) discusses a sufficient condition for strict stationarity by checking the conditions (S1)-(S4) of p. 86 in Taniguchi and Kakizawa (2012). Chen and An (1998) studied the case of GARCH which generalizes ARCH models. Next, a sufficient condition for geometry ergodicity can be obtained by applying a result of Tjøstheim (1990). An et al. (1997) investigate the stationarity property for a particular class of ARCH models nested in model (1.1).

Finally, it is possible that from the theory of Markov chains, other interesting conditions for stationarity and ergodicity be obtained for many models within model (1.1). Ngatchou-Wandji (2008) discuss the stationarity property in a class of CHARN models. For the stationarity and ergodicity of an ExpAR models, in fact, the initial idea of Ozaki and Oda (1977) is to keep the model stationary. Ozaki (1980) and Ozaki (1982) gave some sufficient conditions for the ExpAR model. Further, Chen et al. (2018) discussed the stationarity and ergodicity of an ExpAR models and gave less restrictive conditions. They proved, under some conditions, the stationarity of GExpAR models.

2.3.4 Local asymptotic normality

The random variables X_1, \dots, X_n belongs to the measurable space $(\mathfrak{X}, \mathfrak{A}, \mathbf{P}_\theta)$. Let $\mathbf{X}^{(n)} = (X_1, \dots, X_n)$, then the n -uple $\mathbf{X}^{(n)}$ belong to the space $(\mathfrak{X}^n, \mathfrak{A}^n)$ with a probability law $\mathbf{P}_\theta^{(n)}$, where $\mathbf{P}_\theta^{(n)}$ is the product of n copies of \mathbf{P}_θ .

The experience $X^{(n)} = (X_1, \dots, X_n)$ represents the observation that we have and it can be expressed by the statistical experience $\zeta^{(n)}(\theta) = (\mathfrak{X}^n, \mathfrak{A}^n, \mathcal{P}_\theta^{(n)})$, where $\mathcal{P}_\theta = \{\mathbf{P}_\theta, \theta \in \Theta\}$ and $\mathcal{P}^{(n)}$ is the product of n copies of \mathbf{P}_θ .

In the case $\mathfrak{X} = \mathbb{R}$, we have $\mathfrak{A} = \mathcal{B}$, where \mathcal{B} is the Borelian field.

For every $\theta \in \Theta$, we define the local sequence $\theta^{(n)}$ in a neighborhood of θ as

$$\theta^{(n)} = \theta + \vartheta(n)\tau,$$

where $\vartheta(n)$ is a $k \times k$ matrix and $\tau \in \mathbb{R}^k$. We assume that this sequence $\theta^{(n)}$ converges to θ as $n \rightarrow +\infty$.

We also assume that when n is large enough, $\theta^{(n)} \in \Theta$. The matrix $\vartheta(n)$ is not singular with $\|\vartheta(n)\| \rightarrow 0$ when $n \rightarrow +\infty$, where $\|\cdot\|$ is the Euclidean norm. In general, for $\tau \in \mathbb{R}^k$, we write:

$$\theta^{(n)} = \theta + \vartheta(n)\tau^{(n)}, \text{ where } \sup_n [(\tau^{(n)})^\top \tau^{(n)}] < +\infty.$$

The choice of the sequence $\vartheta(n)$ depends on the problem to be studied. Many examples are cited in Dreesbeke and Fine (1996).

Consider P and Q two probability measures in the measurable space $(\mathfrak{X}, \mathfrak{A})$. There exist a non-negative function f \mathfrak{A} -measurable and an event N where $P(N) = 0$, such that for any $A \in \mathfrak{A}$, we have:

$$Q(A) = \int_A f dP + Q(A \cap N).$$

This equality is "Lesbegue decomposition of Q with respect to P ", and $Q(A \cap N)$ is the singularity part of this decomposition.

In the case where Q is absolutely continuous with respect to P , the density f is known as the likelihood ratio of Q with respect to P , and define by dQ/dP .

In the case the concern the statistical experience $\zeta^{(n)}(\theta)$, we consider the likelihood ratio $\frac{d\mathbb{P}_{\theta+\vartheta(n)\tau^{(n)}}^{(n)}}{d\mathbb{P}_\theta^{(n)}}$, this choice is motivated by Neyman-Pearson lemma which shows the interest of this ratio in the test statistics problems, by studying the logarithm of the likelihood ratio and by setting

$$\Lambda_{\theta+\vartheta(n)\tau^{(n)}} = \log \left(\frac{d\mathbb{P}_{\theta+\vartheta(n)\tau^{(n)}}^{(n)}}{d\mathbb{P}_\theta^{(n)}} \right).$$

There are many versions of the LAN property. We recall that of Dreesbeke and Fine (1996).

Definition 2.3.9. *The sequence $\zeta^{(n)}(\theta) = (\mathbb{R}^n, \mathcal{B}^n, \mathbf{P}^{(n)}(\theta))$ is locally asymptotic normal LAN, if for any $\theta \in \Theta$, exist a sequence $\Pi_n^{(n)}(\theta)$ of random vectors \mathcal{B}^n -measurable, and a positive semi-definite matrix $\Gamma(\theta)$ continuous in terms of θ , such that:*

- i. For every sequence $\tau^{(n)}$ such that $\sup[(\tau^{(n)})'\tau^{(n)}] < +\infty$, and under $\mathbf{P}^{(n)}(\theta)$, as $n \rightarrow +\infty$, we have :*

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

- ii. Under $\mathbf{P}^{(n)}(\theta)$ and when $n \rightarrow +\infty$, $\Pi_n^{(n)}(\theta) \xrightarrow{D} \mathcal{N}(0, \Gamma(\theta))$. $\Pi_n^{(n)}(\theta)$ is called central sequence.*

Many families provides the LAN property. Dreesbeke and Fine (1996) give five examples of model having a LAN property (under the null hypothesis) under classical regularity conditions and Noether conditions in Hájek (1961).

2.3.5 Contiguity

Let $P^{(n)}$ and $Q^{(n)}$ be two probability measures in the measurable space $(\mathfrak{X}^{(n)}, \mathfrak{A}^{(n)})$. From Dreesbeke and Fine (1996), we have:

Definition 2.3.10. *The sequences $Q^{(n)}$ is contiguous to the sequence $P^{(n)}$ if and only if for every $A^{(n)} \in \mathfrak{A}^{(n)}$ we have:*

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

Definition 2.3.11. *The sequence $Q^{(n)}$ and $P^{(n)}$ are mutually contiguous if and only if for every $A^{(n)} \in \mathfrak{A}^{(n)}$ we have:*

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

If two measures are mutually contiguous, then they are contiguous. The contiguity is a type of asymptotic absolute form continuity.

Remark 2.3.3. *In case that we have the convergence in probability for a sequence of random variables with respect to the measure $P^{(n)}$, we can conclude the convergence in probability of this sequence with respect to another measure $Q^{(n)}$ contiguous to $P^{(n)}$.*

It is not always easy to establish the contiguity criteria basing on the definition. *Corollary (4.2)* of Dreesbeke and Fine (1996) gives the contiguity criterion strongly linked to the LAN property.

Corollary 2.3.1. (*Corollary (4.2) of Dreesbeke and Fine (1996)*)

The sequence $\Lambda^{(n)} = \log \left(\frac{dQ^{(n)}}{dP^{(n)}} \right)$ is asymptotically normal with respect to the measure $P^{(n)}$. That is, there exist $\mu^{(n)}$ and $\sigma^{(n)}$ such that, under $P^{(n)}$ as $n \rightarrow +\infty$, $\frac{\Lambda^{(n)} - \mu^{(n)}}{\sigma^{(n)}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$ (*Remark (4.2) of Dreesbeke and Fine (1996)*), $P^{(n)}$ and $Q^{(n)}$ are mutually contiguous if and only if:

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

2.3.6 Le Cam's Third Lemma

Recall $P^{(n)}$ and $Q^{(n)}$ as two probability measures in the measurable space $(\mathfrak{X}^{(n)}, \mathfrak{A}^{(n)})$.

Proposition 2.3.1. (*Proposition (4.2) of Dreesbeke and Fine (1996)*)

Let $S^{(n)}$ be a sequence of random variables and let $\Lambda^{(n)} = \log \left(\frac{dQ^{(n)}}{dP^{(n)}} \right)$.

If under $P^{(n)}$, $\begin{pmatrix} S^{(n)} \\ \Lambda^{(n)} \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N} \left[\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \right]$, where $\mu_2 = -\frac{\sigma_2^2}{2}$, then:

$P^{(n)}$ and $Q^{(n)}$ are mutually contiguous, and under $Q^{(n)}$, $S^{(n)} \xrightarrow{\mathcal{D}} \mathcal{N}(\mu_1 + \sigma_{12}, \sigma_1^2)$.

Le Cam's third lemma is rich in practical consequences. For example, it may allows for the computation of the theoretical asymptotical power of tests.

Local asymptotical power

We consider the sequence of tests φ^n define as:

$$\varphi^n = I \left\{ \frac{S^{(n)} - \mu^{(n)}}{\sigma_1^{(n)}} > Z_\alpha \right\}$$

where Z_α is the $(1 - \alpha)$ -quantile of a standard normal random variable. If under the measure $P^{(n)}$ we have:

$$\begin{pmatrix} \frac{S^{(n)} - \mu^{(n)}}{\sigma^{(n)}} \\ \Lambda^{(n)} \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N} \left[\begin{pmatrix} 0 \\ -\frac{1}{2}\sigma_2^2 \end{pmatrix}, \begin{pmatrix} 1 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \right],$$

then, under the null hypothesis $P^{(n)}$, the asymptotic level of φ^n is α and we have:

$$\lim_{n \rightarrow +\infty} \mathbb{E}_{P^{(n)}}(\varphi^n) = \alpha$$

As we said before, Le Cam's third lemma helps with computing the asymptotic power of the sequence of tests φ^n under the sequence $Q^{(n)}$ and we obtain:

$$\lim_{n \rightarrow +\infty} \mathbb{E}_{Q^{(n)}}(\varphi^n) = 1 - \Phi(Z_\alpha - \sigma_{12}), \quad (2.3)$$

where Φ is the cumulative distribution function of a standard normal random variable.

2.3.7 Central limit theorem of martingales

Let, for $i = 1, \dots, n$, $\mathcal{F}_i = \sigma(X_1, \dots, X_i)$ be the σ -algebra generated by X_1, \dots, X_i .

Theorem 1. (Theorem (3.2) of Hall and Heyde (2014))

Let $\{S_{n,i}, \mathcal{F}_{n,i}, 1 \leq i \leq k_n, n \geq 1\}$ be a centered and square integrable martingale, and $X_{n,i} = S_{n,i} - S_{n,i-1}$. Let η be a random variable a.s. finite.

Assume that

$$\max_i (|X_{n,i}|) \xrightarrow{P} 0, \quad (2.4)$$

$$\sum_{i=2}^n X_{n,i}^2 \xrightarrow{P} \eta^2, \quad (2.5)$$

$$\mathbb{E}(\max_i (X_{n,i}^2)) \text{ is bounded in } n, \quad (2.6)$$

and these σ -algebras are ordered, (i, e)

$$\mathcal{F}_{n,i} \subseteq \mathcal{F}_{n+1,i} \text{ for } 1 \leq i \leq k_n, n \geq 1. \quad (2.7)$$

Then, $S_{n,k_n} = \sum_{i=1}^{k_n} X_{n,i} \xrightarrow{D} Z$, where the random variable Z admits as a characteristic function $\mathbb{E}\left[\exp -\frac{\eta^2}{2}t^2\right]$.

Corollary 2.3.2. (Corollary (3.1) of Hall and Heyde (2014))

We keep in the statement of the theorem (1) the condition (2.7). If, instead of the condition (2.4) and (2.6), we replace them by the Linderberg condition, and if the condition (2.5) is replaced by a condition with the same type of the conditional variance:

$$V_{nk_n}^2 = \sum_{i=1}^{k_n} \mathbb{E}(X_{n,i}^2 | \mathcal{F}_{n,i-1}) \xrightarrow{\mathbb{P}} \eta^2, \quad (2.8)$$

the previous theorem remains verified.

Chapter 3

Testing the conditional mean of a time-dependent coefficients CHARN model: Part I

3.1 Introduction

In this chapter, we consider the testing problem presented in Chapter 1. We recall it to be self contained.

We consider n observations, X_1, \dots, X_n , generated by the following CHARN model

$$X_t = T(\boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(t); \mathbf{X}_{t-1}) + V(\mathbf{X}_{t-1})\varepsilon_t, t \in \mathbb{Z}$$

where $(X_t)_{t \in \mathbb{Z}}$ is a piece-wise stationary and ergodic process in each segment $[\tau_{j-1}, \tau_j[$, where the τ_j , $j = 0, \dots, k+1$, are potential instants of change with $\tau_0 = 1$ and $\tau_{k+1} = n$ and n denotes the sample size; $(\varepsilon_t)_{t \in \mathbb{Z}}$ is a standard white noise with density f ; $\mathbf{X}_t = (X_t, \dots, X_{t-d+1})^\top$; T and $V > 0$ are real smooth functions; $\boldsymbol{\rho}_0^\top \in \mathbb{R}^p$, $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_{k+1}^\top)^\top$, $\boldsymbol{\gamma}_j \in \mathbb{R}^p$ for any $j = 1, \dots, k+1$; $\boldsymbol{\omega}(t) = (\mathbf{1}_{[\tau_0, \tau_1)}(t), \mathbf{1}_{[\tau_1, \tau_2)}(t), \dots, \mathbf{1}_{[\tau_{k-1}, \tau_k)}(t), \mathbf{1}_{[\tau_k, \tau_{k+1})}(t))^\top = (\omega_1, \dots, \omega_{k+1}) \in \{0, 1\}^{k+1}$; for any $\boldsymbol{\gamma} = (\boldsymbol{\gamma}_1^\top, \dots, \boldsymbol{\gamma}_{k+1}^\top)^\top \in \mathbb{R}^{p(k+1)}$ and $\boldsymbol{\omega} = (\omega_1, \dots, \omega_{k+1})^\top \in \mathbb{R}^{k+1}$,

$$\boldsymbol{\gamma} \odot \boldsymbol{\omega} = \boldsymbol{\gamma}_1 \omega_1 + \dots + \boldsymbol{\gamma}_{k+1} \omega_{k+1} \in \mathbb{R}^p,$$

where for $i = 1, \dots, k+1$,

$$\boldsymbol{\gamma}_i \omega_i = (\gamma_{i,1} \omega_i, \dots, \gamma_{i,p} \omega_i) \in \mathbb{R}^p.$$

We construct a log-likelihood ratio test for testing

$$H_0 : \gamma = \gamma_0 \quad \text{against} \quad H_{\beta}^{(n)} : \gamma = \gamma_n = \gamma_0 + \frac{\beta}{\sqrt{n}}, \quad (3.1)$$

for some $\gamma_0 = (\gamma_{0,1}^\top, \dots, \gamma_{0,k+1}^\top)^\top$, $\beta = (\beta_1^\top, \dots, \beta_{k+1}^\top)^\top \in \mathbb{R}^{p(k+1)}$, with $\gamma_{0,i}, \beta_i \in \mathbb{R}^{p(k+1)}$, $i = 1, \dots, k+1$. The class of models (1.1) is very large. It contains models such as ARMA, ARCH, GARCH, ExpAR, GExpAR, whose statistical properties can be found in Guégan et al. (1994) and Tong (1990). This chapter is devoted to the study when all the parameters of the model are assumed to be known. As noted in Chapter 1, this work is an extension of that of Ltaifa (2021) and Ngatchou-Wandji and Ltaifa (2021). The latters present a new approach to change-points study, which is particularly adapted to weak changes, an important issue which is not very much studied in the literature.

In the remainder of this chapter, we define the notation and list some technical assumptions. We define our likelihood-ratio test and study its properties under the null and under the sequence of local alternatives. A LAN result is established as well as a contiguity result. An expression of the theoretical power is given.

3.2 Notation and technical assumptions

3.2.1 Notation

We denote by $\mathcal{M}_{m,n}(\mathbb{R})$ the space of real $m \times n$ matrices and $\mathcal{M}_n(\mathbb{R}) = \mathcal{M}_{n,n}(\mathbb{R})$. For $M \in \mathcal{M}_{m,n}(\mathbb{R})$, we denote by M^\top its transpose and by $\|M\|_{m \times n}$ its Euclidean matrix norm.

For $U \in \mathbb{R}^{p(k+1)}$, we write U as follows

$$U = (U_1^\top, \dots, U_{k+1}^\top)^\top$$

and for any $i \in \{1, \dots, k+1\}$,

$$U_i = (U_{i1}, \dots, U_{ip})^\top.$$

For $M \in \mathcal{M}_{p(k+1)}(\mathbb{R})$, we write M as follows

$$M = \begin{pmatrix} M_{1,1} & \dots & M_{1,k+1} \\ \vdots & \ddots & \vdots \\ M_{k+1,1} & \dots & M_{k+1,k+1} \end{pmatrix},$$

and for every $i, j \in \{1, \dots, k+1\}$, $M_{i,j} \in \mathcal{M}_p(\mathbb{R})$. M is a block matrix.

Let $\mathfrak{S} : \mathbb{R}^{p(k+1)} \rightarrow \mathbb{R}$ be a differentiable function on $\mathbb{R}^{p(k+1)}$. For any $\gamma \in \mathbb{R}^{p(k+1)}$, we denote by $D_\gamma[\mathfrak{S}]$ the matrix function defined, for any $x \in \mathbb{R}^{p(k+1)}$, by

$$D_\gamma[\mathfrak{S}(x)] = \left(\partial_{\gamma_1}[\mathfrak{S}(x)], \dots, \partial_{\gamma_{k+1}}[\mathfrak{S}(x)] \right)^\top \in \mathbb{R}^{p(k+1)},$$

with $\partial_{\gamma_i}[\mathfrak{S}(x)]$ standing for the gradient of \mathfrak{S} with respect to γ_i defined as

$$\partial_{\gamma_i}[\mathfrak{S}(x)] = \left(\frac{\partial \mathfrak{S}}{\partial \gamma_{i,1}}(x), \frac{\partial \mathfrak{S}}{\partial \gamma_{i,2}}(x), \dots, \frac{\partial \mathfrak{S}}{\partial \gamma_{i,p}}(x) \right)^\top \in \mathbb{R}^p,$$

where, for any $i = 1, \dots, k+1$ and $j = 1, \dots, p$, $\frac{\partial \mathfrak{S}}{\partial \gamma_{i,j}}$ stands for the partial derivatives of \mathfrak{S} with respect to $\gamma_{i,j}$.

We also denote

$$H_\gamma[\mathfrak{S}(x)] = \begin{pmatrix} \partial_{\gamma_1 \gamma_1}^2 \mathfrak{S}(x) & \dots & \partial_{\gamma_1 \gamma_{k+1}}^2 \mathfrak{S}(x) \\ \vdots & \ddots & \vdots \\ \partial_{\gamma_{k+1} \gamma_1}^2 \mathfrak{S}(x) & \dots & \partial_{\gamma_{k+1} \gamma_{k+1}}^2 \mathfrak{S}(x) \end{pmatrix} \in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

where, for any $i \in \{1, \dots, k+1\}$,

$$\partial_{\gamma_i \gamma_j}^2 \mathfrak{S}(x) = \begin{pmatrix} \frac{\partial^2 \mathfrak{S}}{\partial \gamma_{i,1} \partial \gamma_{j,1}}(x) & \dots & \frac{\partial^2 \mathfrak{S}}{\partial \gamma_{i,p} \partial \gamma_{j,1}}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 \mathfrak{S}}{\partial \gamma_{i,1} \partial \gamma_{j,p}}(x) & \dots & \frac{\partial^2 \mathfrak{S}}{\partial \gamma_{i,p} \partial \gamma_{j,p}}(x) \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}).$$

For any differentiable function g with derivative g' , we denote by

$$\phi_g = -\frac{g'}{g} \text{ and } I(g) = \int_{\mathbb{R}} \phi_g^2(x) g(x) dx.$$

For any $t = 1, \dots, n$, let $\mathcal{F}_t = \sigma(\mathbf{X}_1, \dots, \mathbf{X}_t)$ be the σ -algebra generated by $\mathbf{X}_1, \dots, \mathbf{X}_t$ such that ε_t is independent of \mathcal{F}_{t-1} . Let $\mathbf{X}^{(n)} = (X_n, \dots, X_1)$, and let $\ell(X_j)$ be the law of X_j . From Million (2007), we define the Hadamard product as follows

Definition 3.2.1. For any vectors β and $\tilde{\beta}$,

$$\beta = \left((\beta_{1,1}, \dots, \beta_{1,p})^\top, \dots, (\beta_{k+1,1}, \dots, \beta_{k+1,p})^\top \right)^\top \in \mathbb{R}^{p(k+1)}$$

$$\tilde{\beta} = \left((\tilde{\beta}_{1,1}, \dots, \tilde{\beta}_{1,p})^\top, \dots, (\tilde{\beta}_{k+1,1}, \dots, \tilde{\beta}_{k+1,p})^\top \right)^\top \in \mathbb{R}^{p(k+1)},$$

the Hadamard product of β and $\tilde{\beta}$ is:

$$\beta \circ \tilde{\beta} = \tilde{\beta} \circ \beta = \begin{pmatrix} \begin{pmatrix} \beta_{1,1} \\ \vdots \\ \beta_{1,p} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \beta_{k+1,1} \\ \vdots \\ \beta_{k+1,p} \end{pmatrix} \end{pmatrix} \circ \begin{pmatrix} \begin{pmatrix} \tilde{\beta}_{1,1} \\ \vdots \\ \tilde{\beta}_{1,p} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \tilde{\beta}_{k+1,1} \\ \vdots \\ \tilde{\beta}_{k+1,p} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} \beta_{1,1}\tilde{\beta}_{1,1} \\ \vdots \\ \beta_{1,p}\tilde{\beta}_{1,p} \end{pmatrix} \\ \vdots \\ \begin{pmatrix} \beta_{k+1,1}\tilde{\beta}_{k+1,1} \\ \vdots \\ \beta_{k+1,p}\tilde{\beta}_{k+1,p} \end{pmatrix} \end{pmatrix} \in \mathbb{R}^{p(k+1)}.$$

For matrices, $A \in \mathcal{M}_{p(k+1)}(\mathbb{R})$ and $B \in \mathcal{M}_{p(k+1)}(\mathbb{R})$:

$$\begin{aligned} A \circ B &= \begin{pmatrix} A_{1,1} & \dots & A_{1,k+1} \\ \vdots & \ddots & \vdots \\ A_{k+1,1} & \dots & A_{k+1,k+1} \end{pmatrix} \circ \begin{pmatrix} \beta_{1,1} & \dots & \beta_{1,k+1} \\ \vdots & \ddots & \vdots \\ \beta_{k+1,1} & \dots & \beta_{k+1,k+1} \end{pmatrix} \\ &= \begin{pmatrix} A_{1,1} \circ \beta_{1,1} & \dots & A_{1,k+1} \circ \beta_{1,k+1} \\ \vdots & \ddots & \vdots \\ A_{k+1,1} \circ \beta_{k+1,1} & \dots & A_{k+1,k+1} \circ \beta_{k+1,k+1} \end{pmatrix}, \end{aligned}$$

where for any $i, j, m, n \in \{1, \dots, \inf\{p, k+1\}\}$, $A_{i,j} \circ \beta_{m,n} \in \mathcal{M}_p(\mathbb{R})$.

3.3 General assumptions

In this section, we present our main assumptions.

(\mathcal{A}_1) For any $j = 1, \dots, k+1$, $n_j(n)$ designate the number of observations between the two instants τ_j and τ_{j-1} , such that $n_j(n) \xrightarrow{n \rightarrow +\infty} +\infty$ and $n_j(n)/n \xrightarrow{n \rightarrow +\infty} \alpha_j$.

(\mathcal{A}_2) $\int_{\mathbb{R}} xf(x)dx = 0$ and $\int_{\mathbb{R}} x^2f(x)dx = 1$.

(\mathcal{A}_3) f is differentiable with derivative f' .

(\mathcal{A}_4) The sequence of the \mathbf{X}_t 's whose components are associated with indices within $[\tau_{j-1}, \tau_j]$, $j = 1, \dots, k+1$, is stationary and ergodic with stationary cumulative distribution function F_j .

(\mathcal{A}_5) For any $j = 1, \dots, k+1$ and $1 \leq h \leq m \leq p$,

$$\eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = I(f) \int_{\mathbb{R}} \left(\frac{1}{V(x)} \right)^2 \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) dF_j(x) < \infty.$$

(\mathcal{A}_6) ϕ_f is differentiable with derivative ϕ'_f and is c_ϕ -Lipschitz where $0 < c_\phi < +\infty$.

$$(\mathcal{A}_7) \max \left\{ \int_{\mathbb{R}} |\phi_f(x)|^3 f(x) dx, \int_{\mathbb{R}} |\phi'_f(x)| f(x) dx \right\} < \infty.$$

$$(\mathcal{A}_8) \lim_{x \rightarrow +\infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0 = \lim_{x \rightarrow +\infty} f'(x) = \lim_{x \rightarrow -\infty} f'(x).$$

$$(\mathcal{A}_9) \max \left\{ \sup_{\gamma} |T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x)|, \sup_{\gamma} \|\partial_{\boldsymbol{\gamma}}[T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x)]\|_{p(k+1)}, \right. \\ \left. \sup_{\gamma} \|\|H_{\boldsymbol{\gamma}}[T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x)]\|\|_{p(k+1)} \right\} < \nu(x).$$

$$(\mathcal{A}_{10}) \text{ for } j = 1, \dots, k+1, a, b \in \{1, 2, 3\}, \int_{\mathbb{R}^d} \frac{|\nu(x)|^b}{|V(x)|^a} dF_j(x) < \infty.$$

Remark 3.3.1. [1] For $j = 1, \dots, k+1$, on each $[\tau_{j-1}, \tau_j]$, there are at most p random vectors \mathbf{X}_t whose components are associated with overlapping indices. They may have different stationary distributions. But as their number is negligible behind $n_j(n)$, their distributions do not affect the asymptotic results.

[2] Assumption (\mathcal{A}_4) holds at least for AR and some ARCH models.

[3] Assumptions (\mathcal{A}_7) and (\mathcal{A}_9) are satisfied by usual models as parametric AR, ARCH, TARARCH, ExpAR models with Gaussian noise.

3.4 The functions involved in the model and the parameters are known.

The likelihood function of $\mathbf{X}^{(n)} = (X_n, \dots, X_1)$ given $X_0 \in \mathbb{R}^d$ can be calculated by successive applications of the Bayes formula as follow:

$$\ell_{X_n, X_{n-1}, \dots, X_1 | \mathbf{X}_0 = x_0}(x_n, \dots, x_1) = \ell_{X_n | (X_{n-1}, \dots, X_1, \mathbf{X}_0) = (x_{n-1}, \dots, x_1, \mathbf{X}_0)}(x_n) \dots \ell_{X_1 | \mathbf{X}_0 = x_0}(x_1) \\ = \ell_{X_n | X_{n-1} = x_{n-1}}(x_n) \ell_{X_{n-1} | X_{n-2} = x_{n-2}}(x_{n-1}) \dots \ell_{X_1 | \mathbf{X}_0 = x_0}(x_1),$$

where $\ell_{X|\mathbf{X}}$ is the law of X knowing \mathbf{X} .

3.4.1 Likelihood ratio test

Proposition 3.4.1. The likelihood of n observation, generated by CHARN model (1.1), is given by:

$$L(\mathbf{X}^{(n)} | \mathbf{X}_0 = x_0) = \prod_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})].$$

Proof: To prove this equality, we should determine the distribution function of X_t . To simplify the expressions, let

$$\varepsilon_i(\boldsymbol{\rho}_0, \boldsymbol{\gamma}) = \frac{X_i - T(\boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(i), \mathbf{X}_{i-1})}{V(\mathbf{X}_{i-1})}, \quad (3.2)$$

$$T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) = T(\boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(t), \mathbf{X}_{t-1}). \quad (3.3)$$

We have

$$\begin{aligned} \mathbb{P}_{X_1|\mathbf{X}_0=x_0}(x_1) &= \mathbb{P}(X_1 < x_1 | \mathbf{X}_0 = x_0) = \mathbb{P}\left(\frac{X_1 - T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x_0)}{V(x_0)} < \frac{x_1 - T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x_0)}{V(x_0)}\right) \\ &= \mathbb{P}\left(\varepsilon_1 < \frac{x_1 - T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x_0)}{V(x_0)}\right) \\ &= F\left(\frac{x_1 - T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x_0)}{V(x_0)}\right), \end{aligned}$$

where F is the cumulative distribution function of ε_t .

By differentiating the right term with respect to x_1 , we obtain

$$\ell_{X_1|\mathbf{X}_0=x_0}(x_1) = \frac{1}{V(x_0)} f\left(\frac{x_1 - T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x_0)}{V(x_0)}\right), x_1 \in \mathbb{R}.$$

For any $i = 1, \dots, n$,

$$\ell_{X_i|\mathbf{X}_i=x_i}(x_i) = \frac{1}{V(x_{i-1})} f\left(\frac{x_i - T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, x_{i-1})}{V(x_{i-1})}\right), x_i \in \mathbb{R}.$$

Then

$$L(\mathbf{X}^{(n)} | \mathbf{X}_0 = x_0) = \prod_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})].$$

□

We can write the conditional log-likelihood as follow:

$$\mathcal{L}(\mathbf{X}^{(n)} | \mathbf{X}_0 = x_0) = \sum_{t=1}^n \{ \log[f(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}))] - \log[V(\mathbf{X}_{t-1})] \}.$$

Going back to our primary goal, we aim to test

$$H_0 : \boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \text{ against } H_{\boldsymbol{\beta}}^{(n)} : \boldsymbol{\gamma} = \boldsymbol{\gamma}_n = \boldsymbol{\gamma}_0 + \boldsymbol{\beta} / \sqrt{n},$$

where $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$.

The conditional log-likelihood function is

$$\mathcal{L}(\mathbf{X}^{(n)} | \mathbf{X}_0 = x_0) = \sum_{t=1}^n \{ \log[f(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0))] - \log[V(\mathbf{X}_{t-1})] \},$$

and under the alternatives hypothesis, it is

$$\mathcal{L}(\mathbf{X}^{(n)} | \mathbf{X}_0 = x_0) = \sum_{t=1}^n \{ \log[f(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_n))] - \log[V(\mathbf{X}_{t-1})] \}.$$

Then, the likelihood-ratio is given by

$$A_n = \prod_{t=1}^n \frac{f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_n)]}{f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]}. \quad (3.4)$$

Proposition 3.4.2. *Assume that (\mathcal{A}_1) – (\mathcal{A}_{10}) hold. Then, the log-likelihood ratio Θ_n of H_0 against $H_{\boldsymbol{\beta}}^{(n)}$ decomposes into*

$$\Theta_n = \Pi_n - \Delta_n + o_P(1),$$

where

- $\Delta_n = \frac{1}{2n} \sum_{t=1}^n \left\{ \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right. \\ \left. - \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\}$
- $\Pi_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\}$
- $M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) N^\top(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})$
- $N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = \omega(t) \circ D_\gamma [T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]$
- $\mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) = \omega(t) \omega(t)^\top \circ H_\gamma [T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]$.

Proof: Taking the logarithm of the likelihood ratio (3.4), we obtain:

$$\Theta_n = \log A_n = \sum_{t=1}^n \{ \log[f(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_n))] - \log[f(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0))] \}. \quad (3.5)$$

It is easy to see that we can apply a Taylor expansion to $\log \{f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]\}$ in a neighborhood of $\boldsymbol{\gamma}_0$ to obtain, for some $\tilde{\boldsymbol{\gamma}}$ lying between $\boldsymbol{\gamma}_0$ and $\boldsymbol{\gamma}_n$:

$$\log \{f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_n)]\} - \log \{f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} = (\boldsymbol{\gamma}_n - \boldsymbol{\gamma}_0)^\top (D_\gamma [\log \{f[\varepsilon_t(\boldsymbol{\gamma})]\}]) \Big|_{\boldsymbol{\gamma}=\boldsymbol{\gamma}_0} \\ + \frac{1}{2} (\boldsymbol{\gamma}_n - \boldsymbol{\gamma}_0)^\top H_\gamma [\log \{f[\varepsilon_t(\boldsymbol{\gamma})]\}] \Big|_{\boldsymbol{\gamma}=\tilde{\boldsymbol{\gamma}}} (\boldsymbol{\gamma}_n - \boldsymbol{\gamma}_0).$$

To simplify our work, we calculate all the expressions we need.

- $D_\gamma [\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] = -\frac{1}{V(\mathbf{X}_{t-1})} N(\boldsymbol{\gamma}, \mathbf{X}_{t-1}),$

- $H_\gamma[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)] = -\frac{1}{V(\mathbf{X}_{t-1})}H_\gamma[T(\boldsymbol{\rho}_0, \gamma, \mathbf{X}_{t-1})],$
- $D_\gamma[\log\{f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]\}] = \frac{1}{V(\mathbf{X}_{t-1})}N(\gamma, \mathbf{X}_{t-1})\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]$
 $= -D_\gamma[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)],$
- $H_\gamma(\log\{f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]\}) = -\frac{1}{V^2(\mathbf{X}_{t-1})}M(\gamma, \mathbf{X}_{t-1})\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]$
 $+\frac{1}{V(\mathbf{X}_{t-1})}H_\gamma[T(\boldsymbol{\rho}_0, \gamma, \mathbf{X}_{t-1})]\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]$
 $= -\frac{1}{V^2(\mathbf{X}_{t-1})}M(\gamma, \mathbf{X}_{t-1})\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]$
 $-H_\gamma[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)],$

where

- $N(\gamma, \mathbf{X}_{t-1}) = \left(\omega_1\partial_{\gamma_1}[T(\boldsymbol{\rho}_0, \gamma, \mathbf{X}_{t-1})], \dots, \omega_{k+1}\partial_{\gamma_{k+1}}[T(\boldsymbol{\rho}_0, \gamma, \mathbf{X}_{t-1})]\right)^\top \in \mathbb{R}^{p(k+1)}$
- $\omega_i \in \{0, 1\}.$

Then,

$$\begin{aligned} \log\{f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_n)]\} - \log\{f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]\} &= \frac{1}{\sqrt{n}}\frac{\boldsymbol{\beta}^\top N(\gamma_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})}\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \\ &\quad - \frac{1}{2n}\left[\frac{1}{V^2(\mathbf{X}_{t-1})}\boldsymbol{\beta}^\top M(\tilde{\gamma}, \mathbf{X}_{t-1})\right]\boldsymbol{\beta}\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] \\ &\quad - \frac{1}{2n}\left[\frac{1}{V(\mathbf{X}_{t-1})}\boldsymbol{\beta}^\top H_\gamma[T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1})]\right]\boldsymbol{\beta}\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})]. \end{aligned}$$

Now,

$$\begin{aligned} \Theta_n &= \frac{1}{\sqrt{n}}\sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})}\boldsymbol{\beta}^\top N(\gamma_0, \mathbf{X}_{t-1})\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] - \frac{1}{2n}\sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})}\boldsymbol{\beta}^\top M(\tilde{\gamma}, \mathbf{X}_{t-1})\boldsymbol{\beta} \\ &\quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] + \frac{1}{2n}\sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})}\boldsymbol{\beta}^\top \mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1})\boldsymbol{\beta}\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})], \end{aligned}$$

where

$$\begin{aligned} N(\tilde{\gamma}, \mathbf{X}_{t-1}) &= \omega(t) \circ D_\gamma[T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1})] \\ &= \begin{pmatrix} \omega_1\partial_{\gamma_1}[T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1})] \\ \vdots \\ \omega_{k+1}\partial_{\gamma_{k+1}}[T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1})] \end{pmatrix} \in \mathbb{R}^{p(k+1)}, \end{aligned}$$

$$M(\tilde{\gamma}, \mathbf{X}_{t-1}) = N(\tilde{\gamma}, \mathbf{X}_{t-1})N^\top(\tilde{\gamma}, \mathbf{X}_{t-1}) =$$

$$\begin{pmatrix} \omega_1^2 \partial_{\gamma_1} T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \partial_{\gamma_1}^\top T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \omega_1 \omega_{k+1} \partial_{\gamma_1} T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \partial_{\gamma_{k+1}}^\top T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \omega_{k+1} \omega_1 \partial_{\gamma_{k+1}} T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \partial_{\gamma_1}^\top T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \omega_{k+1}^2 \partial_{\gamma_{k+1}} T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \partial_{\gamma_{k+1}}^\top T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix}$$

$$\in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

and

$$\mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1}) = \begin{pmatrix} \omega_1^2 \partial_{\gamma_1}^2 T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \omega_1 \omega_{k+1} \partial_{\gamma_1 \gamma_{k+1}}^2 T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \omega_{k+1} \omega_1 \partial_{\gamma_{k+1} \gamma_1}^2 T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \omega_{k+1}^2 \partial_{\gamma_{k+1}}^2 T(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

with $\omega_i \in \{0, 1\}$.

Remark 3.4.1. *At time t and for any $i \neq j$, $\omega_i \omega_j = 0$. Thus*

$$M(\tilde{\gamma}, \mathbf{X}_{t-1}) = \begin{pmatrix} M_1(\tilde{\gamma}, \mathbf{X}_{t-1}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & M_2(\tilde{\gamma}, \mathbf{X}_{t-1}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & M_{k+1}(\tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

where $\mathbf{0} \in \mathcal{M}_p(\mathbb{R})$ is a nil matrix and

$$M_i(\tilde{\gamma}, \mathbf{X}_{t-1}) = \omega_i^2 \begin{pmatrix} \frac{\partial T}{\partial \gamma_{i,1}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,1}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \frac{\partial T}{\partial \gamma_{i,1}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,p}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \frac{\partial T}{\partial \gamma_{i,p}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,1}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \frac{\partial T}{\partial \gamma_{i,p}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,p}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix}$$

$$\in \mathcal{M}_p(\mathbb{R}),$$

and

$$\mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1}) = \begin{pmatrix} \mathcal{H}_1(\tilde{\gamma}, \mathbf{X}_{t-1}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathcal{H}_2(\tilde{\gamma}, \mathbf{X}_{t-1}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathcal{H}_{k+1}(\tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

where, for $i = 1, \dots, k + 1$,

$$\mathcal{H}_i(\tilde{\gamma}, \mathbf{X}_{t-1}) = \omega_i^2 \begin{pmatrix} \frac{\partial^2 T}{\partial \gamma_{i,1}^2}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \cdots & \frac{\partial^2 T}{\partial \gamma_{i,p} \partial \gamma_{i,1}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 T}{\partial \gamma_{i,1} \partial \gamma_{i,p}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \cdots & \frac{\partial^2 T}{\partial \gamma_{i,p}^2}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}).$$

Let

$$\begin{aligned} \chi(\tilde{\gamma}, \mathbf{X}_{t-1}) &= -\frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\tilde{\gamma}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] \\ &\quad + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \left(\mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1}) \right) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})]. \end{aligned}$$

Convergence of $\chi(\tilde{\gamma}, \mathbf{X}_{t-1})$ to $\chi(\gamma_0, \mathbf{X}_{t-1})$

We have for any $t = 1, \dots, n$,

$$\begin{aligned} &|\chi(\tilde{\gamma}, \mathbf{X}_{t-1}) - \chi(\gamma_0, \mathbf{X}_{t-1})| \\ &= \left| -\frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\tilde{\gamma}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] \right. \\ &\quad + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] \\ &\quad + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \\ &\quad \left. - \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \right| \\ &= \left| -\frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\tilde{\gamma}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \right. \\ &\quad \times \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] - \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \\ &\quad \times \boldsymbol{\beta}^\top \mathcal{H}(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] - \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] \\ &\quad + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})] + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \\ &\quad \left. \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] - \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \right| \end{aligned}$$

$$\begin{aligned}
 &= \left| \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \right. \\
 &\times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] - \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \\
 &\times \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \\
 &- \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \\
 &\times \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \left. \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right| \\
 &\leq \left| \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \{ \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \boldsymbol{\beta}^\top M(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \} \right| \\
 &+ \left| \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \{ \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] - \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \} \right| \\
 &+ \left| \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \{ \boldsymbol{\beta}^\top \mathcal{H}(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \} \right| \\
 &+ \left| \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \{ \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \} \right| \\
 &\leq \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| |\boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} - \boldsymbol{\beta}^\top M(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) \boldsymbol{\beta}| \\
 &+ \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} |\boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] - \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| \\
 &+ \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| |\boldsymbol{\beta}^\top \mathcal{H}(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) \boldsymbol{\beta} - \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}| \\
 &+ \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} |\boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 &= A_{1n} + A_{2n} + A_{3n} + A_{4n}.
 \end{aligned}$$

Let

$$\begin{aligned}
 A_{1n} &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| |\boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} - \boldsymbol{\beta}^\top M(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) \boldsymbol{\beta}| \\
 A_{2n} &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} |\boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] - \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]|
 \end{aligned}$$

$$A_{3n} = \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| |\boldsymbol{\beta}^\top \mathcal{H}(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})\boldsymbol{\beta} - \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\boldsymbol{\beta}|$$

$$A_{4n} = \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} |\boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\boldsymbol{\beta}| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.$$

Now, we study each term.

Remark 3.4.2. $\mathbb{E}\{\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]\} = 0$.

Proof:
$$\mathbb{E}\{\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]\} = \int_{\mathbb{R}} \phi_f(x) f(x) dx = \int_{\mathbb{R}} -\frac{f'}{f}(x) f(x) dx$$

$$= \int_{\mathbb{R}} f'(x) dx = 0.$$

Remark 3.4.3. $\mathbb{E}\{\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]\} = I(f)$.

Proof:
$$\mathbb{E}\{\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]\} = \int_{\mathbb{R}} \phi'_f(x) f(x) dx = - \int_{\mathbb{R}} \left[\frac{f''(x)}{f(x)} - \frac{f'^2(x)}{f^2(x)} \right] f(x) dx$$

$$= - \int_{\mathbb{R}} f''(x) dx + \int_{\mathbb{R}} \phi_f^2(x) f(x) dx = I(f).$$

Since ϕ'_f is c_ϕ -lipschitz, we have

$$A_{2n} = \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} |\boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\boldsymbol{\beta}| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] - \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]|$$

$$\leq \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} c_\phi |\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) - \varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})| \|\boldsymbol{\beta}\|_{p(k+1)}^2 \nu^2(\mathbf{X}_{t-1}).$$

We have

$$|\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) - \varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})| = \left| \frac{T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1}) - T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|.$$

Applying a Taylor expansion on $\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})$ in a neighborhood of $\boldsymbol{\gamma}_0$, and in view of (\mathcal{A}_9) , we obtain

$$|\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) - \varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})| = \left| \frac{(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)^\top D_\gamma [T(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})]}{V(\mathbf{X}_{t-1})} \right| \tag{3.6}$$

$$\leq \frac{\|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} \left\| D_\gamma [T(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})] \right\|_{p(k+1)}}{|V(\mathbf{X}_{t-1})|}$$

$$\leq \frac{|\nu(\mathbf{X}_{t-1})| \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)}}{|V(\mathbf{X}_{t-1})|}.$$

Then,

$$\begin{aligned}
 A_{2n} &\leq \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V^3(\mathbf{X}_{t-1})|} c_\phi \|\tilde{\gamma} - \gamma_0\|_{p(k+1)} \left\| D_\gamma [T(\tilde{\gamma}, \mathbf{X}_{t-1})] \right\|_{p(k+1)} \|\beta\|_{p(k+1)}^2 \nu^2(\mathbf{X}_{t-1}) \\
 &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V^3(\mathbf{X}_{t-1})|} c_\phi \|\tilde{\gamma} - \gamma_0\|_{p(k+1)} \|\beta\|_{p(k+1)}^2 |\nu(\mathbf{X}_{t-1})|^3 \\
 &= \frac{1}{2} \sum_{j=1}^{k+1} c_\phi \|\tilde{\gamma} - \gamma_0\|_{p(k+1)} \|\beta\|_{p(k+1)}^2 \frac{n_j(n)}{n} \times \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^3.
 \end{aligned}$$

We have

$$\|\tilde{\gamma} - \gamma_0\|_{p(k+1)} \leq \|\gamma_n - \gamma_0\|_{p(k+1)} = \|\beta/\sqrt{n}\|_{p(k+1)} \xrightarrow{n \rightarrow +\infty} 0.$$

By (\mathcal{A}_{10}) and by the ergodic theorem, for $j = 1, \dots, n$, almost surely, we have

$$\lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^3 = \int_{\mathbb{R}^d} \left| \frac{\nu(x)}{V(x)} \right|^3 dF_j(x) < +\infty.$$

Then,

$$\lim_{n \rightarrow +\infty} A_2 = 0.$$

By (\mathcal{A}_9) , we can write

$$\begin{aligned}
 A_{3n} &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})]| \left| \beta^\top \mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1})\beta - \beta^\top \mathcal{H}(\gamma_0, \mathbf{X}_{t-1})\beta \right| \\
 &\leq \frac{1}{2} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})]| \|\beta_j\|_p^2 \|\mathcal{H}_j(\tilde{\gamma}, \mathbf{X}_{t-1}) - \mathcal{H}_j(\gamma_0, \mathbf{X}_{t-1})\|_p.
 \end{aligned}$$

Let $\mathcal{J}(\tilde{\gamma} - \gamma_0, \mathbf{X}_{t-1}) = \mathcal{H}_j(\tilde{\gamma}, \mathbf{X}_{t-1}) - \mathcal{H}_j(\gamma_0, \mathbf{X}_{t-1})$. Applying a first-order Taylor expansion on ϕ_f in a neighborhood of $\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)$, for some $\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})$ lying between $\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)$ and $\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma})$, we obtain

$$\begin{aligned}
A_{3n} &\leq \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| \left[|\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| + |\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}) - \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)| \phi_f'[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \right] \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\leq \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\quad + \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| |\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}) - \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)| |\phi_f'[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
&\quad \times \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p + \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| |\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}) - \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)|^2 \\
&\quad \times \left| \phi_f''[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \right| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\leq \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\quad + \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} \|D_{\boldsymbol{\gamma}}[T(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]\|_{p(k+1)} \\
&\quad \times |\phi_f'[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\quad + \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 \|D_{\boldsymbol{\gamma}}[T(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]\|_{p(k+1)}^2 \\
&\quad \times \left| \phi_f''[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \right| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\leq \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\quad + \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} |\phi_f'[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p \\
&\quad + \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu^2(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 c_{\phi} \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p.
\end{aligned}$$

Remark 3.4.4. $\mathbb{E}\{|\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]|\} < \infty$.

Proof: In view of (\mathcal{A}_8) , $\mathbb{E}\{|\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]|\} = \int_{\mathbb{R}} |\phi_f(x)| f(x) dx = \int_{\mathbb{R}} \left[\frac{|f'(x)|}{f(x)} f(x) \right] dx = \int_{\mathbb{R}} |f'(x)| dx < \infty$.

For $j = 1, \dots, k + 1$,

$$\beta_j^\top \mathcal{H}_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \beta_j = \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial^2 T}{\partial \gamma_{j,h} \partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}),$$

the function $\partial^2 T / \partial \gamma_{j,h} \partial \gamma_{j,m}$ is a continuous function and since $\lim_{n \rightarrow +\infty} \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} = 0$, by continuity, it is easy to conclude that $\lim_{n \rightarrow +\infty} \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p = 0$.

By the ergodic theorem, for $j = 1, \dots, n$, almost surely, we obtain

- $\lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{1}{V(\mathbf{X}_{t-1})} \right| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p$
 $= 0 \times \int_{\mathbb{R}^d} \left| \frac{1}{V(x)} \right| dF_j(x) \mathbb{E}\{|\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|\} = 0$
- $\lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p$
 $= \lim_{n \rightarrow +\infty} \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} \int_{\mathbb{R}^d} \left| \frac{\nu(x)}{V(x)} \right| dF_j(x) I(f) \times \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p = 0$
- $\lim_{n \rightarrow +\infty} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})^2}{V(\mathbf{X}_{t-1})} \right| \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 |\phi''_f[\widetilde{\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})}]| \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p$
 $= \lim_{n \rightarrow +\infty} \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 \int_{\mathbb{R}^d} \left| \frac{\nu(x)}{V(x)} \right|^2 dF_j(x) \times c_\phi \times \|\mathcal{J}(\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p = 0$

Applying a first-order Taylor expansion to A_{4n} in a neighborhood of $\boldsymbol{\gamma}_0$, we obtain

$$\begin{aligned} A_{4n} &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} |\boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}| |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} |\boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}| \left| (\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0)^\top D_\gamma [T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})] \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] \right| \\ &\leq \frac{1}{2n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \|\boldsymbol{\beta}\|_{p(k+1)}^2 |\nu(\mathbf{X}_{t-1})|^2 \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} \\ &\quad \times (\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] + \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]) \\ &= \frac{1}{2} \sum_{j=1}^{k+1} \|\boldsymbol{\beta}\|_{p(k+1)}^2 \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} \frac{n_j(n)}{n} \frac{1}{n_j} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{|V(\mathbf{X}_{t-1})|} |\nu(\mathbf{X}_{t-1})|^2 \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\ &\quad + \frac{1}{2} \sum_{j=1}^{k+1} \|\boldsymbol{\beta}\|_{p(k+1)}^2 \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} \frac{n_j(n)}{n} \frac{1}{n_j} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{|V(\mathbf{X}_{t-1})|} |\nu(\mathbf{X}_{t-1})|^2 \\ &\quad \times (\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})] - \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]). \end{aligned}$$

In view of (\mathcal{A}_6) and by the ergodic theorem, we can easily show that

$$\lim_{n \rightarrow +\infty} A_{4n} = 0.$$

Now,

$$\begin{aligned} A_{1n} &= \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| |\boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\boldsymbol{\beta} - \boldsymbol{\beta}^\top M(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})\boldsymbol{\beta}| \\ &= \frac{1}{2} \sum_{j=1}^{k+1} n_j \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| |\beta_j^\top [M_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) - M_j(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})] \beta_j|. \end{aligned}$$

Applying a first-order Taylor expansion to ϕ'_f and ε_t in a neighborhood of $\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)$, for some $\widetilde{\varepsilon}_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})$ lying between $\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)$ and $\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})$, we obtain

$$\begin{aligned} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| &= \left| \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] + (\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}) - \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)) \phi''_f(\varepsilon_t(\tilde{\boldsymbol{\gamma}})) \right| \\ &\leq |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| + |(\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}) - \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0))| |\phi''_f(\varepsilon_t(\widetilde{\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}}))|. \end{aligned}$$

By the assumptions (\mathcal{A}_9) and (\mathcal{A}_8) , we obtain

$$|\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}})]| \leq |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| + c_\phi \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|.$$

Furthermore, we have

$$\begin{aligned} A_{1n} &\leq \frac{1}{2} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| |\beta_j^\top [M_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) - M_j(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})] \beta_j| \\ &\quad + \frac{1}{2} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{c_\phi \|\tilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0\|_{p(k+1)}}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})} \right| |\beta_j^\top [M_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) - M_j(\tilde{\boldsymbol{\gamma}}, \mathbf{X}_{t-1})] \beta_j|. \end{aligned}$$

We have

$$\begin{aligned} &|\beta_j^\top M_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\beta_j| = \\ &\omega_j^2 \begin{pmatrix} \beta_{j,1} \\ \vdots \\ \beta_{j,p} \end{pmatrix}^\top \begin{pmatrix} \frac{\partial T}{\partial \gamma_{j,1}}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,1}}(\boldsymbol{\gamma}_{0,j,1}, \mathbf{X}_{t-1}) & \cdots & \frac{\partial T}{\partial \gamma_{j,1}}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,p}}(\boldsymbol{\gamma}_{0,j,p}, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \frac{\partial T}{\partial \gamma_{j,p}}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,1}}(\boldsymbol{\gamma}_{0,j,1}, \mathbf{X}_{t-1}) & \cdots & \frac{\partial T}{\partial \gamma_{j,p}}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,p}}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \end{pmatrix} \begin{pmatrix} \beta_{j,1} \\ \vdots \\ \beta_{j,p} \end{pmatrix} \\ &= \sum_{1 \leq h \leq m \leq p} \omega_j^2 \beta_{j,h} \beta_{j,m} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}). \end{aligned}$$

Remark 3.4.5. $\omega_j^2 = \omega_j \in \{0, 1\}$.

Then

$$\begin{aligned}
 & |\beta_j^\top [M_j(\gamma_0, \mathbf{X}_{t-1}) - M_j(\tilde{\gamma}_j, \mathbf{X}_{t-1})] \beta_j| \\
 &= \sum_{1 \leq h \leq m \leq p} \omega_j^2 \beta_{j,h} \beta_{j,m} \\
 &\times \left\{ \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) - \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right\}.
 \end{aligned}$$

We have

$$\begin{aligned}
 & \left| \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) - \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right| \\
 &= \left| \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) - \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right. \\
 &+ \left. \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) - \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right| \\
 &\leq \left| \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \left[\frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) - \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right] \right| \\
 &+ \left| \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \left[\frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) - \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right] \right|.
 \end{aligned}$$

A first-order Taylor expansion of $\partial T / \partial \gamma_{j,h}$ and $\partial T / \partial \gamma_{j,m}$, for $h, m = 1, \dots, p$, in a neighborhood of γ_0 yields

$$\begin{aligned}
 & \left| \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) - \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right| \\
 &\leq \left| \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \left[(\gamma_0 - \tilde{\gamma}) D_\gamma \left[\frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right] \right] \right| \\
 &+ \left| \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \left[(\gamma_0 - \tilde{\gamma}) D_\gamma \left[\frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right] \right] \right| \\
 &\leq \left| \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \right| \|\gamma_0 - \tilde{\gamma}\|_{p(k+1)} \left\| D_\gamma \left[\frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right] \right\|_{p(k+1)} \\
 &+ \left| \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right| \|\gamma_0 - \tilde{\gamma}\|_{p(k+1)} \left\| D_\gamma \left[\frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \right] \right\|_{p(k+1)} \\
 &\leq 2|\nu^2(\mathbf{X}_{t-1})| \|\gamma_0 - \tilde{\gamma}\|_{p(k+1)}.
 \end{aligned}$$

Then,

$$|\beta_j^\top [M_j(\gamma_0, \mathbf{X}_{t-1}) - M_j(\tilde{\gamma}_j, \mathbf{X}_{t-1})] \beta_j| \leq 2|\nu^2(\mathbf{X}_{t-1})| \|\gamma_0 - \tilde{\gamma}\|_{p(k+1)} \sum_{1 \leq h \leq m \leq p} \omega_j^2 \beta_{j,h} \beta_{j,m}.$$

Finally we obtain

$$|A_{1n}| \leq \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{\|\gamma_0 - \tilde{\gamma}\|_{p(k+1)} \sum_{1 \leq h \leq m \leq p} \omega_j^2 \beta_{j,h} \beta_{j,m}}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^2 |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| \\ + \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{c_\phi \|\tilde{\gamma} - \gamma_0\|_{p(k+1)}^2 \sum_{1 \leq h \leq m \leq p} \omega_j^2 \beta_{j,h} \beta_{j,m}}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^3.$$

The summation $\sum_{1 \leq h \leq m \leq p} \omega_j^2 \beta_{j,h} \beta_{j,m}$ is finite. By the ergodic theorem, for $j = 1, \dots, n$, almost surely, we have

$$\lim_{n \rightarrow +\infty} \frac{\|\gamma_0 - \tilde{\gamma}\|_{p(k+1)}}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^2 |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| \\ = \lim_{n \rightarrow +\infty} \|\gamma_0 - \tilde{\gamma}\|_{p(k+1)} \int_{\mathbb{R}^d} \left| \frac{\nu(x)}{V(x)} \right|^2 dF_j(x) \mathbb{E} \{ |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| \} = 0$$

and

$$\lim_{n \rightarrow +\infty} \frac{\|\gamma_0 - \tilde{\gamma}\|_{p(k+1)}^2}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^3 = 0 \times \int_{\mathbb{R}^d} \left| \frac{\nu(x)}{V(x)} \right|^3 dF_j(x) = 0.$$

Thus,

$$|\chi(\tilde{\gamma}, \mathbf{X}_{t-1}) - \chi(\gamma_0, \mathbf{X}_{t-1})| \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

It results from above that

$$\Theta_n = \Pi_n - \Delta_n + o_P(1),$$

where

- $\Delta_n = \frac{1}{2n} \sum_{t=1}^n \left\{ \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \right. \\ \left. - \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\gamma_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \right\},$
- $\Pi_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \right\}.$

□

3.4.2 Asymptotic behavior of the test statistic

Theorem 2. *Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) hold. Then, as $n \rightarrow +\infty$, for any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, under H_0 ,*

i. $\Delta_n \longrightarrow \frac{\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{2},$

ii. $\Pi_n \longrightarrow \mathcal{N}(0, \eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})),$

where, for $1 \leq h \leq m \leq p,$

$$\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0),$$

$$\eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) dF_j(x).$$

Proof: We first Study of the asymptotic behavior of Δ_n under H_0 .

$$\begin{aligned} \Delta_n &= \frac{1}{2n} \sum_{t=1}^n \left\{ \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right. \\ &\quad \left. - \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top \mathcal{H}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\ &= \frac{1}{2} \sum_{j=1}^{k+1} \frac{1}{n} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}_j^\top M_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}_j \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\ &\quad - \frac{1}{2} \sum_{j=1}^{k+1} \frac{1}{n} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}_j^\top \mathcal{H}_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}_j \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\ &= \frac{1}{2} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]}{V^2(\mathbf{X}_{t-1})} \\ &\quad \times \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \\ &\quad - \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]}{V(\mathbf{X}_{t-1})} \left[\sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial^2 T}{\partial \gamma_{j,h} \partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \right], \end{aligned}$$

where

$$\boldsymbol{\beta}_j^\top M_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}_j = \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})$$

and

$$\boldsymbol{\beta}_j^\top \mathcal{H}_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \boldsymbol{\beta}_j = \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial^2 T}{\partial \gamma_{j,h} \partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}).$$

By the piece-wise stationarity and ergodicity, in view of 3.4.3, for any $j = 1, \dots, n,$ we

can write

$$\begin{aligned}
& \lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^2(\mathbf{X}_{t-1})} \left[\sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \right] \\
& \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\
& = \mathbb{E} \left\{ \frac{1}{V^2(\mathbf{X}_0)} \left[\sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_0) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_0) \right] \right\} \mathbb{E} \left\{ \phi'_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\
& = \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) dF_j(x) \\
& = \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) < \infty
\end{aligned}$$

and

$$\begin{aligned}
& \lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V(\mathbf{X}_{t-1})} \left(\sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial^2 T}{\partial \gamma_{j,h} \partial \gamma_{j,m}}(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \right) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\
& = \int_{\mathbb{R}^d} \frac{1}{V(x)} \left[\sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial^2 T}{\partial \gamma_{j,h} \partial \gamma_{j,m}}(\boldsymbol{\gamma}_0, x) \right] dF_j(x) \times \mathbb{E} \left\{ \phi_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\
& = 0.
\end{aligned}$$

Then, almost surely, we have

$$\lim_{n \rightarrow +\infty} \Delta_n = \frac{1}{2} \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = \frac{\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{2}.$$

Thus we can write

$$\Theta_n = \Pi_n - \frac{\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{2} + o_P(1).$$

Now we prove that under H_0 ,

$$\Pi_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, \eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})).$$

We consider the sequence

$$\Theta_{n,j} = \frac{1}{\sqrt{n}} \sum_{t=1}^j \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)], \quad j = 1, \dots, n$$

and we define for every $t = 1, \dots, n$,

$$\psi_{n,t} = \frac{1}{\sqrt{n}} \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)].$$

We use Corollary 2.3.2 of Hall and Heyde (2014) to study of the asymptotic behavior of $\Theta_{n,j}$.

First, we prove that $\{(\Theta_{n,j}, G_j), j = 1, \dots, n\}$ is a martingale sequence.

Let $q, w \in \mathbb{Z}$ such that $q < w$, then

$$\begin{aligned} \mathbb{E}\left(\Theta_{n,w} | \mathcal{F}_q\right) &= \mathbb{E}\left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^w \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_q \right\} \\ &= \mathbb{E}\left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^q \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_q \right\} \\ &\quad + \mathbb{E}\left\{ \frac{1}{\sqrt{n}} \sum_{t=q+1}^w \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_q \right\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^q \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=q+1}^w \mathbb{E}\left\{ \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \middle| \mathcal{F}_q \right\} \mathbb{E}\left\{ \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\}. \end{aligned}$$

By Remark (3.4.2), we obtain

$$\mathbb{E}\left(\Theta_{n,w} | \mathcal{F}_q\right) = \Theta_{n,q}.$$

Next, we check the first condition of Corollary 2.3.2 of Hall and Heyde (2014):

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(\psi_{n,t}^2 | \mathcal{F}_{t-1}) &= \frac{1}{n} \sum_{t=1}^n \mathbb{E}\left\{ \frac{[\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]^2}{V^2(\mathbf{X}_{t-1})} \phi_f^2[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_{t-1} \right\} \\ &= \frac{1}{n} \sum_{t=1}^n \frac{[\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]^2}{V^2(\mathbf{X}_{t-1})} \mathbb{E}\left\{ \phi_f^2[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_{t-1} \right\}. \end{aligned}$$

Since ε_t is independent of \mathcal{F}_{t-1} for $t = 1, \dots, n$, then

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(\psi_{n,t}^2 | \mathcal{F}_{t-1}) &= \frac{1}{n} \sum_{t=1}^n \frac{[\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]^2}{V^2(\mathbf{X}_{t-1})} \mathbb{E}\left\{ \phi_f^2[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\ &= \frac{1}{n} \sum_{j=1}^{k+1} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{[\boldsymbol{\beta}_j^\top N_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]^2}{V^2(\mathbf{X}_{t-1})} \mathbb{E}\left\{ \phi_f^2[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\ &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{I(f)}{V^2(\mathbf{X}_{t-1})} \\ &\quad \times \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}). \end{aligned}$$

By the ergodic theorem, for any $t = 1, \dots, n$, we obtain almost surely

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{I(f)}{V^2(\mathbf{X}_{t-1})} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \\ &= \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) < \infty. \end{aligned}$$

Then,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \sum_{t=1}^n \mathbb{E}(\psi_{n,t}^2 | \mathcal{F}_{t-1}) &= \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) \\ &= \eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) < \infty. \end{aligned}$$

Finally, we check the Linderberg condition.

Let $\epsilon > 0$, by the Hölder inequality we can write

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(\psi_{n,t}^2 \mathbf{1}_{|\psi_{n,t}| > \epsilon} | \mathcal{F}_{t-1}) &\leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}}(\psi_{n,t}^3 | \mathcal{F}_{t-1}) \mathbb{E}^{\frac{1}{3}}(\mathbf{1}_{|\psi_{n,t}| > \epsilon} | \mathcal{F}_{t-1}) \\ &= \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ n^{-\frac{3}{2}} \frac{[\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]^3}{V^3(\mathbf{X}_{t-1})} \phi_f^3[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_{t-1} \right\} \\ &\quad \times \mathbb{P}^{\frac{1}{3}} \left(|\psi_{n,t}| > \epsilon \middle| \mathcal{F}_{t-1} \right). \end{aligned}$$

By Markov inequality, we can write

$$\begin{aligned} \sum_{t=1}^n \mathbb{E}(\psi_{n,t}^2 \mathbf{1}_{|\psi_{n,t}| > \epsilon} | \mathcal{F}_{t-1}) &\leq \frac{1}{n} \sum_{t=1}^n \left\{ \frac{[\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})]^2}{V^2(\mathbf{X}_{t-1})} \mathbb{E}^{\frac{2}{3}} \{ \phi_f^3[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_{t-1} \} \right. \\ &\quad \left. \times \frac{\mathbb{E}^{\frac{1}{3}} \left(|\psi_{n,t}|^3 \middle| \mathcal{F}_{t-1} \right)}{\epsilon^{\frac{1}{3}}} \right\} \\ &\leq \frac{1}{n^{\frac{3}{2}} \epsilon^{\frac{1}{3}}} \sum_{t=1}^n \frac{|\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})|^3}{V^3(\mathbf{X}_{t-1})} \mathbb{E}^{\frac{2}{3}} \{ \phi_f^3[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_{t-1} \} \\ &\quad \times \mathbb{E}^{\frac{1}{3}} \left\{ |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|^3 \middle| \mathcal{F}_{t-1} \right\} \\ &\leq \frac{1}{n^{\frac{3}{2}} \epsilon^{\frac{1}{3}}} \sum_{t=1}^n \frac{|\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})|^3}{V^3(\mathbf{X}_{t-1})} \mathbb{E} \left\{ |\phi_f^3[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \middle| \mathcal{F}_{t-1} \right\} \\ &= \frac{1}{\epsilon^{\frac{1}{3}}} \sum_{t=1}^{k+1} \left(\frac{n_j(n)}{n} \right)^{\frac{3}{2}} \frac{1}{n_j(n)^{\frac{3}{2}}} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^3(\mathbf{X}_{t-1})} |\beta_j^\top N_j(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})|^3 \\ &\quad \times \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|^3 \right\}. \end{aligned}$$

By the ergodic theorem, in view of (\mathcal{A}_7) , for any $j = 1, \dots, k + 1$, almost surely we can write

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \frac{1}{n_j(n)^{\frac{3}{2}}} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^3(\mathbf{X}_{t-1})} |\beta_j^\top N_j(\gamma_0, \mathbf{X}_{t-1})|^3 \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| \right\} \\ &= \lim_{n \rightarrow +\infty} \frac{1}{\sqrt{n_j(n)}} \lim_{n \rightarrow +\infty} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^3(\mathbf{X}_{t-1})} |\beta_j^\top N_j(\gamma_0, \mathbf{X}_{t-1})|^3 \mathbb{E} \left\{ |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]|^3 \right\} \\ &= 0 \times \int_{\mathbb{R}^d} \frac{1}{V^3(x)} |\beta_{j-1}^\top N_{j-1}(\gamma_0, x)|^3 dF_j(x) \times \int_{\mathbb{R}} |\phi_f(x)|^3 f(x) dx \\ &= 0. \end{aligned}$$

The conditions of Corollary 2.3.2 of Hall and Heyde (2014) are completely verified. Then under H_0 we have

$$\Pi_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, \eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})). \quad (3.7)$$

Consequently, under H_0 , we have

$$\Theta_n \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{\eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})}{2}, \eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})\right). \quad (3.8)$$

□

Collecting the above results, the **LAN** property is established.

Denote the central sequence by

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \right\}. \quad (3.9)$$

Corollary 3.4.1. *Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) hold. Then, for any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, the sequences $\{H_{\boldsymbol{\beta}}^{(n)} : n \geq 1\}$ and $\{H_0^{(n)} = H_0 : n \geq 1\}$ are contiguous. Also, as $n \rightarrow +\infty$, under $H_{\boldsymbol{\beta}}^{(n)}$, we have*

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} \mathcal{N}(\eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}), \eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})).$$

Proof: For any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, using Theorem 2, as $n \rightarrow +\infty$ and under H_0 , we have

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})).$$

It results that, under H_0 , as $n \rightarrow +\infty$,

$$\Theta_n \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{\eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})}{2}, \eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})\right).$$

According to the above result, we can write, under H_0 , as $n \rightarrow +\infty$,

$$\begin{pmatrix} \Pi_n(\gamma_0, \beta) \\ \Theta_n \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N} \left(\begin{pmatrix} 0 \\ -\frac{\eta(\rho_0, \gamma_0, \beta)}{2} \end{pmatrix}, \begin{pmatrix} \eta(\rho_0, \gamma_0, \beta) & \sigma_{1,2} \\ \sigma_{2,1} & \eta(\rho_0, \gamma_0, \beta) \end{pmatrix} \right),$$

where $\sigma_{1,2} = \sigma_{2,1} = \lim_{n \rightarrow +\infty} \text{Cov}(\Pi_n, \Theta_n) = \lim_{n \rightarrow +\infty} [\mathbb{E}(\Pi_n \Theta_n) - \mathbb{E}(\Pi_n)\mathbb{E}(\Theta_n)]$.

We have $\mathbb{E}(\Pi_n) = 0$. Then

$$\lim_{n \rightarrow +\infty} \mathbb{E}(\Theta_n \Pi_n) = \lim_{n \rightarrow +\infty} \left[\mathbb{E}(\Pi_n^2) + \mathbb{E}(\Pi_n \Delta_n) \right] = \eta(\rho_0, \gamma_0, \beta).$$

Then, under H_0 , we have that

$$\begin{pmatrix} \Pi_n(\rho_0, \gamma_0, \beta) \\ \Theta_n \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N} \left(\begin{pmatrix} 0 \\ -\frac{\eta(\rho_0, \gamma_0, \beta)}{2} \end{pmatrix}, \begin{pmatrix} \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \\ \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \end{pmatrix} \right). \quad (3.10)$$

Using Le Cam (1986) or Dreesbeke and Fine (1996), we obtain that the sequences $\{H_\beta^{(n)} : n \geq 1\}$ and $\{H_0^{(n)} = H_0 : n \geq 1\}$ are contiguous, and that under $H_\beta^{(n)}$, as $n \rightarrow +\infty$,

$$\Pi_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(\eta(\rho_0, \gamma_0, \beta), \eta(\rho_0, \gamma_0, \beta)).$$

□

For known γ_0 and for any $\beta \in \mathbb{R}^{p(k+1)}$, for testing H_0 against $H_\beta^{(n)}$, we base our test on the statistic

$$\mathcal{T}_n(\rho_0, \gamma_0, \beta) = \frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\widehat{\pi}_n(\rho_0, \gamma_0, \beta)}, \quad (3.11)$$

where

- $\widehat{\pi}_n(\rho_0, \gamma_0, \beta) = \sqrt{\widehat{\eta}_n(\gamma_0, \beta)}$,
- $\widehat{\eta}_n(\rho_0, \gamma_0, \beta) = \sum_{j=1}^{k+1} \widehat{\alpha}_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \widehat{\eta}_{j,2}^{(h,m)}(\rho_0, \gamma_0)$,
- $\widehat{\eta}_{j,2}^{(h,m)}(\rho_0, \gamma_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_0, \gamma_0, x) d\widehat{F}_j(x)$, where \widehat{F}_j is the empirical distribution function of vector \mathbf{X}_t with components in $[\tau_{j-1}, \tau_j]$.
- An estimation of $\eta_{j,2}^{(h,m)}(\rho_0, \gamma_0)$ can be taken to be

$$\widehat{\eta}_{j,2}(\rho_0, \gamma_0) = \frac{I(f)}{n_j(n)} \sum_{\tau_{j-1}}^{\tau_j} \frac{1}{V^2(\mathbf{X}_{t-1})} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_0, \gamma_0, \mathbf{X}_{t-1}),$$

$$\text{and } \widehat{\alpha}_j = \frac{n_j(n)}{n}.$$

Theorem 3. Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) hold. Then, as $n \rightarrow +\infty$, for any $\beta \in \mathbb{R}^{p(k+1)}$,

i. Under H_0 , $\mathcal{T}_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$.

ii. Under $H_\beta^{(n)}$, at level of significance $\alpha \in [0, 1]$, the asymptotic power of test based on $\mathcal{T}_n(\rho_0, \gamma_0, \beta)$ is

$$\mathcal{P}_{k, \tau^k} = 1 - \Phi(z_\alpha - \pi(\rho_0, \gamma_0, \beta)),$$

where $\tau^k = (\tau_1, \dots, \tau_k)^\top$ and z_α is the $(1 - \alpha)$ -quantile of a standard Gaussian distribution with cumulative distribution function Φ .

iii. The test based on $\mathcal{T}_n(\rho_0, \gamma_0, \beta)$ is locally asymptotic optimal.

Proof of (i):

From Theorem 2 and Corollary 3.4.1, we can conclude that, under H_0 as $n \rightarrow +\infty$,

$$\begin{pmatrix} \Pi_n(\rho_0, \gamma_0, \beta) \\ \Theta_n \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N} \left(\begin{pmatrix} 0 \\ -\frac{\eta(\rho_0, \gamma_0, \beta)}{2} \end{pmatrix}, \begin{pmatrix} \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \\ \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \end{pmatrix} \right). \quad (3.12)$$

The part (i) is a direct consequence of Theorem 2 and is briefly explained in the proof of Corollary 3.4.1. As explained there, the sequences of hypothesis are contiguous and we have under $H_\beta^{(n)}$, as $n \rightarrow +\infty$:

$$\Pi_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(\eta(\rho_0, \gamma_0, \beta), \eta(\rho_0, \gamma_0, \beta)).$$

By (3.10) and the Le Cam's third lemma (Proposition 4.2 in Le Cam (1986)), under $H_\beta^{(n)}$, as $n \rightarrow +\infty$,

$$\Pi_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(\eta(\rho_0, \gamma_0, \beta), \eta(\rho_0, \gamma_0, \beta)).$$

We have that, under H_0 , as $n \rightarrow +\infty$,

$$\widehat{\pi}_n(\rho_0, \gamma_0, \beta) \longrightarrow \pi(\rho_0, \gamma_0, \beta),$$

where $\pi(\rho_0, \gamma_0, \beta) = \sqrt{\eta(\rho_0, \gamma_0, \beta)}$.

This convergence remains true under $H_\beta^{(n)}$ by contiguity. From Theorem 2, it can be seen that, as $n \rightarrow +\infty$, under H_0 ,

$$\mathcal{T}_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1).$$

Thus, by the Le Cam's third lemma, we can conclude that under $H_\beta^{(n)}$, as $n \rightarrow +\infty$,

$$\frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\widehat{\pi}_n(\rho_0, \gamma_0, \beta)} \xrightarrow{\mathcal{D}} \mathcal{N}(\pi(\rho_0, \gamma_0, \beta), 1).$$

For $n \geq 1$, we can write

$$\frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} = \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \times \frac{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}.$$

It results that, under $H_{\boldsymbol{\beta}}^{(n)}$ and as $n \rightarrow +\infty$,

$$\frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \xrightarrow{\mathcal{D}} \mathcal{N}(\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}), 1).$$

□

Proof of (ii) and (iii):

To calculate the asymptotic power of our test statistic, we should calculate the asymptotic cumulative distribution of $\frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}$ under $H_{\boldsymbol{\beta}}^{(n)}$.

$$\begin{aligned} \lim_{n \rightarrow +\infty} P\left(\frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} > z_{\alpha} \mid H_{\boldsymbol{\beta}}^{(n)}\right) &= \lim_{n \rightarrow +\infty} P\left(\frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} > z_{\alpha} \mid H_{\boldsymbol{\beta}}^{(n)}\right) \\ &= 1 - \Phi(z_{\alpha} - \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})) \\ &= \mathcal{P}_{k, \tau^k} \end{aligned} \quad (3.13)$$

where Φ is the cumulative distribution function of a standard Gaussian law with z_{α} its $(1 - \alpha)$ -quantile.

By Section 4.4.3 of Dreesbeke and Fine (1996), the test based on $\mathcal{T}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ is locally asymptotically optimal. □

3.5 Conclusion

In this chapter, we have proved that the families of likelihood under study have a LAN property. Next, we have proved the locally asymptotically optimality of our test built in the case where the functions T , V f and the parameters $\boldsymbol{\rho}_0$, $\boldsymbol{\gamma}_0$ are known. An explicit expression of the local power of the test is given. The case where these parameters are unknown is considered in the next chapter.

Chapter 4

Testing the conditional mean of a time-dependent coefficients CHARN model: Part II

4.1 Introduction

In this chapter, we place ourselves in the framework of the model (1.1) with functions T, V and f having known parametric forms but with the parameter $\boldsymbol{\rho}_0$ assumed to be unknown. This is the situation encountered in practice. Our idea here is to do the same work as in the preceding chapter, with $\boldsymbol{\rho}_0$ and $\boldsymbol{\gamma}_0$ replaced by their estimators. Doing this, similar results are established.

In the sequel, we discuss the parameter estimation, and we study the test statistic defined in Chapter 3 when the estimators are substituted for the parameters. It is found that the results remain asymptotically unchanged.

4.2 The functions involved in the model and the nuisance parameters are unknown

In this section, we give a brief review on parameter estimation in time series models comprised in CHARN models. We make additional assumptions and we study the asymptotic behavior of the new version of the test statistic defined in the preceding chapter.

4.2.1 Parameters estimation

There are many methods for parameter estimation in time series models. Most of them are of least-squares or maximum likelihood types. Ngatchou-Wandji (2008), Bollerslev (1987) and many others study this problem. For linear models as ARMA, one can refer to Brockwell et al. (1990). Here, we summarize a more recent paper dealing with a special CHARN model. For $\mathbf{X}_t = (X_t, \dots, X_{t-p+1})^\top$, $\boldsymbol{\varepsilon}_t = (\varepsilon_t, 0, \dots, 0)^\top$ and $\mathbf{C}_0 = (c_0, 0, \dots, 0)^\top$, Chen et al. (2018) define many extension of GExpAR models, among which, the following one

$$X_t = c_0 + \sum_{i=1}^p \left(c_i + \pi_i e^{-\gamma_i (X_{t-i} - z_i)^2} \right) X_{t-i} + \varepsilon_t. \quad (4.1)$$

The authors prove the stationarity of this model. They show the implementation of VP (variable projection) method to optimize its parameters. They denote by $\boldsymbol{\theta}_L$ and $\boldsymbol{\theta}_N$, respectively, the set of linear and nonlinear parameters:

$$\begin{aligned} \boldsymbol{\theta}_L &\triangleq (c_0, c_1, \pi_1, \dots, c_p, \pi_p) \in \mathbb{R}^{2p+1} \\ \boldsymbol{\theta}_N &\triangleq (\gamma_1, z_1, \dots, \gamma_p, z_p) \in \mathbb{R}^{2p}. \end{aligned}$$

In order to apply VP method, they write the model (4.1) as

$$X_t = \boldsymbol{\Psi}(\boldsymbol{\theta}_N, \mathbf{X}_{t-1})^\top \cdot \boldsymbol{\theta}_L + \varepsilon_t,$$

where

$$\begin{aligned} \boldsymbol{\Psi}(\boldsymbol{\theta}_N, \mathbf{X}_{t-1}) &= \left(1, X_{t-1}, X_{t-1} \cdot e^{-\gamma_1 (X_{t-1} - z_1)^2}, \right. \\ &\quad \left. X_{t-2}, X_{t-2} \cdot e^{-\gamma_2 (X_{t-2} - z_2)^2}, \dots, X_{t-p}, X_{t-p} \cdot e^{-\gamma_p (X_{t-p} - z_p)^2} \right)^\top. \end{aligned}$$

Then, they derive the parameter optimization problem by computing

$$\left(\widehat{\boldsymbol{\theta}}_L, \widehat{\boldsymbol{\theta}}_N \right) = \arg \max_{\boldsymbol{\theta}_L, \boldsymbol{\theta}_N} V(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N),$$

where

$$V(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N) = \frac{1}{2} \|\mathbf{r}(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N)\|_2^2,$$

and

$$\mathbf{r}(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N) = (\mathbf{r}_m(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N), \mathbf{r}_{m-1}(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N), \dots, \mathbf{r}_{p+1}(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N)),$$

with

$$\mathbf{r}_t(\boldsymbol{\theta}_L, \boldsymbol{\theta}_N) = X_t - \boldsymbol{\Psi}(\boldsymbol{\theta}_N, \mathbf{X}_{t-1})^\top \cdot \boldsymbol{\theta}_L,$$

denoting the residuals, and m being the length of the data.

They estimate the nonlinear parameters by

$$\min_{\theta_N} \frac{1}{2} \left\| (1 - \Phi(\theta_N) \Phi(\theta_N)^\dagger) \cdot \mathbf{x} \right\|_2^2,$$

where $\Phi(\theta_N) = (\Psi(\theta_N, \mathbf{X}_{m-1}), \Psi(\theta_N, \mathbf{X}_{m-2}), \dots, \Psi(\theta_N, \mathbf{X}_p))^\top$, $\Phi(\theta_N)^\dagger$ is the Moore-Penrose inverse of $\Phi(\theta_N)$, and $\mathbf{x} = (x_m, x_{m-1}, \dots, x_{p+1})^\top$. Next, they derive the estimation of the linear parameter by

$$\theta_L = \Phi(\theta_N)^\dagger \cdot \mathbf{x}.$$

Also, Chatterjee and Das (2003) study the asymptotics of parameter estimates in conditional heteroscedastic models. They obtain the estimators by minimizing certain functionals and by solving estimation equations. They establish the consistency of these estimators and they find the asymptotic limit laws of the estimators. Next, they consider a condition under which the asymptotic law is normal. Finally, they illustrate their paper with a real data example.

Giraitis and Robinson (2001) use the Whittle estimation based on squared observations to estimate the parameters of a class of ARCH models. The authors establish the \sqrt{n} -consistency and asymptotical normality of the estimators. Many conditions are introduced for the squares to have short memory autocorrelation, by comparison with the work of Zaffaroni and d'Italia (2003), who establish the same properties on the basis of an alternative class of models with martingale difference levels and long memory autocorrelated squares.

In a class of pure GARCH and ARMA-GARCH process, Francq and Zakoian (2004) use the maximum likelihood method to estimate their parameters. The authors prove the strong consistency and asymptotic normality of the quasi-maximum likelihood estimator of the parameters of pure generalized autoregressive conditional heteroscedastic (GARCH) process, and of autoregressive moving-average models with noise sequence driven by a GARCH model. Also, Francq and Zakoian (2007) establish the asymptotic distribution of the quasi-maximum likelihood (QML) estimator for generalized autoregressive conditional heteroscedastic (GARCH) processes, when the true parameters may have zero coefficients. The authors show that this asymptotic distribution is the projection of a normal vector distribution onto a convex cone. Their results are obtained under mild conditions. The main practical implication of these results concerns the estimation of overidentified GARCH models.

Coming back to our work, recall that in the model (1.1), $\rho_0 \in \mathbb{R}^p$ denotes the true parameter and ρ_n its consistent estimator. For any $t = 1, \dots, n$, $\rho \in \mathbb{R}^p$ and $\gamma \in \mathbb{R}^{p(k+1)}$,

define

$$\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}) = \frac{X_t - T(\boldsymbol{\rho} + \boldsymbol{\gamma} \odot \omega(t), \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})}. \quad (4.2)$$

For the treatment of this chapter, we need the following additional assumptions.

(\mathcal{B}_1) Assume that the model is identifiable, that is $(\boldsymbol{\rho}_1, \boldsymbol{\gamma}_1) \neq (\boldsymbol{\rho}_2, \boldsymbol{\gamma}_2) \implies T(\boldsymbol{\rho}_1 + \boldsymbol{\gamma}_1 \odot \omega(t), \mathbf{X}_{t-1}) \neq T(\boldsymbol{\rho}_2 + \boldsymbol{\gamma}_2 \odot \omega(t), \mathbf{X}_{t-1})$, for $\boldsymbol{\rho}_1, \boldsymbol{\rho}_2 \in \mathbb{R}^p$ and $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2 \in \mathbb{R}^{p(k+1)}$.

(\mathcal{B}_2) The true parameter $\boldsymbol{\rho}_0$ has a consistent estimator $\boldsymbol{\rho}_n$ that satisfies the Bahadur representation (see, e.g., Bahadur et al. (1960)) given by:

$$n^{\frac{1}{2}}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) = n^{-\frac{1}{2}} \sum_{t=1}^n \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \aleph(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)) + o_P(1), \quad (4.3)$$

where

- $\Upsilon(x, \boldsymbol{\rho}_0) = (\Upsilon_1(x, \boldsymbol{\rho}_0), \dots, \Upsilon_p(x, \boldsymbol{\rho}_0))^\top \in \mathbb{R}^p$
- For any $j = 1, \dots, k+1$, $x \in \mathbb{R}^d$, $\exists a \geq 0$ such that $\int_{\mathbb{R}^d} \|\Upsilon(\boldsymbol{\rho}_0, x)\|^{2+a} dF_j(x) < \infty$.
- $\int_{\mathbb{R}} |\aleph(x)|^{2+a} f(x) dx < \infty$ and $\int_{\mathbb{R}} \aleph(x) f(x) dx = 0$,

(\mathcal{B}_3) For any $j \in \{1, \dots, k+1\}$, $h \in \{1, \dots, p\}$, suppose that

$$\int_{\mathbb{R}^d} \frac{\Upsilon(\boldsymbol{\rho}_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) dF_j(x) < \infty,$$

(\mathcal{B}_4) For any $i = 1, \dots, k+1$ and $j = 1, \dots, p$, there exists a ball $B(r)$ of radius r , such that

$$\max \left\{ \sup_{\boldsymbol{\rho} \in B(r)} \|\partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, x)\|_p, \sup_{\boldsymbol{\rho} \in B(r)} \left\| \partial_{\boldsymbol{\rho}} \left(\frac{\partial T}{\partial \gamma_{i,j}} \right) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, x) \right\|_p, \right. \\ \left. \sup_{\boldsymbol{\rho} \in B(r)} \left\| \partial_{\boldsymbol{\rho}}^2 \left(\frac{\partial T}{\partial \gamma_{i,j}} \right) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, x) \right\|_p \right\} \leq \chi(x),$$

(\mathcal{B}_5) Assume that, for $j = 1, 2, 3$ and $a, b \in \{0, 1, 2, 3\}$,

$$\lambda_{a,b}^{(j)} = \int_{\mathbb{R}^d} \frac{\nu^a(x) \chi^b(x)}{V^j(x)} dF_j(x) < \infty,$$

For any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, under H_0 , the central sequence with true parameter is denoted by $\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ and its estimated version by $\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$.

Remark 4.2.1. [1] One can find in literature, numbers of models with functions $T(\boldsymbol{\rho}, \boldsymbol{\gamma}, x)$ satisfying (\mathcal{B}_4) and (\mathcal{B}_5) .

[2] Assumption (\mathcal{B}_2) helps to study the distribution of the test statistic. It has been used before in Ngatchou-Wandji and Ltaifa (2021). It is satisfied by least-squares and likelihood type estimators for some usual models within (1.1).

Remark 4.2.2. Assume that assumption (\mathcal{A}_1) holds. Then denote by $\widehat{\rho}_{j,h}$ a consistent estimator of $\rho_{0,h} + \beta_{j,h}/\sqrt{n}$ on the basis of observations within $[\tau_{j-1}, \tau_j]$. Then one can consider $\widehat{\beta}_{j,h} = \sqrt{n}(\widehat{\rho}_{j,h} - \widehat{\rho}_{0,h})$ as an estimator of $\beta_{j,h}$, where $\widehat{\rho}_{0,h}$ is the estimation of $\rho_{0,h}$ on the basis of observations in $[1, \tau_1]$.

Proposition 4.2.1. Under the assumption (\mathcal{B}_2) , (\mathcal{A}_1) - (\mathcal{A}_{10}) we have:

i. Under H_0 :

$$\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma).$$

ii. Under $H_{\boldsymbol{\beta}}^{(n)}$, we have:

$$\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(\mathcal{C}, \Sigma),$$

where

$$\mathcal{C} = \int_{\mathbb{R}} \aleph(x) \phi_f(x) f(x) dx \sum_{j=1}^n \alpha_j \sum_{h=1}^p \beta_{j,h} \int_{\mathbb{R}^d} \frac{\Upsilon(\boldsymbol{\rho}_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, x) dF_j(x) \in \mathbb{R}^p,$$

and

$$\Sigma = \int_{\mathbb{R}} \aleph^2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \Upsilon(\boldsymbol{\rho}_0, x) \Upsilon^\top(\boldsymbol{\rho}_0, x) dF_j(x) \in \mathcal{M}_p(\mathbb{R}).$$

Proof of (i): From the Bahadur representation (4.3), we consider the sequence:

$$R_{n,j} = n^{-\frac{1}{2}} \sum_{t=1}^j \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)], \quad j = 1, \dots, n.$$

It is easy to see that $R_{n,j} = \sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0)$ if $j = n$. We use the sequence $R_{n,j}$ to prove that $\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0)$ converge to a normal law by applying Corollary 2.3.2 of Hall and Heyde (2014) that represents the modified version of the Central limit theorem of martingales.

We take assistance from Ngatchou-Wandji (2008) and Ngatchou-Wandji (2005) to consider

this new sequence:

Let $u \in \mathbb{R}^p$ such that $u = (u_1, \dots, u_p)^\top$. We observe that:

$$u^\top R_{n,j} = \frac{1}{\sqrt{n}} \sum_{t=1}^j y_t(s) \in \mathbb{R},$$

where

$$y_t(u) = u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)].$$

It is easy to see that $u^\top R_{n,j}$ is centered for every $j = 1, \dots, n$. In this part, we follow the same steps as in the proof of Theorem 2.

• We prove that $\{(u^\top R_{n,j}, \mathcal{F}_j), j = 1, \dots, n\}$ is a martingale sequence. Let $q, w \in \mathbb{Z}$ such that $q < w$, we have

$$\begin{aligned} \mathbb{E}\{u^\top R_{n,w} | \mathcal{F}_q\} &= \mathbb{E}\{u^\top R_{n,q} | \mathcal{F}_q\} + \frac{1}{\sqrt{n}} \sum_{t=q+1}^w \mathbb{E}\{y_t(s) | \mathcal{F}_q\} \\ &= u^\top R_{n,q} + \sum_{t=q+1}^w \mathbb{E}\{u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] | \mathcal{F}_q\} \\ &= u^\top R_{n,q} + \sum_{t=q+1}^w \mathbb{E}\{u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1})\} \mathbb{E}\{\aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} \end{aligned}$$

Since $\mathbb{E}\{\aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} = 0$ for every $t \in \mathbb{N}$, it follows that

$$\mathbb{E}\{u^\top R_{n,w} | \mathcal{F}_q\} = u^\top R_{n,q}.$$

□

• We check now the first condition of Corollary 2.3.2 of Hall and Heyde (2014). Since ε_t is independent of \mathcal{F}_{t-1} for $t = 1, \dots, n$, we can write

$$\begin{aligned} \sum_{t=1}^n \mathbb{E} \left\{ \left[n^{-\frac{1}{2}} y_t(u) \right]^2 | \mathcal{F}_{t-1} \right\} &= \sum_{t=1}^n n^{-1} \mathbb{E} \left\{ \left[u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right]^2 | \mathcal{F}_{t-1} \right\} \\ &= \sum_{t=1}^n n^{-1} \left[u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \right]^2 \mathbb{E} \left\{ \aleph^2[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\ &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left[u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \right]^2 \mathbb{E} \left\{ \aleph^2[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\}. \end{aligned}$$

By the assumptions (\mathcal{B}_2) , (\mathcal{B}_3) and the ergodic theorem, for $j = 1, \dots, k+1$, we can

write

$$\begin{aligned} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} [u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1})]^2 \mathbb{E} \{ \aleph^2[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \} \xrightarrow{a.s.} \int_{\mathbb{R}^d} [u^\top \Upsilon(\boldsymbol{\rho}_0, x)]^2 dF_j(x) \\ \times \int_{\mathbb{R}} \aleph^2(x) f(x) dx < \infty. \end{aligned}$$

Then,

$$\sum_{t=1}^n \mathbb{E} \left\{ \left[\frac{1}{\sqrt{n}} y_t(u) \right]^2 \middle| \mathcal{F}_{t-1} \right\} \xrightarrow[n \rightarrow +\infty]{a.s.} s,$$

where

$$s = \int_{\mathbb{R}} \aleph^2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} [u^\top \Upsilon(\boldsymbol{\rho}_0, x)]^2 dF_j(x).$$

□

• Finally, we check the Linderberg condition, that is the second condition of Corollary 2.3.2 of Hall and Heyde (2014). In this purpose, we will prove that, as $n \rightarrow +\infty$,

$$\sum_{t=1}^n \mathbb{E} \left\{ \left[n^{-\frac{1}{2}} y_t(u) \right]^2 \mathbb{1}_{|n^{-\frac{1}{2}} y_t(u)| > \epsilon} \middle| \mathcal{F}_{t-1} \right\} \xrightarrow{a.s.} 0.$$

Let $\epsilon > 0$, by Hölder inequality, and Markov inequalities, we can write

$$\begin{aligned} & \sum_{t=1}^n \mathbb{E} \left\{ \left[n^{-\frac{1}{2}} y_t(u) \right]^2 \mathbb{1}_{|n^{-\frac{1}{2}} y_t(u)| > \epsilon} \middle| \mathcal{F}_{t-1} \right\} \\ & \leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ \left[n^{-\frac{1}{2}} y_t(u) \right]^3 \middle| \mathcal{F}_{t-1} \right\} \mathbb{E}^{\frac{1}{3}} \left\{ \mathbb{1}_{|n^{-\frac{1}{2}} y_t(u)| > \epsilon} \middle| \mathcal{F}_{t-1} \right\} \\ & \leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ \left[n^{-\frac{1}{2}} y_t(u) \right]^3 \middle| \mathcal{F}_{t-1} \right\} \mathbb{P}^{\frac{1}{3}} \left\{ \left| n^{-\frac{1}{2}} y_t(u) \right| > \epsilon \middle| \mathcal{F}_{t-1} \right\} \\ & \leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ \left[n^{-\frac{1}{2}} y_t(u) \right]^3 \middle| \mathcal{F}_{t-1} \right\} \frac{\mathbb{E}^{\frac{1}{3}} \left\{ \left| n^{-\frac{1}{2}} y_t(u) \right|^3 > \epsilon \middle| \mathcal{F}_{t-1} \right\}}{\sqrt[3]{\epsilon}} \\ & \leq \sum_{t=1}^n \frac{1}{\sqrt[3]{\epsilon} \sqrt[2]{n^3}} \mathbb{E}^{\frac{2}{3}} \left\{ [y_t(u)]^3 \middle| \mathcal{F}_{t-1} \right\} \mathbb{E}^{\frac{1}{3}} \left\{ |y_t(u)|^3 > \epsilon \middle| \mathcal{F}_{t-1} \right\} \\ & \leq \sum_{t=1}^n \frac{1}{\sqrt[3]{\epsilon} \sqrt[2]{n^3}} \|u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1})\|^3 \mathbb{E} \{ |\aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|^3 \} \\ & \leq \frac{1}{\sqrt[3]{\epsilon} \sqrt{n}} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \|u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1})\|^3 \mathbb{E} \{ |\aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|^3 \}. \end{aligned}$$

By piece-wise stationary and ergodic theorem, for $j = 1, \dots, k + 1$, we obtain

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \frac{1}{\sqrt{n}} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \|u^\top \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1})\|^3 \mathbb{E} \{ |\aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|^3 \} \\ &= 0 \times \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \|u^\top \Upsilon(\boldsymbol{\rho}_0, x)\|^3 dF_j(x) \times \int_{\mathbb{R}} |\aleph(x)|^3 f(x) dx \\ &= 0. \end{aligned}$$

□

Then, using Corollary 2.3.2 of Hall and Heyde (2014), we conclude that, under H_0 , we have

$$u^\top \sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, s).$$

Which implies that, under H_0 and as $n \rightarrow +\infty$,

$$\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma)$$

where Σ is the covariance matrix defined as

$$\Sigma = \int_{\mathbb{R}} \aleph^2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \Upsilon(\boldsymbol{\rho}_0, x) \Upsilon^\top(\boldsymbol{\rho}_0, x) dF_j(x) \in \mathcal{M}_p(\mathbb{R}).$$

□

Proof of (ii): We recall that under H_0 , as $n \rightarrow +\infty$,

$$\begin{aligned} & \sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{a.s.} \mathcal{N}(0, \Sigma) \\ & \Theta_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{2}, \eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\right), \end{aligned}$$

where

$$\begin{aligned} \eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) \\ \eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) &= I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) dF_j(x). \end{aligned}$$

We consider the sequence $Q_n = \sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0)$. By Le Cam's third Lemma, under H_0 , as $n \rightarrow +\infty$,

$$\begin{pmatrix} Q_n \\ \Theta_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix} 0 \\ -\frac{\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{2} \end{pmatrix}, \xi\right),$$

where

$$\begin{aligned}
 \xi &= \lim_{n \rightarrow +\infty} \begin{pmatrix} \text{Var}(Q_n) & \text{Cov}(Q_n, \Theta_n) \\ \text{Cov}(Q_n, \Theta_n) & \text{Var}(\Theta_n) \end{pmatrix}, \\
 \text{Cov}(Q_n, \Theta_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})) &= \text{Cov} \left(Q_n, \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) - \frac{\eta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{2} \right) \\
 &= \text{Cov}(Q_n, \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})) \\
 &= \mathbb{E} \{ Q_n \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \} - \mathbb{E} \{ Q_n \} \mathbb{E} \{ \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \} \\
 &= \mathbb{E} \left\{ \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{Q_n}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\
 &\quad - \mathbb{E}(Q_n) \frac{1}{\sqrt{n}} \sum_{t=1}^n \mathbb{E} \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \\
 &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \left[\mathbb{E} \left\{ \frac{Q_n}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \right. \\
 &\quad \left. - \mathbb{E}(Q_n) \mathbb{E} \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \right\} \mathbb{E} \{ \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \} \right].
 \end{aligned}$$

Since $\mathbb{E}\{\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} = 0$ and $\lim_{n \rightarrow +\infty} \mathbb{E}(Q_n) = 0$, then

$$\begin{aligned}
 &\text{Cov}(Q_n, \Theta_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})) \\
 &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \left[\mathbb{E} \left\{ \frac{Q_n}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right\} \right. \\
 &= \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left\{ \sum_{m=1}^n \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{m-1}) \aleph(\varepsilon_m(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)) \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]}{V(\mathbf{X}_{t-1})} \right\} \\
 &= \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left\{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \aleph(\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)) \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]}{V(\mathbf{X}_{t-1})} \right\} \\
 &\quad + \frac{1}{n} \sum_{\substack{t,m=1 \\ t \neq m}}^n \mathbb{E} \left\{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{m-1}) \aleph(\varepsilon_m(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)) \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]}{V(\mathbf{X}_{t-1})} \right\}.
 \end{aligned}$$

Since ε_t is independent of \mathcal{F}_{t-1} and $\mathbb{E}\{\aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} = 0$ and $\mathbb{E}\{\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} = 0$. Then, it follows that, for $t \neq m$, we have

$$\begin{aligned}
 &\frac{1}{n} \sum_{\substack{t,m=1 \\ t \neq m}}^n \mathbb{E} \left\{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{m-1}) \aleph[\varepsilon_m(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]}{V(\mathbf{X}_{t-1})} \right\} \\
 &= \frac{1}{n} \sum_{\substack{t,m=1 \\ t \neq m}}^n \mathbb{E} \{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{m-1}) \} \mathbb{E} \{ \aleph[\varepsilon_m(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \} \mathbb{E} \left\{ \frac{\boldsymbol{\beta}^\top N(\boldsymbol{\gamma}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right\} \mathbb{E} \{ \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \} \\
 &= 0.
 \end{aligned}$$

With this,

$$\begin{aligned}
 & \text{Cov}(Q_n, \Theta_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})) \\
 &= \mathbb{E} \left\{ \aleph[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \right\} \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left\{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \frac{\boldsymbol{\beta}^\top N(\gamma_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right\} \\
 &= \int_{\mathbb{R}} \aleph(x) \phi_f(x) f(x) dx \frac{1}{n} \sum_{t=1}^n \mathbb{E} \left\{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \frac{\boldsymbol{\beta}^\top N(\gamma_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right\} \\
 &= \int_{\mathbb{R}} \aleph(x) \phi_f(x) f(x) dx \underbrace{\sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \mathbb{E} \left\{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \frac{\boldsymbol{\beta}^\top N(\gamma_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right\}}_{V_n}.
 \end{aligned}$$

By stationarity and ergodic theorem, we have

$$\begin{aligned}
 V_n &= \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \omega_j \mathbb{E} \left\{ \Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1}) \frac{\boldsymbol{\beta}_j^\top \partial_j [T(\gamma_0, \mathbf{X}_{t-1})]}{V(\mathbf{X}_{t-1})} \right\} \\
 &= \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \omega_j \mathbb{E} \left\{ \frac{\Upsilon(\boldsymbol{\rho}_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \sum_{h=1}^p \beta_{j,h} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \right\} \\
 &\xrightarrow{a.s.} \omega_j \sum_{h=1}^p \beta_{j,h} \int_{\mathbb{R}^d} \frac{\Upsilon(\boldsymbol{\rho}_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, x) dF_j(x) < \infty.
 \end{aligned}$$

Then,

$$\text{Cov}(Q_n, \Theta_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})) \xrightarrow{n \rightarrow +\infty} \mathcal{C},$$

where

$$\mathcal{C} = \int_{\mathbb{R}} \aleph(x) \phi_f(x) f(x) dx \sum_{j=1}^n \alpha_j \omega_j \sum_{h=1}^p \beta_{j,h} \int_{\mathbb{R}^d} \frac{\Upsilon(\boldsymbol{\rho}_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, x) dF_j(x) < \infty.$$

Then, under H_0 , we have

$$\begin{pmatrix} Q_n \\ \Theta_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N} \left(\begin{pmatrix} 0 \\ -\frac{\eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta})}{2} \end{pmatrix}, \begin{pmatrix} \Sigma & \mathcal{C}^\top \\ \mathcal{C} & \eta(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) \end{pmatrix} \right).$$

From this result and Le Cam's third lemma, under $H_1^{(n)}$, as n tends to $+\infty$,

$$\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) = Q_n \xrightarrow{\mathcal{D}} \mathcal{N}(\mathcal{C}, \Sigma).$$

□

4.2.2 Asymptotic behavior of the new test statistic

Recall from Chapter 3 that the test statistic is, in this case,

$$\mathcal{T}_n(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})},$$

and $\boldsymbol{\rho}$ must be substituted by its estimator. Here, the central sequence is denoted by

$$\Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}, \boldsymbol{\beta}) = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma})] \right\}.$$

Recall that, for every $j = 1, \dots, k+1$, $n_j(n)$ is the number of observations between the instants τ_j and τ_{j-1} and F_j be the common cumulative distribution function of \mathbf{X}_t 's with components within these instants. Also, recall that $\eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)$ has the expression

$$\eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) dF_j(x),$$

which depends on F_j and $\boldsymbol{\rho}$.

An estimated version of $\eta_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)$, denoted by $\widehat{\eta}_{j,2}^{(h,m)}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)$, is given by

$$\widehat{\eta}_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, x) d\widehat{F}_j(x),$$

where \widehat{F}_j is the empirical distribution function of the observations between τ_j and τ_{j-1} .

We can write

$$\widehat{\eta}_{j,b}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = \frac{I(f)}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^b(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}).$$

Then, we denote by

$$\begin{aligned} \widehat{\eta}_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \sum_{j=1}^{k+1} \widehat{\alpha}_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \widehat{\eta}_{j,2}^{(h,m)}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0), \\ \widehat{\pi}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \sqrt{\widehat{\eta}_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}, \end{aligned}$$

where, for $j = 1, \dots, k+1$, the $\widehat{\alpha}_j$ is the estimator of $\alpha_j = \lim_{n \rightarrow +\infty} n_j(n)/n$.

The proof of the convergence of the central sequence is given by the following Proposition.

Proposition 4.2.2. *Under the assumptions (\mathcal{A}_1) - (\mathcal{A}_{10}) and (\mathcal{B}_1) - (\mathcal{B}_5) , for any $n \geq 0$, for any consistent and asymptotic normal estimator $\boldsymbol{\rho}_n$ of $\boldsymbol{\rho}_0$, we have, for any sequence of positive integers $s(n)$ such that $n/s(n) \rightarrow 0$, as $n \rightarrow +\infty$.*

i.

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1), \quad (4.4)$$

ii.

$$\widehat{\pi}_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \longrightarrow \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}). \quad (4.5)$$

Proof of (i): We aim to prove the convergence of the central sequence (3.9) to its estimated version in order to verify that the test still be optimal when we replace the parameter by its estimator.

In this purpose, we must prove that

$$\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) - \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow[n \rightarrow +\infty]{\mathbb{P}} 0.$$

Recalling (3.2), we can write

$$\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}) = \frac{X_t - T(\boldsymbol{\rho} + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(t), \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})}.$$

From which one has,

$$\begin{aligned} \partial_{\boldsymbol{\rho}} \varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}) &= -\frac{1}{V(\mathbf{X}_{t-1})} \partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho} + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(t), \mathbf{X}_{t-1}) \\ &= -\frac{1}{V(\mathbf{X}_{t-1})} \partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}), \end{aligned} \quad (4.6)$$

and

$$\partial_{\boldsymbol{\rho}}^2 \varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}) = -\frac{1}{V(\mathbf{X}_{t-1})} \partial_{\boldsymbol{\rho}}^2 T(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}). \quad (4.7)$$

Using a first-order Taylor expansion of $\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)$ in a neighborhood of $\boldsymbol{\rho}_0$, for some $\tilde{\boldsymbol{\rho}}_n$ lying between $\boldsymbol{\rho}_n$ and $\boldsymbol{\rho}_0$, from a simple calculations, we obtain

$$\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0) = \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) - \frac{(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0)}{V(\mathbf{X}_{t-1})} \partial_{\boldsymbol{\rho}} T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}). \quad (4.8)$$

For any $\boldsymbol{\rho} \in \mathbb{R}^p$ and $\boldsymbol{\gamma}, \boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, we define

$$\Theta_n(\boldsymbol{\rho}, \boldsymbol{\gamma}, \boldsymbol{\beta}) = \log A_n = \sum_{t=1}^n \log \left\{ f \left[\varepsilon_t \left(\boldsymbol{\rho}, \boldsymbol{\gamma} + \frac{\boldsymbol{\beta}}{\sqrt{n}} \right) \right] \right\} - \log \{f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma})]\} + o_P(1).$$

From this definition, we can conclude that the log-likelihood ratio test of H_0 against $H_{\boldsymbol{\beta}}^{(n)}$ is $\Theta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$. For $\tilde{\boldsymbol{\rho}}_n$ lying between $\boldsymbol{\rho}_n$ and $\boldsymbol{\rho}_0$, we write a second-order Taylor expansion of $\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ around $\boldsymbol{\rho}_n$ and obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \\ &\quad + \frac{1}{2} (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n). \end{aligned} \quad (4.9)$$

We wish to prove that, under H_0 , as $n \rightarrow +\infty$,

$$(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = o_P(1) \quad (4.10)$$

$$\frac{1}{2}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n) = o_P(1). \quad (4.11)$$

We start by (4.11), and we observe that

$$\begin{aligned} & \left| \frac{1}{2}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n) \right| \\ & \leq \|\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n\|_p \left\| \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \right\|_p \|\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n\|_p \\ & \leq \left\| \sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n) \right\|_p \times \frac{1}{\sqrt{n}} \left\| \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \right\|_p \|\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n\|_p. \end{aligned}$$

Using Proposition 4.2.1, under H_0 and as $n \rightarrow +\infty$, we have

$$\sqrt{n}(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma).$$

Also, as n tends to infinity, we have

$$(\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \xrightarrow{\mathbb{P}} 0.$$

Then, to prove (4.11), it suffices to show that $(1/\sqrt{n}) \left\| \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \right\|_p$ tends in probability to some positive random variable.

Recall from (3.9), that

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}, \boldsymbol{\beta}) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma})] \right\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \frac{\partial T}{\partial \gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma})]. \end{aligned}$$

In order to simplify the notations, let $T^{\gamma_{m,h}} = \partial T / \partial \gamma_{m,h}$. We have

$$\begin{aligned} \frac{1}{\sqrt{n}} \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \{ \partial_{\boldsymbol{\rho}} T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ & \quad + T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\boldsymbol{\rho}} \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \} \\ &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\rho}} (T^{\gamma_{m,h}})(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ & \quad + \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\boldsymbol{\rho}} \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)]. \end{aligned}$$

Thus,

$$\begin{aligned} \partial_{\rho} \Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\rho} (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ &\quad - \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}} (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\rho} T(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \end{aligned} \quad (4.12)$$

For any $\boldsymbol{\rho} \in \mathbb{R}^p$, we have

$$\begin{aligned} \partial_{\rho}^2 \Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \partial_{\rho} \left(\frac{\partial \Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\partial \rho} \right) \\ &= \partial_{\rho} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\rho} (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \right) \\ &\quad - \partial_{\rho} \left(\frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}} (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\rho} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \right) \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\rho}^2 (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ &\quad + \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\rho} (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) (\partial_{\rho} \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)])^{\top} \\ &\quad - \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_{\rho} (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})) (\partial_{\rho} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}))^{\top} \phi'_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ &\quad - \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}} (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\rho}^2 T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ &\quad - \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^3(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}} (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) (\partial_{\rho} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})) (\partial_{\rho} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}))^{\top} \\ &\quad \times \phi''_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\rho}^2 (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ &\quad - \frac{2}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_{\rho} (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})) (\partial_{\rho} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}))^{\top} \phi'_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\ &\quad - \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}} (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \partial_{\rho}^2 T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \end{aligned}$$

$$\begin{aligned}
 & + \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V^3(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) (\partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})) (\partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}))^\top \\
 & \times \phi_f''[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)].
 \end{aligned}$$

Multiplying by $1/\sqrt{n}$ and replacing $\boldsymbol{\rho}$ by $\tilde{\boldsymbol{\rho}}_n$, we obtain

$$\begin{aligned}
 & \frac{1}{\sqrt{n}} \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \\
 & = \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\rho}}^2 (T^{\gamma_{m,h}}) (\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \\
 & - \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\rho}} (T^{\gamma_{m,h}}) (\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \partial_{\boldsymbol{\rho}}^\top T(\tilde{\boldsymbol{\rho}}_n, \mathbf{X}_{t-1}) \phi_f'[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \\
 & - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \partial_{\boldsymbol{\rho}}^2 T(\tilde{\boldsymbol{\rho}}_n, \mathbf{X}_{t-1}) \phi_f'[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \\
 & + \frac{1}{n} \sum_{t=1}^n \frac{1}{V^3(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \partial_{\boldsymbol{\rho}} T(\tilde{\boldsymbol{\rho}}_n, \mathbf{X}_{t-1}) \partial_{\boldsymbol{\rho}}^\top T(\tilde{\boldsymbol{\rho}}_n, \mathbf{X}_{t-1}) \phi_f''[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \\
 & = \Delta_{1,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + \Delta_{2,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + \Delta_{3,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + \Delta_{4,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}),
 \end{aligned}$$

where

$$\begin{aligned}
 \Delta_{1,n}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) & = \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\rho}}^2 (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\
 \Delta_{2,n}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) & = -\frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\rho}} (T^{\gamma_{m,h}}) (\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \partial_{\boldsymbol{\rho}}^\top T(\boldsymbol{\rho}, \mathbf{X}_{t-1}) \phi_f'[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\
 \Delta_{3,n}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) & = -\frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \partial_{\boldsymbol{\rho}}^2 T(\boldsymbol{\rho}, \mathbf{X}_{t-1}) \phi_f'[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)] \\
 \Delta_{4,n}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) & = \frac{1}{n} \sum_{t=1}^n \frac{1}{V^3(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho}, \mathbf{X}_{t-1}) \partial_{\boldsymbol{\rho}}^\top T(\boldsymbol{\rho}, \mathbf{X}_{t-1}) \\
 & \quad \times \phi_f''[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}_0)].
 \end{aligned}$$

Recall that, for $i = 1, \dots, 4$, we wish to prove that $\|\Delta_{i,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p$ tends to a finite positive number as $n \rightarrow +\infty$.

Study of the asymptotic behavior of $\Delta_{1,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$.

Recall that

$$\Delta_{1,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\rho}}^2 T^{\gamma_{m,h}}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)].$$

We have

$$\|\Delta_{1,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \|\partial_{\boldsymbol{\rho}}^2 T^{\gamma_{m,h}}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p |\phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]|.$$

By the condition (\mathcal{B}_4) , we can write

$$\begin{aligned} \|\Delta_{1,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p &\leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \chi(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]| \\ &\leq \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

Adding and subtracting appropriate terms, we obtain

$$\begin{aligned} \|\Delta_{1,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p &\leq \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &\quad + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

We know that, by a second-order Taylor expansion, for some $\tilde{\boldsymbol{\rho}}_n$ lying between $\boldsymbol{\rho}_0$ and $\tilde{\boldsymbol{\rho}}_n$,

$$\begin{aligned} &\phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\ &= (\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0)^\top \partial_{\boldsymbol{\rho}} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] + \frac{1}{2} (\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0)^\top \partial_{\boldsymbol{\rho}}^2 \phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] (\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0). \end{aligned}$$

Using (4.6) and (4.7), we can write

$$\begin{aligned} \partial_{\boldsymbol{\rho}} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] &= -\frac{1}{|V(\mathbf{X}_{t-1})|} \partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)], \\ \partial_{\boldsymbol{\rho}}^2 \phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] &= \frac{1}{|V(\mathbf{X}_{t-1})|^2} \left(\partial_{\boldsymbol{\rho}} T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \right) \left(\partial_{\boldsymbol{\rho}} T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \right)^\top \phi''_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \\ &\quad - \frac{1}{|V(\mathbf{X}_{t-1})|} \partial_{\boldsymbol{\rho}}^2 T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]. \end{aligned}$$

Using the assumption (\mathcal{B}_4) and (\mathcal{A}_6) , we obtain

$$\begin{aligned} \|\partial_{\boldsymbol{\rho}} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\|_p &\leq \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|, \\ \left\| \partial_{\boldsymbol{\rho}}^2 \phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \right\|_p &\leq c_\phi \frac{\chi^2(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^2} + \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} \left| \phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \right|. \end{aligned}$$

By the assumption (\mathcal{A}_6) , we can write

$$\left| \phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \right| \leq c_\phi \left| \varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0) - \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) \right| + |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.$$

From on the inequality (3.6), we can write

$$\left| \varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0) - \varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) \right| \leq \frac{\left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|}.$$

Then

$$\left| \phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \right| \leq c_\phi \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} + |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \quad (4.13)$$

Consequently,

$$\begin{aligned} \left\| \partial_{\boldsymbol{\rho}}^2 \phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \right\|_p &\leq c_\phi \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} + \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &\quad + c_\phi \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \\ &= c_\phi \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \left(1 + \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \right) + \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

Hence,

$$\begin{aligned} |\phi_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\leq \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &\quad + \frac{1}{2} \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p^2 c_\phi \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \left(1 + \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \right) \\ &\quad + \frac{1}{2} \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p^2 \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

Finally,

$$\begin{aligned} \|\Delta_{1,n}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p &\leq \frac{\left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &\quad + \frac{c_\phi \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p^2 \left(1 + \left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \right) \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{2n} \sum_{t=1}^n \frac{\chi^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ &\quad + \frac{\left\| \tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0 \right\|_p \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{2n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &\quad + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

Applying the ergodic theorem on each term after the summation with respect to t , as

$n \rightarrow +\infty$, since ε_t is independent of \mathcal{F}_{t-1} and in view of Remark (3.4.4), we obtain

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \mathbb{E} \{ |\phi'_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \} \int_{\mathbb{R}^d} \frac{\chi^2(x)}{V^2(x)} dF_j(x) \\ &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \lambda_{0,2}^{(2)} < \infty. \end{aligned}$$

In the same lines, we obtain

$$\frac{1}{n} \sum_{t=1}^n \frac{\chi^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \lambda_{0,3}^{(3)} < \infty,$$

and,

$$\frac{1}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \lambda_{0,2}^{(2)} < \infty,$$

In view of Remark (3.4.4),

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \mathbb{E} \{ |\phi_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \} \int_{\mathbb{R}^d} \frac{\chi(x)}{V(x)} dF_j(x) \\ &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \lambda_{0,1}^{(1)} \int_{\mathbb{R}^d} |\phi_f(x)| f(x) dx < \infty. \end{aligned}$$

Remark 4.2.3. As $n \rightarrow +\infty$, in probability,

$$\|\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0\|_p \leq \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p \rightarrow 0.$$

From Remark 4.2.3, since all the other terms are bounded, we can conclude that, as $n \rightarrow +\infty$, in probability, we have

$$\|\Delta_{1,n}(\boldsymbol{\rho}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p \rightarrow 0.$$

Study of the asymptotic behavior of $\Delta_{2,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$.

Recalling (4.13), we have

$$|\phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]| \leq c_\phi \|\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0\|_p \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| + |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.$$

Thus,

$$\begin{aligned}
 |||\Delta_{2,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p &\leq \frac{2}{n} \sum_{t=1}^n \frac{|\phi'_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)]|}{|V(\mathbf{X}_{t-1})|^2} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \|\partial_{\rho} T^{\gamma_{m,h}}(\tilde{\rho}_n, \gamma_0, \mathbf{X}_{t-1})\|_p \\
 &\quad \times \left\| \partial_{\rho} T(\tilde{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \right\|_p \\
 &\leq \frac{2}{n} \sum_{t=1}^n \frac{|\phi'_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)]|}{|V(\mathbf{X}_{t-1})|^2} |\chi(\mathbf{X}_{t-1})|^2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \\
 &\leq \frac{2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \left| \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^2 |\phi'_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)]| \\
 &\leq \frac{2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| c_{\phi} \|\tilde{\rho}_n - \rho_0\|_p}{n} \sum_{t=1}^n \left| \frac{\chi^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})} \right| \\
 &\quad + \frac{2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \left| \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^2 |\phi'_f[\varepsilon_t(\rho_0, \gamma_0)]|.
 \end{aligned}$$

By the ergodic theorem and in view of the assumption (\mathcal{B}_4) , we have

$$\begin{aligned}
 \frac{1}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})} &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\chi^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})} \\
 &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\chi^2(x) \nu(x)}{V^3(x)} dF_j(x) \\
 &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \lambda_{1,2}^{(3)} < \infty.
 \end{aligned}$$

Also, since ε_t is independent of \mathcal{F}_{t-1} , in view of (\mathcal{A}_7) ,

$$\begin{aligned}
 \frac{1}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\rho_0, \gamma_0)]| &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\rho_0, \gamma_0)]| \\
 &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\chi(x)}{V(x)} dF_j(x) \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx < \infty.
 \end{aligned}$$

Then, as $n \rightarrow +\infty$, $|||\Delta_{2,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p$ tends, in probability, to a positive real number. It results that, as $n \rightarrow +\infty$, $|||\Delta_{2,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p$ tends in probability to a positive finite number.

Study of the asymptotic behavior of $\Delta_{3,n}(\tilde{\rho}_n, \gamma_0, \beta)$.

Recalling that

$$\begin{aligned} \Delta_{3,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= -\frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \partial_{\boldsymbol{\rho}}^2 T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \\ &\quad \times \phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)], \end{aligned}$$

and using $(\mathcal{B}_4), (\mathcal{A}_9)$ and (4.13), we have

$$\begin{aligned} & \|\Delta_{3,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p \\ & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| |T^{\gamma_{m,h}}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})| \|\partial_{\boldsymbol{\rho}}^2 T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1})\|_p |\phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]| \\ & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \nu(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1}) |\phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]| \\ & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \nu(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1}) \left[c_\phi \|\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0\|_p \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| + |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \right] \\ & \leq \frac{c_\phi \|\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0\|_p \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ & \quad + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

By the ergodic theorem,

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} &= \lim_{n \rightarrow +\infty} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\nu^2(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ &= \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\nu^2(x) \chi(x)}{|V(x)|^3} dF_j(x) \\ &= \sum_{j=1}^{k+1} \alpha_j \lambda_{2,1}^{(3)} < \infty. \end{aligned}$$

Again, since ε_t is independent of \mathcal{F}_{t-1} , in view of (\mathcal{A}_7) ,

$$\begin{aligned} & \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^n \frac{\nu(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &= \lim_{n \rightarrow +\infty} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\nu(\mathbf{X}_{t-1}) \chi(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &= \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\nu(x) \chi(x)}{V^2(x)} dF_j(x) \mathbb{E} \{ |\phi'_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \} \end{aligned}$$

$$= \sum_{j=1}^{k+1} \alpha_j \lambda_{1,1}^{(2)} \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx < \infty.$$

Thus, using (4.2.3), as $n \rightarrow +\infty$, $|||\Delta_{3,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p$ tends in probability to a finite positive real number.

It results that, as $n \rightarrow +\infty$, $|||\Delta_{3,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p$ tends in probability to a finite positive number.

Study of the asymptotic behavior of $\Delta_{4,n}(\tilde{\rho}_n, \gamma_0, \beta)$.

Recalling that

$$\begin{aligned} \Delta_{4,n}(\tilde{\rho}_n, \gamma_0, \beta) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\tilde{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \\ &\quad \times (\partial_{\rho} T(\tilde{\rho}_n, \gamma_0, \mathbf{X}_{t-1})) (\partial_{\rho} T(\tilde{\rho}_n, \gamma_0, \mathbf{X}_{t-1}))^{\top} \phi''_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)]. \end{aligned}$$

By assumption (\mathcal{A}_6) ,

$$\begin{aligned} |||\Delta_{4,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p &\leq \frac{c_{\phi}}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| |T^{\gamma_{m,h}}(\tilde{\rho}_n, \gamma_0, \mathbf{X}_{t-1})| \\ &\quad \times \|\partial_{\rho} T(\tilde{\rho}_n, \gamma_0, \mathbf{X}_{t-1})\|_p^2. \end{aligned}$$

By assumptions (\mathcal{B}_4) and (\mathcal{A}_9) , we obtain

$$|||\Delta_{4,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p \leq \frac{c_{\phi} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})}.$$

The ergodic theorem yields

$$\begin{aligned} \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} &= \lim_{n \rightarrow +\infty} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\chi^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \\ &= \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}} \frac{\chi^2(x) \nu(x)}{V^2(x)} dF_j(x) \\ &= \sum_{j=1}^{k+1} \alpha_j \lambda_{1,2}^{(2)} < \infty. \end{aligned}$$

Then, as $n \rightarrow +\infty$, $|||\Delta_{4,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p$ tends in probability to a finite positive number.

Now, we proved that, for $i = 1, \dots, 4$, $|||\Delta_{i,n}(\tilde{\rho}_n, \gamma_0, \beta)|||_p$ is bounded.

We conclude that, as $n \rightarrow +\infty$, $(1/\sqrt{n}) |||\partial_{\rho}^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta)|||_p$ tends in probability to a

finite positive real number, denoted by b .

Recall from (4.11) that, as $n \rightarrow +\infty$, we have

$$\begin{aligned} & \frac{1}{2} |(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_\rho^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \gamma_0, \boldsymbol{\beta})(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)| \\ & \leq \frac{1}{2} \|\sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)\|_p \left\| \left\| \frac{1}{\sqrt{n}} \partial_\rho^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \gamma_0, \boldsymbol{\beta}) \right\| \right\|_p \|\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n\|_p \\ & \leq \frac{1}{2} \|\sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)\|_p \|\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n\|_p \times b, \end{aligned}$$

since under H_0 , as $n \rightarrow +\infty$, we have

$$\sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n) \rightarrow \mathcal{N}(0, \Sigma),$$

by Remark (4.2.3), we conclude that

$$\frac{1}{2} (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_\rho^2 \Pi_n(\tilde{\boldsymbol{\rho}}_n, \gamma_0, \boldsymbol{\beta})(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n) = o_P(1). \quad (4.14)$$

□

Treating (4.10).

From (4.14), we can write

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_\rho \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + o_P(1). \quad (4.15)$$

Now, we aim to prove that

$$(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)^\top \partial_\rho \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = o_P(1).$$

Now, as $n \rightarrow +\infty$, adding and subtracting appropriate terms, we can write

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n + \boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_{s(n)})^\top \partial_\rho \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + o_P(1) \\ &= \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + (\boldsymbol{\rho}_0 - \boldsymbol{\rho}_{s(n)})^\top \partial_\rho \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) \\ &\quad + (\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)^\top \partial_\rho \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + o_P(1), \end{aligned}$$

where $\{s(n)\}_{n \geq 1}$ stands for a sequence of positive integers such that $n/s(n) \rightarrow 0$ as $n \rightarrow +\infty$.

Observing that, as $n \rightarrow +\infty$,

$$\sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_{s(n)})^\top = \sqrt{s(n)}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_{s(n)})^\top \times \frac{\sqrt{n}}{\sqrt{s(n)}} = o_P(1).$$

It is easy to see that,

$$(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_{s(n)})^\top \partial_\rho \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = \sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_{s(n)})^\top \frac{1}{\sqrt{n}} \partial_\rho \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}).$$

Then, it suffices to show that $\partial_{\rho}\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})/\sqrt{n}$ converges in probability to a random vector.

Asymptotic behavior of $\partial_{\rho}\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})/\sqrt{n}$.

We can write the following decomposition

$$\begin{aligned} & \frac{1}{\sqrt{n}}\partial_{\rho}\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) \\ &= \frac{1}{n}\sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\rho}(T^{\gamma_{m,h}})(\boldsymbol{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)] \\ & - \frac{1}{n}\sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \partial_{\rho}T(\boldsymbol{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)] \\ &= \mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + \mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}), \end{aligned}$$

where

$$\begin{aligned} \mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) &= \frac{1}{n}\sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\rho}(T^{\gamma_{m,h}})(\boldsymbol{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)], \\ \mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) &= -\frac{1}{n}\sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \partial_{\rho}T(\boldsymbol{\rho}_n, \gamma_0, \mathbf{X}_{t-1}) \\ & \quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)]. \end{aligned}$$

Study of the asymptotic behavior of $\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})$.

We have

$$\begin{aligned} \|\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})\|_p &\leq \frac{1}{n}\sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \|\partial_{\rho}(T^{\gamma_{m,h}})(\boldsymbol{\rho}_n, \gamma_0, \mathbf{X}_{t-1})\|_p |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)]| \\ &\leq \frac{1}{n}\sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \chi(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)]| \\ &\leq \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)]|. \end{aligned}$$

Adding and subtracting appropriate terms, we obtain,

$$\|\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})\|_p \leq \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \gamma_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]|$$

$$+ \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.$$

We know that

$$\begin{aligned} & \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\ &= -\frac{1}{V(\mathbf{X}_{t-1})} (\boldsymbol{\rho}_n - \boldsymbol{\rho}_0)^\top \partial_\rho T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \\ & \quad - \frac{1}{V(\mathbf{X}_{t-1})} (\boldsymbol{\rho}_n - \boldsymbol{\rho}_0)^\top \partial_\rho^2 T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0) \phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] (\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \\ & \quad + \frac{1}{V^2(\mathbf{X}_{t-1})} (\boldsymbol{\rho}_n - \boldsymbol{\rho}_0)^\top (\partial_\rho T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)) (\partial_\rho T(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0))^\top (\boldsymbol{\rho}_n - \boldsymbol{\rho}_0) \phi''_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)], \end{aligned}$$

where $\tilde{\boldsymbol{\rho}}_n$ lies between $\boldsymbol{\rho}_n$ and $\boldsymbol{\rho}_0$.

Then, by assumption (\mathcal{B}_4) , we obtain

$$\begin{aligned} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\leq \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ & \quad + \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]| \\ & \quad + \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi''_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

By (4.13), we can write

$$|\phi'_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)]| \leq c_\phi \|\tilde{\boldsymbol{\rho}}_n - \boldsymbol{\rho}_0\|_p \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} + |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|,$$

since as $n \rightarrow +\infty$,

$$\phi''_f[\varepsilon_t(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0)] \leq c_\phi.$$

Then

$$\begin{aligned} & |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ & \leq \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| + c_\phi \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^3 \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \\ & \quad + \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| + c_\phi \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \\ & \leq \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| + c_\phi \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 (1 + \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p) \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \\ & \quad + \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

Thus,

$$\begin{aligned} \|\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p &\leq \frac{\|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &+ \frac{c_\phi \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 (1 + \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p) \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ &+ \frac{\|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p^2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &+ \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

By the ergodic theorem, as ε_t is independent of \mathcal{F}_{t-1} , in view of Remark 3.4.4, we have

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\chi(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)]| &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\chi(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)]| \\ &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\chi(x)}{V(x)} dF_j(x) \int_{\mathbb{R}^d} |\phi_f(x)| f(x) dx < \infty. \end{aligned}$$

In view of (\mathcal{A}_7) and following the same procedure for the other terms, we obtain

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \lambda_{0,2}^{(2)} < \infty, \\ \frac{1}{n} \sum_{t=1}^n \frac{\chi^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \lambda_{0,3}^3 < \infty, \\ \frac{1}{n} \sum_{t=1}^n \frac{\chi^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \lambda_{0,2}^{(2)} < \infty. \end{aligned}$$

We know that, as $n \rightarrow +\infty$,

$$\|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p \xrightarrow{\mathbb{P}} 0,$$

then we can easily conclude that, as $n \rightarrow +\infty$,

$$\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{\mathbb{P}} \mathbf{0} \in \mathbb{R}^p.$$

Study of the asymptotic behavior of $\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$.

Using the condition (\mathcal{B}_4) , we have

$$\begin{aligned} & \|\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p \\ & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| |T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})| \|\partial_{\boldsymbol{\rho}} T(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)\|_p |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)]| \\ & \leq \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu(\mathbf{X}_{t-1})\chi(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

Using 4.13, we can write

$$\begin{aligned} \|\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p & \leq \frac{c_\phi \|\boldsymbol{\rho}_n - \boldsymbol{\rho}_0\|_p \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu(\mathbf{X}_{t-1})\chi^2(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ & \quad + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu(\mathbf{X}_{t-1})\chi(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

Since ε_t is independent of \mathcal{F}_{t-1} and using ergodic theorem, in view of (\mathcal{A}_7) , we obtain

$$\begin{aligned} & \frac{1}{n} \sum_{t=1}^n \frac{\nu(\mathbf{X}_{t-1})\chi(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\nu(\mathbf{X}_{t-1})\chi(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ & \xrightarrow[n \rightarrow +\infty]{a.s} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\nu(x)\chi(x)}{V^2(x)} dF_j(x) \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx < \infty. \end{aligned}$$

Following the same procedure, we obtain

$$\frac{1}{n} \sum_{t=1}^n \frac{\nu(\mathbf{X}_{t-1})\chi^2(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \xrightarrow[n \rightarrow +\infty]{a.s} \sum_{j=1}^{k+1} \alpha_j \lambda_{1,2}^{(3)} < \infty.$$

Then, as $n \rightarrow +\infty$, it is easy to see that $\|\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p$ tends in probability to a finite positive real number. Then,

$$\left\| \frac{1}{\sqrt{n}} \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \right\|_p \leq \|\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p + \|\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|_p < \infty.$$

It results that

$$(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_{s(n)})^\top \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = o_P(1).$$

Thus,

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + (\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1). \quad (4.16)$$

In order to treat the above equation (4.16), we need the following lemma.

Lemma 4.2.1. *Assume that (\mathcal{B}_2) hold. Let $\{s(n)\}_{n \geq 1}$ be a sequence of positive integers such that $n/s(n)$ tends to 0 as $n \rightarrow +\infty$. For $\gamma_0, \beta \in \mathbb{R}^{p(k+1)}$, $\rho_{s(n)}$ is asymptotically in the tangent space \mathbb{T}_n to the curve of $\Pi_n(\rho, \gamma, \beta)$ at ρ_n , defined as follow:*

$$\mathbb{T}_n = \left\{ z \in \mathbb{R}^p / \Pi_n(z, \gamma_0, \beta) = \Pi_n(\rho_n, \gamma_0, \beta) + (z - \rho_n)^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) \right\}.$$

Proof. Writing a second-order Taylor expansion of $\Pi_n(\rho_{s(n)}, \gamma_0, \beta)$ in a neighborhood of ρ_n , for some $\tilde{\rho}_{s(n)}$ lies between $\rho_{s(n)}$ and ρ_n we obtain

$$\begin{aligned} \Pi_n(\rho_{s(n)}, \gamma_0, \beta) &= \Pi_n(\rho_n, \gamma_0, \beta) + (\rho_{s(n)} - \rho_n)^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) \\ &\quad + \frac{1}{2} (\rho_{s(n)} - \rho_n)^\top \partial_\rho^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta) (\rho_{s(n)} - \rho_n). \end{aligned}$$

To prove that, as $n \rightarrow +\infty$, $\rho_{s(n)}$ belongs to \mathbb{T}_n , it suffices to show that $(\rho_{s(n)} - \rho_n)^\top \partial_\rho^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta) (\rho_{s(n)} - \rho_n) = o_P(1)$.

To find the asymptotic distribution of $\sqrt{n}(\rho_{s(n)} - \rho_n)$, we add and substract appropriate terms and we obtain

$$\begin{aligned} \sqrt{n}(\rho_{s(n)} - \rho_n) &= \sqrt{n}(\rho_{s(n)} - \rho_0 + \rho_0 - \rho_n) \\ &= \sqrt{n}(\rho_{s(n)} - \rho_0) + \sqrt{n}(\rho_0 - \rho_n) \\ &= \sqrt{s(n)}(\rho_{s(n)} - \rho_0) \frac{\sqrt{n}}{\sqrt{s(n)}} + \sqrt{n}(\rho_0 - \rho_n) \\ &= o_P(1) + \sqrt{n}(\rho_0 - \rho_n). \end{aligned}$$

Then, asymptotically, $\sqrt{n}(\rho_{s(n)} - \rho_n)$ have the same distribution as $\sqrt{n}(\rho_0 - \rho_n)$. This means that $\sqrt{n}(\rho_{s(n)} - \rho_n)$ converges in distribution to a normal law.

Now, to prove that $(\rho_{s(n)} - \rho_n)^\top \partial_\rho^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta) (\rho_{s(n)} - \rho_n) = o_P(1)$, it suffices to show that $\partial_\rho^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta) / \sqrt{n}$ converges in probability to a random vector.

Recall that

$$\begin{aligned} &\frac{1}{\sqrt{n}} \partial_\rho^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta) \\ &= \Delta_{1,n}(\tilde{\rho}_{s(n)}, \gamma_0, \beta) + \Delta_{2,n}(\tilde{\rho}_{s(n)}, \gamma_0, \beta) + \Delta_{3,n}(\tilde{\rho}_{s(n)}, \gamma_0, \beta) + \Delta_{4,n}(\tilde{\rho}_{s(n)}, \gamma_0, \beta), \end{aligned}$$

where $\{s(n)\}_{n \geq 1}$ stands for a sequence of positive integers such that $n/s(n) \rightarrow 0$ as $n \rightarrow +\infty$, ρ_n is given by (\mathcal{B}_2) , and $\tilde{\rho}_{s(n)}$ lies between $\rho_{s(n)}$ and ρ_0 .

We have, as $n \rightarrow +\infty$, $\|\tilde{\rho}_{s(n)} - \rho_0\| \leq \|\rho_{s(n)} - \rho_0\| \xrightarrow{a.s.} 0$, then,

$$\tilde{\rho}_{s(n)} - \rho_0 = o_P(1).$$

For some $\tilde{\rho}_n$ lying between ρ_n and ρ_0 , we proved perviously that $\|\partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta)\|_p / \sqrt{n}$ converge in probability, as $n \rightarrow +\infty$, to a finite positive number. By following the same

strategy, we can prove that $\|\partial_{\rho}^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta)\|_p / \sqrt{n}$ converge in probability to a positive finite number, where $\tilde{\rho}_{s(n)}$ lies between $\rho_{s(n)}$ and ρ_0 .

Consequently,

$$(\rho_{s(n)} - \rho_n)^\top \partial_{\rho}^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta)(\rho_{s(n)} - \rho_n) = o_P(1).$$

□

It results from Lemma 4.2.1 that, as $n \rightarrow +\infty$, $\rho_{s(n)}$ belongs to the tangent space \mathbb{T}_n . Thus, by replacing z by $\rho_{s(n)}$, we obtain

$$\Pi_n(\rho_{s(n)}, \gamma_0, \beta) = \Pi_n(\rho_n, \gamma_0, \beta) + (\rho_{s(n)} - \rho_n)^\top \partial_{\rho} \Pi_n(\rho_n, \gamma_0, \beta) + o_P(1).$$

Finally, recalling (4.16), we obtain

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_{s(n)}, \gamma_0, \beta) + o_P(1).$$

□

Proof of (ii): To prove (4.5), it suffices to show that, as $n \rightarrow +\infty$, $\hat{\eta}_n(\rho_n, \gamma_0, \beta) \rightarrow \eta(\rho_0, \gamma_0, \beta)$. For any $\beta \in \mathbb{R}^{p(k+1)}$, we have

$$\begin{aligned} & \hat{\eta}_n(\rho_n, \gamma_0, \beta) - \eta(\rho_0, \gamma_0, \beta) \\ &= \sum_{j=1}^{k+1} \hat{\alpha}_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \\ &= \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \left[\hat{\alpha}_j \hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \alpha_j \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right]. \end{aligned}$$

We add and subtract appropriate terms and we obtain

$$\begin{aligned} & \hat{\eta}_n(\rho_n, \gamma_0, \beta) - \eta(\rho_0, \gamma_0, \beta) \tag{4.17} \\ &= \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \left[\hat{\alpha}_j \hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) + \alpha_j \hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \alpha_j \hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \alpha_j \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right] \\ &= \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} (\hat{\alpha}_j - \alpha_j) \left[\hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) + \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right] \\ &+ \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \alpha_j \left[\hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right] \\ &= \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \left[(\hat{\alpha}_j - \alpha_j) \left[\hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right] + (\hat{\alpha}_j - \alpha_j) \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right] \\ &+ \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \alpha_j \left[\hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right], \end{aligned}$$

where, for $h = 1, \dots, p$,

$$\widehat{\eta}_{j,2}^{(h,m)}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0) = \frac{I(f)}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, x).$$

For all $j = 1, \dots, k+1$, we have

$$\widehat{\alpha}_j - \alpha_j \xrightarrow[n \rightarrow +\infty]{} 0.$$

For $j = 1, \dots, k+1$ and $h = 1, \dots, p$, using the assumption (\mathcal{A}_9) and the fact that the functions $\frac{1}{V}$, $\frac{\partial T}{\partial \gamma_{j,h}}$ are bounded, it follows from the Lebesgue's convergence theorem that each term in the right-hand side of (4.17) tends to 0.

□

For testing H_0 against $H_{\boldsymbol{\beta}}^{(n)}$, for any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, we consider the following statistic

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}.$$

Theorem 4. (*Optimality*)

Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) and (\mathcal{B}_1) - (\mathcal{B}_5) hold. Then, for any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, we have

- i. Under H_0 , as $n \rightarrow +\infty$, $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$.
- ii. Under local alternatives $H_{\boldsymbol{\beta}}^{(n)}$, at level of significance $\alpha \in]0, 1[$, the asymptotic power of the test based on the statistic $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ is $\mathcal{P}_{k,\tau^k} = 1 - \Phi(z_\alpha - \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}))$, where z_α is the $(1 - \alpha)$ -quantile of a standard Gaussian distribution with cumulative distribution function Φ .
- iii. The test based on the statistic $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ is locally asymptotically optimal.

Proof: The statistic that we have is

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}.$$

We proved in Proposition (4.2.2) that

$$\Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1).$$

Also, we proved that, as $n \rightarrow +\infty$,

$$\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{\mathbb{P}} \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}).$$

Then, under H_0 and as $n \rightarrow +\infty$, we have

$$\lim_{n \rightarrow +\infty} \frac{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} = 1.$$

Then, under H_0 and as $n \rightarrow +\infty$, we have

$$\begin{aligned} \mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \\ &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1)}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \\ &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} + \frac{1}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \times o_P(1) \\ &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \times \frac{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} + \frac{1}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \times o_P(1) \\ &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \times \frac{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} + \frac{1}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \times o_P(1) \\ &\simeq \mathcal{T}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}). \end{aligned}$$

The convergence of $\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ to $\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ remains true under the local alternatives $H_\beta^{(n)}$ by contiguity. Then, under $H_\beta^{(n)}$ and as $n \rightarrow +\infty$, using Lecam's third lemma, we have

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{\widehat{\Pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \xrightarrow{\mathcal{D}} \mathcal{N}(\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}), 1).$$

Thus, we can say that the power of the test remains the same.

With this, we have established that

- The central sequence based on the estimated version has no effect to the power of the test.
- The statistic based on the estimated version is asymptotically equivalent to that based on the true parameter.

Then we conclude that the test remains optimal in the case where the nuisance parameter is unknown with known $\boldsymbol{\gamma}_0$.

□

4.3 The functions involved in the model and the parameters are unknown.

In practice, γ_0 may be unknown and has to be estimated. One of estimation methods is the maximum likelihood. Denote by $\gamma_n = \gamma_0 + \beta/\sqrt{n}$, for some $\beta \in \mathbb{R}^{p(k+1)}$. Let $\hat{\gamma}_n$ and $\hat{\gamma}_{0,n}$ be respectively, the maximum likelihood estimator of γ_n and γ_0 .

We need the following additional assumptions

$$(\mathcal{B}'_1) \text{ For any } m = 1, \dots, k+1, h = 1, \dots, p, \\ \max \left\{ \sup_{\gamma} |T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}, x)|, \sup_{\gamma} \|\partial_{\gamma} T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}, x)\|_{p(k+1)}, \right. \\ \left. \sup_{\gamma} \|\partial_{\gamma}^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}, x)\|_{p(k+1)} \right\} < \kappa(x), \text{ where } T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}, x) = \frac{\partial T}{\partial \gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}, x).$$

$$(\mathcal{B}'_2) \text{ For } j = 1, \dots, k+1, \text{ assume that, for } u = 1, 2, 3 \text{ and } a, b \in \{0, 1, 2, 3\}$$

$$\delta_{a,b}^{(u)} = \int_{\mathbb{R}^d} \frac{\nu^a(x) \kappa^b(x)}{V^u(x)} dF_j(x) < \infty.$$

$$(\mathcal{B}'_3) \text{ Assume that } \sqrt{n}(\hat{\gamma}_{0,n} - \gamma_0) = O_P(1).$$

Remark 4.3.1. Let $\hat{\gamma}_{0,n}$ and $\tilde{\gamma}_{0,n}$ be the maximum likelihood estimator of γ_0 under H_0 and $H_{\beta}^{(n)}$ respectively. In probability, asymptotically, we have

$$\tilde{\gamma}_{0,n} = \hat{\gamma}_{0,n} + \beta/\sqrt{n}.$$

Let $s(n)$ be any sequence of positive integers satisfying $n/s(n) \rightarrow 0$, as $n \rightarrow +\infty$.

When γ_0 is unknown, for testing H_0 against $H_{\beta}^{(n)}$, for any $\beta \in \mathbb{R}^{p(k+1)}$, our test statistic can be expressed as follow

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \hat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \hat{\gamma}_{0,s(n)}, \boldsymbol{\beta})}{\hat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \hat{\gamma}_{0,s(n)}, \boldsymbol{\beta})}.$$

Proposition 4.3.1. Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) , (\mathcal{B}_1) - (\mathcal{B}_5) and (\mathcal{B}'_1) - (\mathcal{B}'_3) hold. Then, for any sequence $s(n)$ of positive integers satisfying, as $n \rightarrow +\infty$, $n/s(n) \rightarrow 0$, for any sequence of consistent and asymptotically normal estimators $\{(\boldsymbol{\rho}_n, \hat{\gamma}_{0,n})\}_{n \geq 1}$ of $(\boldsymbol{\rho}_0, \gamma_0)$ and for any $\beta \in \mathbb{R}^{p(k+1)}$, we have, under H_0 and as $n \rightarrow +\infty$,

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \hat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1).$$

Proof: The proof relies on a number of lemmas that we state and prove.

Lemma 4.3.1. *Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) , (\mathcal{B}_1) - (\mathcal{B}_5) , and (\mathcal{B}'_1) - (\mathcal{B}'_3) hold. Then, for any sequence $s(n)$ of positive integers satisfying, as $n \rightarrow +\infty$, $n/s(n) \rightarrow 0$, for any sequence of consistent and asymptotically normal estimators $\{\widehat{\gamma}_{0,n}\}_{n \geq 1}$ of γ_0 and for any $\beta \in \mathbb{R}^{p(k+1)}$, we have, under H_0 and as $n \rightarrow +\infty$,*

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_0, \widehat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

Proof: For any $\rho \in \mathbb{R}^p$ and $(\gamma, \beta) \in \mathbb{R}^{p(k+1)} \times \mathbb{R}^{p(k+1)}$, for $\widehat{\gamma}_{0,n}$ the maximum likelihood or the least squares estimator of γ_0 , we write a first-order Taylor expansion of $\Pi_n(\rho_0, \gamma_0, \beta)$ in a neighborhood of $\widehat{\gamma}_{0,n}$ and we obtain, for some $\widetilde{\gamma}_{0,n}$ lying between γ_0 and $\widehat{\gamma}_{0,n}$,

$$\begin{aligned} \Pi_n(\rho_0, \gamma_0, \beta) &= \Pi_n(\rho_0, \widehat{\gamma}_{0,n}, \beta) - (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Pi_n(\rho_0, \widehat{\gamma}_{0,n}, \beta) \\ &\quad + \frac{1}{2} (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,n} - \gamma_0), \end{aligned}$$

where

$$\partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) = \begin{pmatrix} \partial_{\gamma_1}^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) & \cdots & \partial_{\gamma_1 \gamma_p}^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) \\ \vdots & \ddots & \vdots \\ \partial_{\gamma_p \gamma_1}^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) & \cdots & \partial_{\gamma_p}^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) \end{pmatrix} \in \mathcal{M}_{p(k+1)}.$$

Our aim is to prove that, under H_0 , as $n \rightarrow +\infty$,

$$(\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Pi_n(\rho_0, \widehat{\gamma}_{0,n}, \beta) = o_P(1), \quad (4.18)$$

$$(\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,n} - \gamma_0) = o_P(1). \quad (4.19)$$

Starting with (4.19), we observe that

$$\begin{aligned} &\left| (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,n} - \gamma_0) \right| \\ &\leq \|\widehat{\gamma}_{0,n} - \gamma_0\|_{p(k+1)} \times \left\| \left\| \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) \right\| \right\|_{p(k+1)} \times \|\widehat{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}. \end{aligned}$$

Multiplying and dividing by \sqrt{n} , we obtain

$$\begin{aligned} &\left| (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) (\widehat{\gamma}_{0,n} - \gamma_0) \right| \\ &\leq \|\sqrt{n}(\widehat{\gamma}_{0,n} - \gamma_0)\|_{p(k+1)} \times \frac{1}{\sqrt{n}} \left\| \left\| \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta) \right\| \right\|_{p(k+1)} \times \|\widehat{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}. \end{aligned}$$

The Proposition 4.2.1 treats the case where ρ_n is an asymptotically normal estimator of ρ_0 . Same thing when $\widehat{\gamma}_{0,n}$ is an estimator of γ_0 . Then, using Proposition 4.2.1, under H_0 , as $n \rightarrow +\infty$, $\sqrt{n}(\widehat{\gamma}_{0,n} - \gamma_0)$ converges in distribution to a normal distribution and $\widehat{\gamma}_{0,n} - \gamma_0$ tends to 0 in probability as $n \rightarrow +\infty$.

Then, to prove (4.19), it suffices to show that $\|\|\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)}/\sqrt{n}$ tends in probability, as $n \rightarrow +\infty$, to some positive random variable. Recalling (3.9), we have

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}, \boldsymbol{\beta}) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top N(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma})] \right\} \\ &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}, \boldsymbol{\gamma}, \boldsymbol{\beta}) \phi_f[\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma})]. \end{aligned}$$

Then,

$$\partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) = (\omega_1 \partial_{\gamma_1} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}), \dots, \omega_{k+1} \partial_{\gamma_{k+1}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}))^\top \in \mathbb{R}^{p(k+1)},$$

and

$$\begin{aligned} &\partial_\gamma^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \\ &= \begin{pmatrix} \omega_1^2 H_{\gamma_1} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \omega_2^2 H_{\gamma_2} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \omega_{k+1}^2 H_{\gamma_{k+1}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \\ &\in \mathcal{M}_{p(k+1)}(\mathbb{R}), \end{aligned}$$

where, for any $i = 1, \dots, k+1$, $H_{\gamma_i} [T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})]$ is the Hessian matrix of $T^{\gamma_{m,h}}$ with respect to γ_i .

We recall (3.2) and write, for any $t \in \mathbb{N}$, $\boldsymbol{\rho} \in \mathbb{R}^p$ and $\boldsymbol{\gamma} \in \mathbb{R}^{p(k+1)}$,

$$\varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}) = \frac{X_t - T(\boldsymbol{\rho} + \boldsymbol{\gamma} \odot \boldsymbol{\omega}(t), \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})},$$

then,

$$\partial_\gamma \varepsilon_t(\boldsymbol{\rho}, \boldsymbol{\gamma}) = -\frac{1}{V(\mathbf{X}_{t-1})} \partial_\gamma T(\boldsymbol{\rho}, \boldsymbol{\gamma}, \mathbf{X}_{t-1}).$$

Remember that we wish to bound $\frac{1}{\sqrt{n}} \|\|\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)}$. Also, recall that

$$\begin{aligned} \frac{1}{\sqrt{n}} \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \boldsymbol{\beta}) &= \frac{1}{n} \partial_\gamma \left(\sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \right) \\ &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\gamma (T^{\gamma_{m,h}})(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\ &\quad - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \partial_\gamma T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f'[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})]. \end{aligned}$$

From this, for any $\boldsymbol{\rho} \in \mathbb{R}^p$ and $(\boldsymbol{\gamma}, \boldsymbol{\beta}) \in \mathbb{R}^{p(k+1)} \times \mathbb{R}^{p(k+1)}$, we have

$$\begin{aligned}
\frac{1}{\sqrt{n}} \partial_{\boldsymbol{\gamma}}^2 \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \boldsymbol{\beta}) &= \frac{1}{\sqrt{n}} \partial_{\boldsymbol{\gamma}} [\partial_{\boldsymbol{\gamma}} \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \boldsymbol{\beta})] \\
&= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\gamma}}^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&\quad - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_{\boldsymbol{\gamma}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}))^{\top} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&\quad - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_{\boldsymbol{\gamma}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}))^{\top} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&\quad - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \partial_{\boldsymbol{\gamma}}^2 T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&\quad + \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}))^{\top} \\
&\quad \times \phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\gamma}}^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&\quad - \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_{\boldsymbol{\gamma}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}))^{\top} \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&\quad - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \partial_{\boldsymbol{\gamma}}^2 T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})] \\
&\quad + \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1})) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}))^{\top} \\
&\quad \times \phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma})].
\end{aligned}$$

Thus,

$$\begin{aligned}
\frac{1}{\sqrt{n}} \partial_{\boldsymbol{\gamma}}^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) - \Xi_{2,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \\
&\quad - \Xi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + \Xi_{4,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}),
\end{aligned} \tag{4.20}$$

where

$$\Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) = \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\gamma}}^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})],$$

$$\begin{aligned}\Xi_{2,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_{\boldsymbol{\gamma}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})) (\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}))^{\top} \\ &\quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})], \\ \Xi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}) \partial_{\boldsymbol{\gamma}}^2 T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}) \\ &\quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})], \\ \Xi_{4,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}) \left(\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}) \right) \\ &\quad \times \left(\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}) \right)^{\top} \phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})].\end{aligned}$$

Here, our main purpose is to show that, for $i = 1, \dots, 4$, $\left\| \Xi_{i,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)}$ converges to a finite positive number.

Study of the asymptotic behavior of $\Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

We have

$$\begin{aligned}& \left\| \Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} \\ &= \left\| \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\gamma}}^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] \right\|_{p(k+1)} \\ &\leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \left\| \partial_{\boldsymbol{\gamma}}^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \right\|_{p(k+1)} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})]|.\end{aligned}$$

By assumption (\mathcal{B}'_1) , we obtain

$$\left\| \Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})]|.$$

Adding and subtracting appropriate terms, we obtain

$$\begin{aligned}& \left\| \Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} \\ &\leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] + \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &\leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ &+ \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.\end{aligned}$$

By a second-order Taylor expansion, for some $\check{\gamma}_{0,n}$ lying between $\tilde{\gamma}_{0,n}$ and γ_0 , we have

$$\begin{aligned} & \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \\ &= -\frac{1}{V(\mathbf{X}_{t-1})}(\tilde{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma T(\boldsymbol{\rho}_0, \gamma_0, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)] \\ & - \frac{1}{2V(\mathbf{X}_{t-1})}(\tilde{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 T(\boldsymbol{\rho}_0, \check{\gamma}_{0,n}, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n})](\tilde{\gamma}_{0,n} - \gamma_0) \\ & + \frac{1}{2V^2(\mathbf{X}_{t-1})}(\tilde{\gamma}_{0,n} - \gamma_0)^\top (\partial_\gamma T(\boldsymbol{\rho}_0, \check{\gamma}_{0,n}, \mathbf{X}_{t-1})) (\partial_\gamma T(\boldsymbol{\rho}_0, \check{\gamma}_{0,n}, \mathbf{X}_{t-1}))^\top (\tilde{\gamma}_{0,n} - \gamma_0) \\ & \times \phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n})]. \end{aligned}$$

Then, using the assumption (\mathcal{A}_9) , we obtain

$$\begin{aligned} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| &\leq \|\tilde{\gamma}_{0,n} - \gamma_0\|_{p(k+1)} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| \\ & + \|\tilde{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}^2 \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n})]| \\ & + \|\tilde{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}^2 \left| \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \right| |\phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n})]|. \end{aligned}$$

Using the fact that ϕ'_f is c_ϕ -lipshitz, we have, for any $\gamma \in \mathbb{R}^{p(k+1)}$,

$$|\phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma)]| \leq c_\phi,$$

and

$$|\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n})]| \leq c_\phi |\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n}) - \varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)| + |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]|.$$

From equality (3.6), we obtain

$$|\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n}) - \varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)| \leq \frac{|\nu(\mathbf{X}_{t-1})| \|\check{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}}{|V(\mathbf{X}_{t-1})|},$$

from which we have

$$|\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \check{\gamma}_{0,n})]| \leq c_\phi \|\check{\gamma}_{0,n} - \gamma_0\|_{p(k+1)} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| + |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]|. \quad (4.21)$$

Thus,

$$\begin{aligned} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\gamma}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| &\leq \|\tilde{\gamma}_{0,n} - \gamma_0\|_{p(k+1)} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| \quad (4.22) \\ & + c_\phi \|\tilde{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}^3 \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^2 \\ & + \|\tilde{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}^2 \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \gamma_0)]| \\ & + c_\phi \|\tilde{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}^2 \left| \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \right|. \end{aligned}$$

Then,

$$\begin{aligned}
& \|\|\Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)} \\
& \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
& + \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
& \leq \frac{\|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}}{n} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
& + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| c_\phi \|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^3}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1}) \nu^2(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\
& + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
& + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| c_\phi \|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2}{n} \sum_{t=1}^n \frac{\kappa^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\
& + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.
\end{aligned}$$

Applying suitably the ergodic theorem on each term in the right-hand side of the above inequality, by assumption (\mathcal{B}'_2) , since ε_t is independent of \mathcal{F}_{t-1} , in view of (\mathcal{A}_7) , we obtain

$$\begin{aligned}
& \frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
& = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\kappa(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
& \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\kappa(x) \nu(x)}{V^2(x)} dF_j(x) \mathbb{E} \{ |\phi'_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \} \\
& \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \delta_{1,1}^{(2)} < \infty.
\end{aligned}$$

Also,

$$\frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1}) \nu^2(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\kappa(\mathbf{X}_{t-1}) \nu^2(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})}$$

$$\begin{aligned} & \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\kappa(x)\nu^2(x)}{V^3(x)} dF_j(x) \\ & \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{2,1}^{(3)} \\ & < \infty. \end{aligned}$$

Furthermore, since ε_t is independent of \mathcal{F}_{t-1} and based on (\mathcal{A}_7) ,

$$\begin{aligned} & \frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})\nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ & = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\kappa(\mathbf{X}_{t-1})\nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ & \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\kappa(x)\nu(x)}{V^2(x)} \mathbb{E} \{ |\phi'_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \} \\ & = \sum_{j=1}^{k+1} \alpha_j \delta_{1,1}^{(2)} \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \\ & < \infty. \end{aligned}$$

In the same lines, one shows that

$$\frac{1}{n} \sum_{t=1}^n \frac{\kappa^3(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})} \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{0,3}^{(3)} < \infty.$$

Also,

$$\frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} |\phi_f(x)| f(x) dx \delta_{0,1}^{(1)} < \infty.$$

Since $\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|$ and c_ϕ are finite, we have

$$\|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \leq \|\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \xrightarrow[n \rightarrow +\infty]{} 0,$$

we conclude that, as $n \rightarrow +\infty$,

$$\left\| \Xi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} \rightarrow 0.$$

Study of the asymptotic behavior of $\Xi_{2,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

By assumption (\mathcal{B}'_1) , we have

$$\begin{aligned} & \left\| \Xi_{2,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} \\ & \leq \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \left\| \partial_{\boldsymbol{\gamma}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \right\|_{p(k+1)}^2 |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})]| \\ & \leq \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \kappa^2(\mathbf{X}_{t-1}) |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})]|. \end{aligned}$$

Using the inequality (4.21), we obtain

$$\begin{aligned} \left\| \Xi_{2,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} & \leq \frac{2c_\phi \|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ & \quad + \frac{2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|. \end{aligned}$$

By the piece-wise stationarity and the ergodic theorem, using (\mathcal{B}'_2) , we obtain

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\kappa^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} & = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\kappa^2(\mathbf{X}_{t-1}) \nu(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ & \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\kappa^2(x) \nu(x)}{V^3(x)} dF_j(x) \\ & \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{1,2}^{(3)} < \infty. \end{aligned}$$

Since $\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|$ is a finite summation, since as $n \rightarrow +\infty$, $\|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \rightarrow 0$, its result that

$$\frac{2c_\phi \|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{j=1}^{k+1} \alpha_j \delta_{1,2}^{(3)} \xrightarrow{a.s.} 0.$$

Also, since ε_t is independent of \mathcal{F}_{t-1} , in view of (\mathcal{A}_7) and (\mathcal{B}'_2) , we have

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\kappa^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| & = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\kappa^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\ & \xrightarrow{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \delta_{0,2}^{(2)} < \infty. \end{aligned}$$

Hence, as $n \rightarrow +\infty$,

$$\left\| \Xi_{2,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} < \infty.$$

Study of the asymptotic behavior of $\Xi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

Recall that

$$\begin{aligned} \Xi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \partial_{\boldsymbol{\gamma}}^2 T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \\ &\quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})]. \end{aligned}$$

Using the assumptions (\mathcal{B}'_1) and (\mathcal{A}_9) , we have

$$\begin{aligned} &\left\| \Xi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{\mathcal{M}_{p(k+1)}} \\ &\leq \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \left| T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \right| \left\| \partial_{\boldsymbol{\gamma}}^2 T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \right\|_{p(k+1)} \\ &\quad \times \left| \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] \right| \\ &\leq \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \left| \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] \right|. \end{aligned}$$

Using the inequality (4.21), we obtain

$$\begin{aligned} \left\| \Xi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} &\leq \frac{c_\phi \|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^3(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})} \\ &\quad + \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \left| \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] \right|. \end{aligned}$$

By the ergodic theorem, using (\mathcal{B}'_2) , as $n \rightarrow +\infty$, we obtain

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\nu^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\nu^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\ &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\nu^3(x)}{V^3(x)} dF_j(x) \\ &= \sum_{j=1}^{k+1} \alpha_j \delta_{3,0}^{(3)} < \infty. \end{aligned}$$

Furthermore, as $n \rightarrow +\infty$,

$$\frac{c_\phi \|\tilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^3(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})^3} \xrightarrow{a.s.} 0.$$

Again, by the ergodic theorem, since ε_t is independent of \mathcal{F}_{t-1} , in view of (\mathcal{A}_7) ,

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\nu^2(x)}{V^2(x)} dF_j(x) \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx \\ &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{2,0}^{(2)} \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx < \infty. \end{aligned}$$

We conclude that, as $n \rightarrow +\infty$,

$$\|\|\Xi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)} < \infty.$$

Study of the asymptotic behavior of $\Xi_{4,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

Using the assumptions (\mathcal{A}_9) and (\mathcal{A}_6) , we obtain

$$\begin{aligned} &\|\|\Xi_{4,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)} \\ &\leq \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| |T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1})| \left\| \partial_{\gamma} T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}) \right\|_{p(k+1)}^2 \\ &\times \left| \phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] \right| \\ &\leq \frac{c_{\phi} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^3(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})}. \end{aligned}$$

By piece-wise stationarity and the ergodic theorem, in view of (\mathcal{B}'_2) , we obtain

$$\begin{aligned} \frac{1}{n} \sum_{t=1}^n \frac{\nu^3(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}} \frac{\nu^3(x)}{V^2(x)} dF_j(x) \\ &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{3,0}^{(2)} < \infty. \end{aligned}$$

Then, as $n \rightarrow +\infty$,

$$\|\|\Xi_{4,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)} < \infty.$$

Then, for $i = 1, \dots, 4$, we proved that $\|\|\Xi_{i,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|\|_{p(k+1)}$ converges to a finite positive number. From this, one will have that $\left\| \left\| \partial_{\gamma}^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\| \right\|_{p(k+1)} / \sqrt{n}$ converges to a finite positive number.

For proving (4.19), one can write

$$\begin{aligned} &\left| (\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)^\top \partial_{\gamma}^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) (\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) \right| \\ &\leq \left\| \frac{1}{\sqrt{n}} (\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) \right\|_{p(k+1)} \frac{1}{\sqrt{n}} \|\|\partial_{\gamma}^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)} \|\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}. \end{aligned}$$

Using Proposition 4.2.1 we can see that, under H_0 , $\|\widehat{\gamma}_{0,n} - \gamma_0\|_{p(k+1)} / \sqrt{n}$ converges to a finite positive number. Since $\widehat{\gamma}_{0,n} - \gamma_0$ tends to $\mathbf{0} \in \mathbb{R}^{p(k+1)}$ in probability, and $\|\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta})\|_{p(k+1)} / \sqrt{n}$ converges under H_0 to some finite positive number, it follows that

$$(\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta}) (\widehat{\gamma}_{0,n} - \gamma_0) = o_P(1). \quad (4.23)$$

Treating (4.18).

Based on (4.14), we can write

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) - (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1).$$

Now, we prove that

$$(\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) = o_P(1).$$

By adding and subtracting appropriate terms, we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\gamma_0 - \widehat{\gamma}_{0,n} + \widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1) \\ &= \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\gamma_0 - \widehat{\gamma}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) \\ &\quad + (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1), \end{aligned}$$

where $\{s(n)\}_{n \geq 1}$ stands for a sequence of positive integers such that $n/s(n) \rightarrow 0$ as $n \rightarrow +\infty$.

Observing that, as $n \rightarrow +\infty$,

$$\sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,s(n)})^\top = \sqrt{s(n)}(\gamma_0 - \widehat{\gamma}_{0,s(n)})^\top \times \frac{\sqrt{n}}{\sqrt{s(n)}} = o_P(1), \quad (4.24)$$

it is easy to see that,

$$(\gamma_0 - \widehat{\gamma}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) = \sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,s(n)})^\top \frac{1}{\sqrt{n}} \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}).$$

Then, it suffices to show that $(1/\sqrt{n}) \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta})$ converge in probability, as $n \rightarrow +\infty$, to a random vector.

Study of the asymptotic behavior of $\partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) / \sqrt{n}$.

Recalling that

$$\begin{aligned}
 & \frac{1}{\sqrt{n}} \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \\
 &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})] \\
 & - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}) \partial_\gamma T(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})] \\
 &= \psi_{1,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) - \psi_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}),
 \end{aligned}$$

where

$$\begin{aligned}
 \psi_{1,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})] \\
 \psi_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}) \partial_\gamma T(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})].
 \end{aligned}$$

Study of the asymptotic behavior of $\psi_{1,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

By assumption (\mathcal{B}'_1) , we have

$$\begin{aligned}
 & \|\psi_{1,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_p \\
 & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \|\partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})\|_{p(k+1)} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})]| \\
 & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})]|.
 \end{aligned}$$

Adding and subtracting appropriate terms, we obtain

$$\begin{aligned}
 & \|\psi_{1,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_p \\
 & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)] + \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 & \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 & + \frac{1}{n} \sum_{t=1}^n \frac{1}{|V(\mathbf{X}_{t-1})|} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| \kappa(\mathbf{X}_{t-1}) |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.
 \end{aligned}$$

Using (4.22), we can write

$$\begin{aligned}
 |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})] - \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\leq \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 &+ c_\phi \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^3 \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right|^2 \\
 &+ \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 \left| \frac{\nu(\mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \right| |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 &+ c_\phi \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 \left| \frac{\kappa^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} \right|.
 \end{aligned}$$

Then, as $n \rightarrow +\infty$,

$$\begin{aligned}
 \|\Psi_{1,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_p &\leq \frac{\|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})\nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 &+ \frac{c_\phi \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^3 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})\nu^2(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\
 &+ \frac{\|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})\nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 &+ \frac{c_\phi \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}^2 \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} \\
 &+ \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} \phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)].
 \end{aligned}$$

By a suitable application of the ergodic theorem, as ε_t is independent of \mathcal{F}_{t-1} , in view of (\mathcal{A}_7) , we obtain

$$\begin{aligned}
 \frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})\nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\kappa(\mathbf{X}_{t-1})\nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \\
 &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\kappa(x)\nu(x)}{V^2(x)} dF_j(x) \mathbb{E}\{|\phi'_f[\varepsilon_{\tau_{j-1}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|\} \\
 &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{1,1}^{(2)} \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx < \infty.
 \end{aligned}$$

Following the same procedure, we obtain

$$\begin{aligned}
 \frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})\nu^2(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})} &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{2,1}^{(3)} < \infty, \\
 \frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})\nu(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| &\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{1,1}^{(2)} \int_{\mathbb{R}^d} |\phi'_f(x)| f(x) dx < \infty,
 \end{aligned}$$

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{t=1}^n \frac{\kappa^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} = \sum_{j=1}^{k+1} \alpha_j \delta_{0,3}^{(3)} < \infty,$$

$$\frac{1}{n} \sum_{t=1}^n \frac{\kappa(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|} |\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \delta_{0,1}^{(1)} \int_{\mathbb{R}^d} |\phi_f(x)| f(x) dx < \infty.$$

We know that

$$\lim_{n \rightarrow +\infty} \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} = 0,$$

then, it is easy to see that, as $n \rightarrow +\infty$,

$$\|\boldsymbol{\psi}_{1,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_p \rightarrow 0.$$

Study of the asymptotic behavior of $\boldsymbol{\psi}_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

We have

$$\|\boldsymbol{\psi}_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_{p(k+1)} \leq \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}| |T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1})|$$

$$\times \|\partial_{\boldsymbol{\gamma}} T(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1})\|_{p(k+1)} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})]|.$$

Using the assumption (\mathcal{A}_9) , we can write

$$\|\boldsymbol{\psi}_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_{p(k+1)} \leq \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n})]|.$$

Using (4.13), we obtain

$$\|\boldsymbol{\psi}_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_{p(k+1)} \leq \frac{c_{\phi} \|\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)} \sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^3(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})}$$

$$+ \frac{\sum_{m=1}^{k+1} \sum_{h=1}^p |\beta_{m,h}|}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]|.$$

Using the ergodic theorem and (\mathcal{B}_5) , we obtain

$$\frac{1}{n} \sum_{t=1}^n \frac{\nu^3(\mathbf{X}_{t-1})}{|V(\mathbf{X}_{t-1})|^3} = \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{\nu^3(\mathbf{X}_{t-1})}{V^3(\mathbf{X}_{t-1})}$$

$$\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \frac{\nu^3(x)}{V^3(x)} dF_j(x)$$

$$\xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \lambda_{0,3}^{(3)} < \infty.$$

In a similar way, we obtain

$$\frac{1}{n} \sum_{t=1}^n \frac{\nu^2(\mathbf{X}_{t-1})}{V^2(\mathbf{X}_{t-1})} |\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]| \xrightarrow[n \rightarrow +\infty]{a.s.} \sum_{j=1}^{k+1} \alpha_j \lambda_{0,2}^{(2)} \int_{\mathbb{R}^d} \phi'_f(x) f(x) dx < \infty.$$

It is easy to see that, under H_0 , as $n \rightarrow +\infty$,

$$\|\boldsymbol{\psi}_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_{p(k+1)} < \infty.$$

From the asymptotic behavior of $\boldsymbol{\psi}_{1,n}$ and $\boldsymbol{\psi}_{2,n}$, we can see that $\|\partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_{p(k+1)} / \sqrt{n}$ is bounded. It results that

$$(\boldsymbol{\gamma}_0 - \widehat{\boldsymbol{\gamma}}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) = o_P(1).$$

Thus,

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + o_P(1). \quad (4.25)$$

To treat the previous equation, we need the following lemma.

Lemma 4.3.2. *Let $\widehat{\boldsymbol{\gamma}}_{0,n}$ be a consistent and asymptotically normal estimator of $\boldsymbol{\gamma}_0 \in \mathbb{R}^{p(k+1)}$. Let $\{s(n)\}_{n \geq 1}$ be a sequence of positive integers such that $n/s(n)$ tends to 0 as $n \rightarrow +\infty$. For $\boldsymbol{\rho}_0 \in \mathbb{R}^p$ and $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, $\widehat{\boldsymbol{\gamma}}_{0,s(n)}$ is asymptotically in the tangent space $\widetilde{\mathbb{T}}_n$ to the curve $\Pi_n(\boldsymbol{\rho}, \boldsymbol{\gamma}, \boldsymbol{\beta})$ at $\widehat{\boldsymbol{\gamma}}_{0,n}$, defined as follow:*

$$\widetilde{\mathbb{T}}_n = \left\{ y \in \mathbb{R}^{p(k+1)} \mid \Pi_n(\boldsymbol{\rho}_0, y, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + (y - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\}.$$

Proof: Writing a second-order Taylor expansion of $\Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$ in a neighborhood of $\widehat{\boldsymbol{\gamma}}_{0,n}$, for some $\widetilde{\boldsymbol{\gamma}}_{0,s(n)}$ lying between $\widehat{\boldsymbol{\gamma}}_{0,s(n)}$ and $\widehat{\boldsymbol{\gamma}}_{0,n}$, we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \\ &\quad + \frac{1}{2} (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n}). \end{aligned}$$

To prove that, as $n \rightarrow +\infty$, $\widehat{\boldsymbol{\gamma}}_{0,n}$ belongs to $\widetilde{\mathbb{T}}_n$, it suffices to show that

$$(\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n}) = o_P(1).$$

Then, we study the asymptotic distribution of $\sqrt{n} (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})$. By adding and subtracting appropriate terms, we obtain

$$\begin{aligned} \sqrt{n}(\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n}) &= \sqrt{n}(\widehat{\boldsymbol{\gamma}}_{0,s(n)} + \boldsymbol{\gamma}_0 - \boldsymbol{\gamma}_0 - \widehat{\boldsymbol{\gamma}}_{0,n}) \\ &= \sqrt{n}(\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \boldsymbol{\gamma}_0) + \sqrt{n}(\boldsymbol{\gamma}_0 - \widehat{\boldsymbol{\gamma}}_{0,n}) \\ &= \sqrt{s(n)}(\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \boldsymbol{\gamma}_0) \frac{\sqrt{n}}{\sqrt{s(n)}} + \sqrt{n}(\boldsymbol{\gamma}_0 - \widehat{\boldsymbol{\gamma}}_{0,n}) \\ &= o_P(1) + \sqrt{n}(\boldsymbol{\gamma}_0 - \widehat{\boldsymbol{\gamma}}_{0,n}). \end{aligned}$$

It is easy to see that, as $n \rightarrow +\infty$, $\sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})$ has the same distribution as $\sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n})$ and then, $\sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})$ converge in distribution to a normal law.

Now, to prove that $(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) = o_P(1)$, it suffices to show that $\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) / \sqrt{n}$ converges in probability to a random vector.

Using (4.26), we can write

$$\begin{aligned} & \frac{1}{\sqrt{n}} \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) \\ &= \Xi_{1,n}(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) - \Xi_{2,n}(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) - \Xi_{3,n}(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + \Xi_{4,n}(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}), \end{aligned}$$

where $\{s(n)\}_{n \geq 1}$ stands for a sequence of positive integers, $\widehat{\gamma}_{0,n}$ is an asymptotically normal estimator of γ_0 and $\widetilde{\gamma}_{0,s(n)}$ lies between $\widehat{\gamma}_{0,s(n)}$ and γ_0 .

Asymptotically, we have $\left\| \widetilde{\gamma}_{0,s(n)} - \gamma_0 \right\|_{p(k+1)} \xrightarrow{a.s.} 0$, which implies that

$$\widetilde{\gamma}_{0,s(n)} - \gamma_0 = o_P(1).$$

Previously, we proved that $\left\| \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} / \sqrt{n}$ converges in probability to a positive random variable where $\widetilde{\gamma}_{0,n}$ lies between $\widehat{\gamma}_{0,n}$ and γ_0 .

Following the same techniques, we can prove that $\left\| \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) \right\|_{p(k+1)} / \sqrt{n}$ converges in probability to a positive random variable, where $\widetilde{\gamma}_{0,s(n)}$ lies between $\widehat{\gamma}_{0,s(n)}$ and γ_0 .

It results that

$$(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) = o_P(1).$$

□

It follows from Lemma 4.3.2 that, as $n \rightarrow +\infty$, $\widehat{\gamma}_{0,s(n)}$ belongs to the tangent space $\widetilde{\mathbb{T}}_n$. Replacing y by $\widehat{\gamma}_{0,s(n)}$, we obtain

$$\Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1)$$

Recalling (4.25), we obtain

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \widehat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1).$$

□

Lemma 4.3.3. *Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) , (\mathcal{B}_1) - (\mathcal{B}_5) and (\mathcal{B}'_1) - (\mathcal{B}'_3) hold. Let $\{\widehat{\gamma}_{0,n}\}_{n \geq 1}$ be a sequence of consistent and asymptotically normal estimators of γ_0 . Let $s(n)$ be any sequence of positive integers such that $n/s(n) \rightarrow 0$ as $n \rightarrow +\infty$. Then, for any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, under H_0 , as $n \rightarrow +\infty$, we have,*

$$\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1).$$

Proof: Following the same technique as above and by applying Taylor expansion of $\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ in a neighborhood of $\widehat{\boldsymbol{\gamma}}_{0,n}$, for some $\widetilde{\boldsymbol{\gamma}}_{0,n}$ lying between $\boldsymbol{\gamma}_0$ and $\widehat{\boldsymbol{\gamma}}_{0,n}$, we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) - (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \\ &\quad + \frac{1}{2} (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0). \end{aligned}$$

We have

$$\begin{aligned} &\left| (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) \right| \\ &= \left| \sqrt{n} (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)^\top \frac{1}{\sqrt{n}} \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) \right| \\ &\leq \left\| \sqrt{n} (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) \right\|_{p(k+1)} \times \frac{1}{\sqrt{n}} \left\| \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} \left\| \widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0 \right\|_{p(k+1)}. \end{aligned}$$

Recall that, under H_0 , $\sqrt{n}(\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)$ converges to a finite Gaussian random vector and, almost surely, as $n \rightarrow +\infty$, $\widetilde{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0$ tends to $\mathbf{0} \in \mathbb{R}^{p(k+1)}$; we study the convergence of $\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) / \sqrt{n}$.

Based on the proof of Lemma 4.3.1, we have

$$\begin{aligned} \frac{1}{\sqrt{n}} \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \Xi_{1,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) - \Xi_{2,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \\ &\quad - \Xi_{3,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + \Xi_{4,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}), \end{aligned} \quad (4.26)$$

where

$$\begin{aligned} \Xi_{1,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\gamma^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n})], \\ \Xi_{2,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1})) (\partial_\gamma T(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}))^\top \\ &\quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n})], \\ \Xi_{3,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \partial_\gamma^2 T(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \\ &\quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n})], \\ \Xi_{4,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \\ &\quad \times \left(\partial_\gamma T(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \right) \left(\partial_\gamma T(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \right)^\top \phi''_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n})]. \end{aligned}$$

Based on the assumptions (\mathcal{B}_1) - (\mathcal{B}_5) , (\mathcal{B}'_1) - (\mathcal{B}'_3) and using the same techniques, we can prove that $(1/\sqrt{n}) \left\| \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)}$ converges to a finite positive number.

Thus, $(1/\sqrt{n}) \|\|\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)}$ converges, almost surely to a finite random positive number.

From these results, we can conclude that

$$(\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})(\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) = o_P(1).$$

Then we can write

$$\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + (\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + o_P(1).$$

Adding and subtracting appropriate term, we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + (\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,s(n)} + \hat{\boldsymbol{\gamma}}_{0,s(n)} - \hat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + o_P(1) \\ &= \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + (\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \\ &\quad + (\hat{\boldsymbol{\gamma}}_{0,s(n)} - \hat{\boldsymbol{\gamma}}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + o_P(1). \end{aligned}$$

We start with the study of the asymptotic behavior of $(\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$. Note that,

$$\sqrt{n} (\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,s(n)})^\top = \sqrt{s(n)} (\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,s(n)})^\top \sqrt{\frac{n}{s(n)}} = o_P(1).$$

We have

$$(\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) = \sqrt{n} (\boldsymbol{\gamma}_0 - \hat{\boldsymbol{\gamma}}_{0,s(n)})^\top \frac{1}{\sqrt{n}} \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}).$$

Then, to prove that this term tends to 0 as $n \rightarrow +\infty$, it suffices to prove that, as $n \rightarrow +\infty$, $\|\partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|_{p(k+1)} / \sqrt{n}$ tends in probability to a finite positive number.

We have

$$\begin{aligned} &\frac{1}{\sqrt{n}} \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) \\ &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n})] \\ &\quad - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \partial_\gamma T(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n})] \\ &= \Psi_{1,n}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) - \Psi_{2,n}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}), \end{aligned}$$

where

$$\begin{aligned} \Psi_{1,n}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n})] \\ \Psi_{2,n}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \partial_\gamma T(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \\ &\quad \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \hat{\boldsymbol{\gamma}}_{0,n})]. \end{aligned}$$

Study of the asymptotic behavior of $\psi_{1,n}(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

We have

$$\psi_{1,n}(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) = \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_{\boldsymbol{\gamma}} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n})].$$

Applying a first order Taylor expansion on $\psi_{1,n}$ around $\boldsymbol{\gamma}_0$, we obtain

$$\psi_{1,n}(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) = \psi_{1,n}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + (\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) \psi_{1,n}(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}),$$

where $\widetilde{\boldsymbol{\gamma}}_{0,n}$ lies between $\widehat{\boldsymbol{\gamma}}_{0,n}$ and $\boldsymbol{\gamma}_0$.

First of all, as $n \rightarrow +\infty$, we have

$$\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0 \xrightarrow{\mathbb{P}} 0.$$

By the assumptions (\mathcal{B}_1) - (\mathcal{B}_5) and by the ergodic theorem, following the same techniques as before, we can easily prove that

$$(\widehat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) \psi_{1,n}(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) = o_P(1).$$

Study of the asymptotic behavior of $\psi_{2,n}(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$.

From the same calculation as for $\psi_{2,n}(\boldsymbol{\rho}_0, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$, we can easily show that, as $n \rightarrow +\infty$, $\psi_{2,n}(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})$ tends to a finite constant. Then, we can write

$$\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_{\boldsymbol{\gamma}} \Pi_n(\boldsymbol{\rho}_n, \widehat{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + o_P(1).$$

Now, as in a preceding situation, we must prove that $\widehat{\boldsymbol{\gamma}}_{0,s(n)}$ belongs to the tangent space to the curve of $\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}, \boldsymbol{\beta})$ at $\widehat{\boldsymbol{\gamma}}_{0,n}$, as $n \rightarrow +\infty$.

By applying a second-order Taylor expansion on $\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_{0,s(n)}, \boldsymbol{\beta})$ in a neighborhood of $\widehat{\boldsymbol{\gamma}}_{0,n}$, we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_{0,s(n)}, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_{0,n}, \boldsymbol{\beta}) + (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n}) \partial_{\boldsymbol{\gamma}} \Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_{0,n}, \boldsymbol{\beta}) \\ &\quad + \frac{1}{2} (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_{\boldsymbol{\gamma}}^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n}), \end{aligned}$$

where $\widetilde{\boldsymbol{\gamma}}_{0,s(n)}$ lies between $\widehat{\boldsymbol{\gamma}}_{0,s(n)}$ and $\widehat{\boldsymbol{\gamma}}_{0,n}$.

Then, to prove that $\widehat{\boldsymbol{\gamma}}_{0,s(n)}$ belongs to the tangent space to the curve of $\Pi_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}, \boldsymbol{\beta})$ at $\widehat{\boldsymbol{\gamma}}_{0,n}$, it suffices to prove that

$$(\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n})^\top \partial_{\boldsymbol{\gamma}}^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) (\widehat{\boldsymbol{\gamma}}_{0,s(n)} - \widehat{\boldsymbol{\gamma}}_{0,n}) = o_P(1).$$

We have

$$\begin{aligned}\sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) &= \sqrt{s(n)}(\widehat{\gamma}_{0,s(n)} - \gamma_0 + \gamma_0 - \widehat{\gamma}_{0,n}) \frac{\sqrt{n}}{\sqrt{s(n)}} \\ &= \sqrt{s(n)}(\widehat{\gamma}_{0,s(n)} - \gamma_0) \frac{\sqrt{n}}{\sqrt{s(n)}} + \sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n}) \\ &= o_P(1) + \sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n}).\end{aligned}$$

That means, as $n \rightarrow +\infty$, $\sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})$ converges in law to a normal distribution. Also, we have

$$\begin{aligned}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta})(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) \\ = \sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \frac{1}{\sqrt{n}} \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta})(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}).\end{aligned}$$

We proved above that $(1/\sqrt{n}) \|\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta})\|_{p(k+1)}$ tends to a positive random variable. It results that $(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,s(n)}, \boldsymbol{\beta})(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) = o_P(1)$. Then, we conclude that

$$\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \gamma_{0,n}, \boldsymbol{\beta}) + o_P(1),$$

from which it follows that $\widehat{\gamma}_{0,s(n)}$ belongs to the tangent space to the curve of $\Pi_n(\boldsymbol{\rho}_n, \gamma, \boldsymbol{\beta})$ at $\widehat{\gamma}_{0,n}$.

Therefore,

$$\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1),$$

where $\{s(n)\}_{n \geq 1}$ is a sequence of positive integers such that, as $n \rightarrow +\infty$, $n/s(n) \rightarrow 0$.

□

Returning to the proof of Proposition 4.3.1 and by using Lemma 4.3.1, we have

$$\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1).$$

We can write

$$\Pi_n(\boldsymbol{\rho}_{s(n)}, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \widehat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1).$$

Using Proposition 4.2.2, we have

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \gamma_0, \boldsymbol{\beta}) + o_P(1).$$

Also, we have

$$\begin{aligned} & \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) - \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_{s(n)}, \boldsymbol{\beta}) \\ &= \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) - \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) - \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_{s(n)}, \boldsymbol{\beta}) \\ &= \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) - \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1). \end{aligned}$$

Finally, using the above results, we can write

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1)$$

□

For testing H_0 against $H_{\boldsymbol{\beta}}^{(n)}$, for any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, we consider the following statistic

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}.$$

Theorem 5. (*Optimality*)

Assume that (\mathcal{A}_1) - (\mathcal{A}_{10}) , (\mathcal{B}_1) - (\mathcal{B}_5) and (\mathcal{B}'_1) - (\mathcal{B}'_3) hold. Let $s(n)$ be any sequence of positive integers such that $n/s(n) \rightarrow 0$ as $n \rightarrow +\infty$. Then, for any $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$, we have

- i. Under H_0 , as $n \rightarrow +\infty$, $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) \xrightarrow{D} \mathcal{N}(0, 1)$.
- ii. Under local alternatives hypothesis $H_{\boldsymbol{\beta}}^{(n)}$, at level of significance $\alpha \in]0, 1[$, the asymptotic power of the test based on the statistic $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})$ is $\mathcal{P}_{k,\tau^k} = 1 - \Phi(z_\alpha - \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}))$, where z_α is the $(1 - \alpha)$ -quantile of a standard Gaussian distribution with cumulative distribution function Φ .
- iii. The test based on the statistic $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})$ is locally asymptotically optimal.

Proof: The test statistic is

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) = \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})}.$$

We proved in Proposition 4.3.1, that

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1).$$

We also proved that, as $n \rightarrow +\infty$,

$$\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) \xrightarrow{\mathbb{P}} \pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}).$$

Thus, under H_0 , as $n \rightarrow +\infty$, we have

$$\lim_{n \rightarrow +\infty} \frac{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} = 1,$$

and

$$\begin{aligned}
 \mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta}) &= \frac{\Pi_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} \\
 &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1)}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} \\
 &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} + \frac{1}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} \times o_P(1) \\
 &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} \times \frac{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} + \frac{1}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} \times o_P(1) \\
 &= \frac{\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})} \times \frac{\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} + \frac{1}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} \times o_P(1) \\
 &\xrightarrow{\mathbb{P}} \mathcal{T}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}).
 \end{aligned}$$

The convergence of $\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})$ to $\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$ remains true under the local alternative $H_{\boldsymbol{\beta}}^{(n)}$ by contiguity. Then, under $H_{\boldsymbol{\beta}}^{(n)}$ and as $n \rightarrow +\infty$, using Lecam's third lemma, we have

$$\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \frac{\widehat{\Pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})}{\widehat{\pi}_n(\boldsymbol{\rho}_{s(n)}, \widehat{\boldsymbol{\gamma}}_{0,s(n)}, \boldsymbol{\beta})} \xrightarrow{\mathcal{D}} \mathcal{N}(\pi(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}), 1).$$

Then, we can say that the power of the test remains the same.

With this, we have established that

- The central sequence based on the estimated version has no effect to the power of the test.
- The statistic based on the estimated version is asymptotically equivalent to that based on the true parameter.

Then we conclude that the test remains optimal in the case where the nuisance parameter and γ parameters are unknown.

□

4.4 Conclusion

In this chapter, we have treated the case where the functions T and V have known forms and the parameters are assumed to be unknown. We have assumed the existence of

consistent and asymptotically normal estimators for $\boldsymbol{\rho}_0$ and $\boldsymbol{\gamma}_0$ and we have proved the asymptotic equivalence between the central sequence based on the estimators of these parameters and that based on the true parameters. Furthermore, we have proved theoretically the local asymptotic optimality of our test, and we have derived an explicit expression for its local power. The results of this chapter will be illustrated by a simulation experiment in chapter 6. They will also be used to change-point study in the next chapter.

Chapter 5

Application to change-point study

5.1 Introduction

In this chapter, we apply the results of Chapter 4 to change-point study. We follow the idea of Ltaifa (2021) and Ngatchou-Wandji and Ltaifa (2021). However, the sequential algorithm that we propose here is different from that proposed in these papers. The algorithm proposed there are based on the use of the chronogram from which potential change-points are found. The locations are then found around these change locations. This presents an exhaustive search which is known to be time-consuming.

The new algorithm proposed here finds change-points at blind. It is more automatic and do not require the eyes of the user. However, it is more exhaustive and therefore, time-consuming.

This chapter is organized as follows. In Section 5.2 and 5.3, we give a brief review of some existing algorithm on change-points study. We describe our algorithm in the last section.

5.2 Some algorithms for change-point detection and the location estimation

Many papers study strategies for change-points detection and/or for estimating their locations. Here, we review some of them, that propose algorithms based on these strategies for change-points study.

5.2.1 Automatic procedure based on parametric autoregressive model (*Auto-PARM*)

In Davis et al. (2006) an automatic procedure called *Auto-PARM* is proposed for modeling a nonstationary time series by segmenting the series into blocks of different autoregressive processes.

Consider k_j as the change-point between the j th and the $(j + 1)$ th *AR* processes, with $j = 1, \dots, m$, $k_0 = 1$ and $k_m < T$. Thus, the j th piece of the series is modelled as:

$$X_t = x_{t,j}, \quad k_{j-1} \leq t < k_j,$$

where $\{x_{t,j}\}$ is an $AR(p_j)$ process:

$$x_{t,j} = \gamma_j + \phi_{j,1}x_{t-1,j} + \dots + \phi_{j,p_j,j} + \sigma_j \varepsilon_t,$$

with $\theta_j := (\gamma_j, \phi_{j,1}, \dots, \phi_{j,p_j,j}, \sigma_j^2)$ the parameter vector corresponding to this $AR(p_j)$ process, and $\{\varepsilon_t\}$ as iid sequence with mean 0 and variance 1. This model assumes that the behavior of the time series is changing at various times. Such a change might be a shift in the mean, a change in the variance, and/or a change in the dependence structure of the process.

The best segmentation is the one that makes the maximum compression of the data possible measured by the *MDL* principle of Rissanen (1998)). *MDL* is defined as

$$MDL(m, k_1, \dots, k_m, p_1, \dots, p_{m+1}) = \log m + (m + 1) \log T + \sum_{j=1}^{m+1} \log p_j + \sum_{j=1}^{m+1} \frac{p_j + 2}{2} \log T_j + \sum_{j=1}^{m+1} \frac{T_j}{2} \log(2\pi \hat{\sigma}_j^2),$$

where m is the number of the change-points located at k_1, \dots, k_m , T_j is the number of observations in each segment j , p_j is the order of the autoregressive model fitted to the segment j , and $\hat{\sigma}_j^2$ is the Yule Walker estimator of σ_j^2 (Brockwell et al. (1991)).

5.2.2 Smooth localized complex exponentials (*Auto-SLEX*) functions approach

In Adak (1998) and Donoho et al. (1998), the segmentation is performed by using a cost function based on the spectrum, called the evolutionary spectrum, because the calculation is done by the spectrum of each stationary interval. Ombao et al. (2002) created

SLEX vectors (Smooth Localized complex Exponential) which are calculated by applying a projection operator on the Fourier vectors, to get a basis that is simultaneously orthogonal and localized in time and frequency and is useful to compute the spectrum of nonstationary time series.

The cost function of the block $S_j = [k_j, k_{j+1}]$ is given by

$$\text{cost}(S) = \sum_{S_j} \log(\hat{\alpha}_{S_j}) + \beta \sqrt{m_j},$$

where $\hat{\alpha}_{S_j}$ is the *SLEX* periodogram, β is a penalty parameter generally equal to 1 (see Donoho et al. (1998)), and m_j is the number of breaks in the block. The cost for a particular segmentation of the time series is the sum of the costs at all the blocks defining that segmentation is the one having the smallest cost.

5.2.3 Informational approaches

Information criteria that are used for the measure of goodness-of-fit of a model, can also be used to detect change-points and estimate their locations. The most popular of the information criteria is the Akaike Information Criteria (*AIC*), introduced in 1973 for model selection in statistics. This criterion has found many applications in time series, outliers detection, robustness, and regression analysis. *AIC* is defined as

$$AIC = T \log(\hat{\sigma}_{MV}^2) + 2p,$$

where $\hat{\sigma}_{MV}^2$ is the maximum likelihood estimator of σ^2 , and p is the number of free parameters.

A model that minimizes the *AIC* is considered the appropriate model.

Another information criterion was introduced in Schwarz et al. (1978) and is commonly referred to as *BIC* or *SIC*. The fundamental difference with the *AIC* is the penalization function, which penalizes more the number of model parameters and leads to an asymptotically consistent estimate of the order of the true model.

BIC is defined as:

$$BIC = T \log(\hat{\sigma}_{MV}^2) + p \log(T),$$

where $\hat{\sigma}_{MV}^2$ is the maximum likelihood estimator of σ^2 , p is the number of free parameters, and T is the length of time series.

5.2.4 Binary segmentation

Another method for change detection is the binary segmentation (BS). In case where the number of breaks is known, and in a multivariate space, Vostrikova (1981) proves the consistency of this method. The test used is the CUSUM test. When the test detects the break, the proposed stopping criterion is not easy. For this reason, Olshen et al. (2004) introduce a modification of the (BS) called circular segmentation method (CBS). They test the hypothesis of no break in the mean of the data against the hypothesis of the presence of breaks at the ends of the data.

Fryzlewicz (2014) proposes a new technique for estimating the number of breaks and their locations in a univariate and non-stationary stochastic series. This method based on CUSUM test is called Wild Binary Segmentation (WBS). Fryzlewicz (2014) proves its consistency, and shows that it works even when the breaks are very close and/or of small amplitude, unlike the BS method. Fryzlewicz (2014) proposes two stopping criterions for WBS. The first is based on the critical threshold, and the second one based on the "strengthened Schwarz Information Criterion (sSIC)". He applies his method for detecting breaks in financial data, more precisely, in S&P 500 indexes.

Baranowski et al. (2019) introduce a new method called NOT, which is a modification of WBS. They apply their method to real data such as the global anomaly temperature and the housing data in London. We can use the NOT method with "R" program *not* (see Baranowski et al. (2019)). The NOT method estimates the number of breaks in piece-wise stationary time series. Baranowski et al. (2019) use the same model as Fryzlewicz (2014) given by the following stochastic equation

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

where $(Y_t)_t$ stands for the series of observations, f_t is an unknown function that verifies some conditions, $(\varepsilon_t)_t$ is a centered white noise and n is the number of observations.

When f_t is piece-wise constant, they consider the breaks in f_t , and they study the breaks in f'_t when f_t is linear and piece-wise continuous. They study the discontinuity of f_t when f_t is discontinuous, and they study the breaks in f'_t when f_t is piece-wise linear without the continuity assumption. The NOT method has low computational complexity and precision. NOT method is one of the most efficient method for the estimation and the localization of the breaks. The calculation of NOT method is fast which gives it an advantage regarding WBS method of Fryzlewicz (2014), SMVCE method of Frick et al. (2014) and the FDRSeg method of Li et al. (2016).

Furthermore, we can see many other update for BS method, see Cho and Fryzlewicz (2020), Anastasiou and Fryzlewicz (2022). We can see a detailed discussion of these

methods in Ltaifa (2021).

5.3 Some recent methods

Some more recent methods focus on the study of change detection at the beginning and the end of a time series, and on weak change detection study.

5.3.1 Horváth et al. (2020)

Horváth et al. (2020) propose a new class of change-point test statistics that utilizes a modified version of CUSUM test. It is called weighting and trimming scheme for the cumulative sum (CUSUM) process inspired by Rényi (1953). Let X_1, \dots, X_n denoted a series of observations, n the sample size and consider the following shifted model

$$X_t = \mu_t + e_t, \quad 1 \leq t \leq n,$$

where $\mathbb{E}(e_t) = 0$. The author aim to test

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_n \text{ against } H_A : \mu_1 = \mu_2 = \dots = \mu_{t^*} \neq \mu_{t^*+1} = \dots = \mu_n,$$

with some $1 < t^* < n$. For that, Horváth et al. (2020) generalize the statistic of Rényi (1953) and the test statistic considered has the following expression

$$D_n = D_n(t_n) = \max_{t_n \leq t \leq n-t_n} \left| \frac{1}{t} \sum_{s=1}^t X_s - \frac{1}{n-t} \sum_{s=t+1}^n X_s \right|,$$

where the standard CUSUM has the following expression

$$CUSUM = \frac{1}{\sqrt{n}} \max_{1 \leq t \leq n} \left| \sum_{s=1}^t X_s - \frac{t}{n} \sum_{s=1}^n X_s \right|.$$

This new class of statistics possesses superior power compared to traditional change-point statistics based on the CUSUM process when the change-point is near the beginning or the end of the sample. On the website <https://github.com/haeran-cho/wem.gsc>, the authors propose an "R" code for detecting the change-points and estimating their locations.

5.3.2 Ngatchou-Wandji and Ltaifa (2021)

Now, we present the methods in Ngatchou-Wandji and Ltaifa (2021). These are for the model (1.1) with the autoregression function (the conditional mean) of the form

$$T(\boldsymbol{\rho}_0 + \boldsymbol{\gamma} \odot \omega(t), \mathbf{X}_{t-1}) = U(\mathbf{X}_{t-1}) + \boldsymbol{\gamma} \odot \omega(t),$$

for some smooth function U . The authors give an explicit expression to the local power of the test studied, which is a particular case of that obtained in this thesis. This power is then used for detecting changes and estimating their locations. For the change detection, an estimated version of the power is obtained as a function of potential change-points to which several values are attributed. If the obtained power is significantly different from the nominal level of the test, then the data contains at least one change-point. A sequential algorithm is further constructed for estimating the locations of all possible changes. Note that the potential change-points locations are determined from the chronogram.

5.4 Algorithm for break detection and location estimation

5.4.1 The algorithm

There is at least one change in the time series at hand if the parameter of its distribution is constant up to the time index τ_1 when it changes for the first time. Considering the test constructed in this paper for testing the null hypothesis of no change against at least one change, for any model from the class (1.1) adjusted to the time series under study, all the components of $\boldsymbol{\gamma}_0$ in the null model will always be taken to be the same, say $\boldsymbol{\rho}_1$, the parameter of the stationary model on $[\tau_0, \tau_1)$, while the first component of $\boldsymbol{\beta}$ would be nil.

The test constructed in this work can do more than testing no change against at least one change. To understand this, assume changes have been detected in the data by a given method, and their locations have been estimated. Our test can serve as a screening method for finding possible missing changes by this method. In this situation, one can assume the changes already detected as well as their locations as known. With this, all the components of $\boldsymbol{\gamma}_0$ will no longer be the same, and some of the τ_j 's in the model would be considered as known. Thus, our test can be used for testing the null hypothesis of ι changes against at least $\iota + 1$ changes, for some given $\iota \in \mathbb{N}$.

Recall that for any $k \geq 1$, \mathcal{P}_{k,τ^k} denotes the theoretical power of this test at $\tau^k = (\tau_1, \dots, \tau_k)$, with the convention that $\mathcal{P}_{0,\tau^0} = \alpha$, $\alpha \in (0, 1)$, the level of significance. Let $\zeta \in (0, .1)$ and X_1, X_2, \dots, X_m , ($m \ll n$), the m first stationary observations. With our assumptions, m may be taken to be $n_1(n)$.

Our procedure for detecting changes in the time series X_1, X_2, \dots, X_n and estimating their locations is described in the following algorithm.

Location 1 :

Put $j = 1$ and Do

(A1): Take any t between 1 and $m + j$,

so that there is a large number of indices before and after t (for example $t = [(m + j)/2]$).

Adjust model (1.1) to X_1, \dots, X_{m+j} with a potential change located at the time index t and apply the testing procedure studied.

If $|\mathcal{P}_{1,\tau^1} - \mathcal{P}_{0,\tau^0}| > \zeta$,

Put $\tau_1 = m + j$ and Go to **Location 2** (first change location estimated)

Else

Do $j = j + 1$ and Go to (A1).

Location 2 :

Consider the next h observations to X_{τ_1} :

$X_{\tau_1+1}, \dots, X_{\tau_1+h}$

Put $j = 1$ and Do

(A2): Take any t between $\tau_1 + 1$ and $\tau_1 + h + j + 1$

so that there is a large number of indices before and after t .

Adjust model (1.1) to $X_{\tau_1+1}, \dots, X_{\tau_1+h+j}$ with a potential change located at the time index t and apply the testing procedure studied.

If $|\mathcal{P}_{1,\tau^1} - \mathcal{P}_{0,\tau^0}| > \zeta$,

Put $\tau_2 = \tau_1 + h + j$ and Go to **Location 3** (second change location estimated)

Else

Do $j = j + 1$ and Go to **(A2)**.

Location i :

Let τ_{i-1} be the change location at step i-1

Put $j = 1$ and Do

(Ai): Take any t between $\tau_{i-1} + 1$ and $\tau_{i-1} + h + j$

so that there is a large number of indices before and after t

Adjust model (1.1) to $X_{\tau_{i-1}+1}, \dots, X_{\tau_{i-1}+h+j+1}$ with a potential change located at the time index t and apply the testing procedure studied

If $|\mathcal{P}_{1,\tau^1} - \mathcal{P}_{0,\tau^0}| > \zeta$,

Put $\tau_i = \tau_{i-1} + h + j$, $i = i + 1$ and Go to **Location i** (i th change location estimated)

Else

Do $j = j + 1$ and Go to **(Ai)**.

5.4.2 Comments

The change-point detection algorithm above is motivated by the idea that in the case of no change, the magnitude of any eventual change would be nil or very close to 0, and the estimated power would be close to the level of the test. Also, if there is at least one change, then one can find one τ_1 for which the value of the parameter associated to the observations before and after that time index are different. The test would then reject the null hypothesis of no break. Next, at the correct locations, the power of the test should be the largest.

Chapter 6

Practical considerations

6.1 Introduction

In this chapter, we study the finite sample properties of our results. We apply them to simulated data using "R" software. We first study the power of our test as a function of the magnitudes of the breaks, when the breaks locations are assumed to be known. Next, using an estimated version of the power where the magnitudes are replaced by their estimators, we apply the algorithm described in Chapter 5 for detecting change-points and estimating their locations.

6.2 Numerical study

In this part, we use the following particular CHARN model

$$X_t = \rho_{0,1} + \gamma_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \gamma_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} e^{\left(\rho_{0,3} + \gamma_{0,3} + \frac{\beta_{1,3}}{\sqrt{n}} \right) X_{t-1}^2} + (\theta_1 + \theta_2 X_{t-1}^2)^{\frac{1}{2}} \varepsilon_t, \quad (6.1)$$

where n denotes the number of observations and $(\varepsilon_t)_t$ is a standard white noise. Here, $\boldsymbol{\rho}_0 = (\rho_{0,1}, \rho_{0,2}, \rho_{0,3}) \in \mathbb{R}^3$, $\boldsymbol{\gamma}_0 = (\gamma_{0,1}, \gamma_{0,2}, \gamma_{0,3}) \in \mathbb{R}^3$, $\boldsymbol{\beta}_1 = (\beta_{1,1}, \beta_{1,2}, \beta_{1,3}) \in \mathbb{R}^3$ and $\boldsymbol{\theta} = (\theta_1, \theta_2)$ is assumed to be known.

The number of replications is $N = 1000$ and the nominal levels considered are $\alpha = 1\%$, 5% and 10% .

First of all, we study the local power of the test constructed in chapters 3 and 4. For

fixed breaks, this power is a function of $\boldsymbol{\beta}$. Next, we apply our methods to change-point study.

6.2.1 Power study

Case where $\gamma_0 = 0$

We start with the case $\gamma_0 = 0$, for $\boldsymbol{\rho}_0 = (\rho_{0,1}, 0, 0)$, $\theta_1 = 1$ and $\theta_2 = 0$. It corresponds to an $AR(1)$ model, which reduces to

$$X_t = \rho_{0,1} + \gamma \odot \omega(t) + \varepsilon_t.$$

In this case, the hypotheses to be tested are:

$$H_0 : \gamma = 0 \quad \text{against} \quad H_1 : \gamma = \frac{\boldsymbol{\beta}}{\sqrt{n}},$$

where $\boldsymbol{\beta} = (\beta_{1,1}, \beta_{2,1}) \in \mathbb{R}^2$.

For $t = 1, \dots, 100$, and a fixed break location, for example, at $t = 40$, the simulated data, in this case, can be obtained from

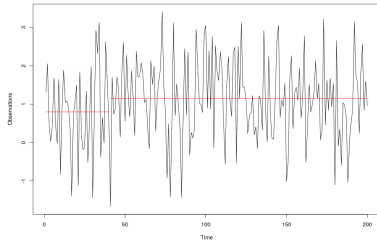
$$\begin{cases} X_t = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \varepsilon_t & t = 1, \dots, 40, \\ X_t = \rho_{0,1} + \frac{\beta_{2,1}}{\sqrt{n}} + \varepsilon_t & t = 41, \dots, 100, \end{cases}$$

where (ε_t) is a sequence of standard Gaussian white noise.

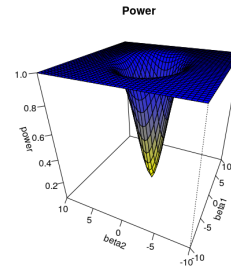
Recall the expression of the power function (3.13). We use $\hat{\pi}_n = \sqrt{\hat{\eta}_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta})}$ an estimator of π where $\hat{\eta}_n(\hat{\boldsymbol{\rho}}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \sum_{j=1}^{k+1} \hat{\alpha}_j \sum_{1 \leq h, m \leq p} \beta_{j,h} \beta_{j,m} \hat{\eta}_{j,2}^{(h,m)}(\hat{\boldsymbol{\rho}}_0, \boldsymbol{\gamma}_0)$ with $\hat{\alpha}_j = n_j(n)/n$ and $\hat{\eta}_{j,2}^{(h,m)}(\hat{\boldsymbol{\rho}}_0, \boldsymbol{\gamma}_0)$ is the empirical estimator of $\eta_{j,2}^{(h,m)}(\hat{\boldsymbol{\rho}}_0, \boldsymbol{\gamma}_0)$ given by

$$\hat{\eta}_{j,2}^{(h,m)}(\hat{\boldsymbol{\rho}}_0, \boldsymbol{\gamma}_0) = \frac{I(f)}{n_j(n)} \sum_{\tau_{j-1}}^{\tau_j} \frac{1}{V^2(\mathbf{X}_{t-1})} \frac{\partial T}{\partial \gamma_{j,h}}(\hat{\boldsymbol{\rho}}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{j,m}}(\hat{\boldsymbol{\rho}}_0, \boldsymbol{\gamma}_0, \mathbf{X}_{t-1}).$$

Here, $I(f) = 1$ since (ε_t) is a standard white noise, $n_1(n) = 40$ and $n_2(n) = 60$, $V(x) = 1$ and the partial derivative $\partial T / \partial \gamma_{j,h}$ and $\partial T / \partial \gamma_{j,m}$ are equals to 1. For $\rho_{0,1} = 0.8$, $\beta_{1,1} = 0$ and $\beta_{2,1} = 5$, the chronogram of the data is represented on Figure (6.1a).



(a) Model 6.1 with $\rho_{0,2} = \rho_{0,3} = \theta_2 = \gamma_{0,2} = \gamma_{0,3} = 0$



(b) Power of the test.

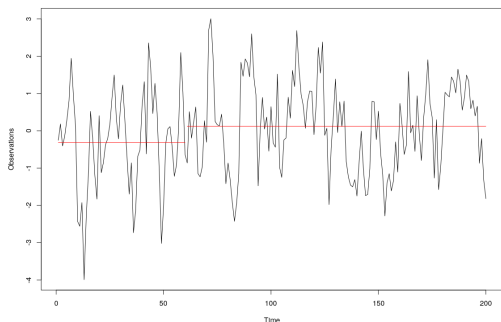
Figure 6.1: Power of test against a shifted white noise.

It is easy to see that the mean has a break at $t = 40$. Figure (6.1b) presents the power of the test as a function of $\beta_{1,1}$ and $\beta_{2,1}$. It is clear that the power increases when $|\beta_{1,1}|$ and $|\beta_{2,1}|$ increase. When $\beta_{1,1} = \beta_{2,1} = 0$, the power is close to $1 - \Phi(Z_{0.05} - 0) \simeq 0.05$. Next, we study the case where $\boldsymbol{\rho}_0 = (0, \rho_{0,2}, 0)$, $\theta_1 = 1$ and $\theta_2 = 0$. The test concerns the change in the autocorrelation. Here, the model has the following expression

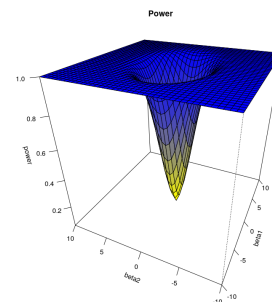
$$\begin{cases} X_t = \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t & t = 1, \dots, 40, \\ X_t = \left(\rho_{0,2} + \frac{\beta_{2,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t & t = 41, \dots, 100, \end{cases}$$

where (ε_t) is a sequence of standard Gaussian white noise.

In this case, the partial derivatives $\partial T / \partial \gamma_{j,h}(\boldsymbol{\gamma}_0, X_{t-1})$ and $\partial T / \partial \gamma_{j,m}(\boldsymbol{\gamma}_0, X_{t-1})$ are equals to X_{t-1} . For $\rho_{0,2} = 0.5$, the power as a function of magnitudes change $\beta = (\beta_{1,2}, \beta_{2,2})$ is presented in Figure (6.2b).



(a) Model 6.1 with $\rho_{0,1} = \rho_{0,3} = \theta_2 = \gamma_{0,1} = \gamma_{0,3} = 0$.



(b) Power of the test.

Figure 6.2: Power of test in case of change in autocorrelation.

Here, $\boldsymbol{\beta} = (\beta_{1,2}, \beta_{2,1})$. The power as a function of $\boldsymbol{\beta}$, for $\beta_{1,2}$ and $\beta_{2,2}$ varying between -10 and 10 , is given in Figure (6.2b). It is clear from this figure that the power increases when $|\beta_{1,2}|$ and $|\beta_{2,2}|$ increase. In the third case, we take $\boldsymbol{\rho}_0 = (\rho_{0,1}, \rho_{0,2}, 0)$, $\theta_1 = 1$ and $\theta_2 = 0$. This corresponds to the following model

$$\begin{cases} X_t = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t & t = 1, \dots, 40, \\ X_t = \rho_{0,1} + \frac{\beta_{2,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{2,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t & t = 41, \dots, 100, \end{cases} \quad (6.2)$$

where (ε_t) is a sequence of iid standard Gaussian variables.

Here, $\partial T / \partial \gamma_{j,1}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, X_{t-1}) = 1$ and $\partial T / \partial \gamma_{j,2}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, X_{t-1}) = X_{t-1}$. In this case, $\boldsymbol{\beta}$ have 4 components ($\beta_1 = (\beta_{1,1}, \beta_{1,2})$ and $\beta_2 = (\beta_{2,1}, \beta_{2,2})$). To obtain a clear figure, we fix one component of β_1 and one of β_2 , and we study the power as a function of the remaining components. With this, it results 4 figures. Figure (6.3) shows that the power converges to 1 when the magnitude of change increases.

Now, we take $\theta_1 = 1$ and $\theta_2 = 0.02$. Then, we have $V(x) = \sqrt{1 + 0.02x^2}$. This corresponds to the following model

$$\begin{cases} X_t = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \sqrt{1 + 0.02X_{t-1}^2} \varepsilon_t & t = 1, \dots, 40, \\ X_t = \rho_{0,1} + \frac{\beta_{2,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{2,2}}{\sqrt{n}} \right) X_{t-1} + \sqrt{1 + 0.02X_{t-1}^2} \varepsilon_t & t = 41, \dots, 100, \end{cases} \quad (6.3)$$

where (ε_t) is a sequence of standard Gaussian white noise.

Here, $\boldsymbol{\rho}_0 = (\rho_{0,0}, \rho_{0,1}) \in \mathbb{R}^2$, $\boldsymbol{\beta} = (\beta_1, \beta_2) \in \mathbb{R}^{2 \times 2}$ with $\beta_1 = (\beta_{1,1}, \beta_{1,2})$ and $\beta_2 = (\beta_{2,1}, \beta_{2,2})$, $V(X_t) = \sqrt{1 + 0.02X_{t-1}^2}$, $\partial T / \partial \gamma_{j,1}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, X_{t-1}) = 1$ and $\partial T / \partial \gamma_{j,2}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, X_{t-1}) = X_{t-1}$. As before, following the strategy explained above, we obtain Figure 6.4.

Case where $\boldsymbol{\gamma}_0$ and f are unknown

As said before, in practice, $\boldsymbol{\gamma}_0$ and f are unknown and must be estimated. The theoretical study of this case seems long and cumbersome. So, we do not try to tackle it here. However, as in Ltaifa (2021) and Ngatchou-Wandji and Ltaifa (2021), we don't study theoretically the case where f is unknown. We use the least-squares method to estimate $\boldsymbol{\gamma}_0$ and we estimate f by Parzen-Rosenbatt estimator (see Parzen (1962)) by

$$\hat{f}_n(x) = \frac{1}{nh_n^2} \sum_{t=1}^n K \left(\frac{x - \hat{\varepsilon}_t}{h_n} \right), \quad x \in \mathbb{R},$$

where

$$\widehat{\varepsilon}_t = \frac{X_t - T(\boldsymbol{\rho}_n + \widehat{\boldsymbol{\gamma}}_n \odot \omega(t), \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})},$$

where $\boldsymbol{\rho}_n$ and $\widehat{\boldsymbol{\gamma}}_n$ denote respectively the least-squares estimator of $\boldsymbol{\rho}_0$ and $\boldsymbol{\gamma}_0$, h_n is the smoothing parameter and K the kernel function having the following properties:

- i. $K(x) > 0$ for any $x \in \mathbb{R}$ (Positivity).
- ii. $\int_{\mathbb{R}} K(x)dx = 1$ (density).
- iii. $\int_{\mathbb{R}} xK(x)dx = 0$ (symmetry).

Now we calculate the power of the test using these estimators. We choose a Gaussian kernel K and the band width $h_n = \sigma_n^2.n^{-1/5}$ where σ_n^2 stands for the sample variance. We consider a sample of $n = 100$ observations, $n_1(n) = 40$, $n_2(n) = 60$ and we generate the observations using model (6.1). The results for $\rho_{0,1} = 0.8$, $\rho_{0,2} = 0.2$, $\rho_{0,3} = 0$, $\theta_1 = 1$ and $\theta_2 = 0.02$ at level of significance 5% are given by Figure (6.5), and those at level of significance 1% by Figure (6.6). It is clear that the test is more powerful in case of exponential noise than in case of Gaussian noise.

It is important to know that we obtain the same result if we choose another kernel like Epanechnikov, Uniform kernel, quadratic kernel, *etc.*

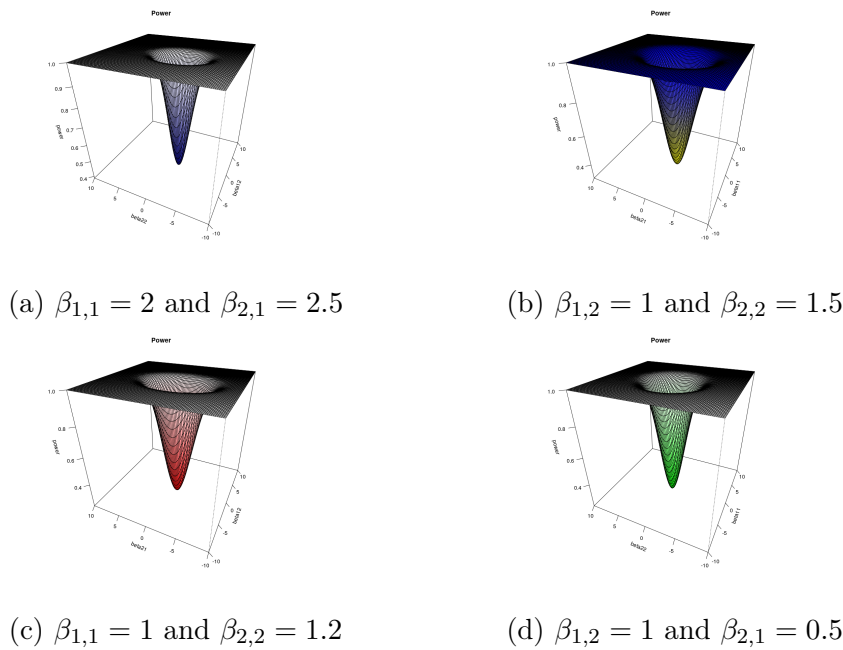
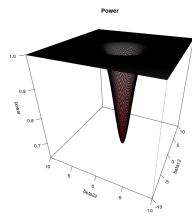
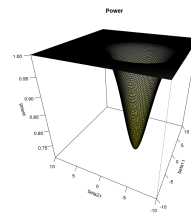


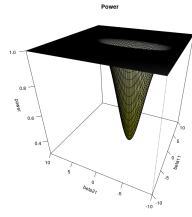
Figure 6.3: Power against an AR(1).



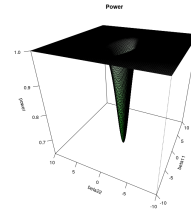
(a) $\rho_0 = (0.6, 0.2), \beta_{1,1} = 2, \beta_{2,1} = 2.5$



(b) $\rho_0 = (0.6, 0.2), \beta_{1,2} = 1$ and $\beta_{2,2} = 1.5$

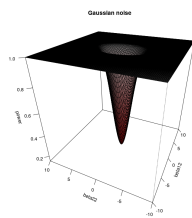


(c) $\rho_0 = (0.6, 0.2), \beta_{1,1} = 1$ and $\beta_{2,2} = 1.5$

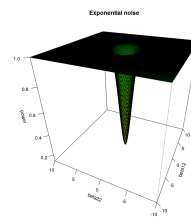


(d) $\rho_0 = (0.6, 0.2), \beta_{1,2} = 1$ and $\beta_{2,1} = 1.5$

Figure 6.4: Power against an AR(1).

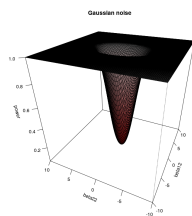


(a) $\rho_{0,1} = 0.6, \rho_{0,2} = 0.2, \rho_{0,3} = 0, \theta_1 = 1$ and $\theta_2 = 0.02$

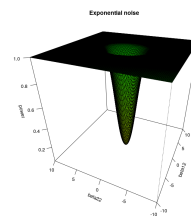


(b) $\rho_{0,1} = 0.6, \rho_{0,2} = 0.2, \rho_{0,3} = 0, \theta_1 = 1$ and $\theta_2 = 0.02$

Figure 6.5: Level of significance 5%.



(a) $\rho_{0,1} = 0.6, \rho_{0,2} = 0.2, \rho_{0,3} = 0, \theta_1 = 1$ and $\theta_2 = 0.02$



(b) $\rho_{0,1} = 0.6, \rho_{0,2} = 0.2, \rho_{0,3} = 0, \theta_1 = 1$ and $\theta_2 = 0.02$

Figure 6.6: Level of significance 1%.

6.2.2 Breaks detection

In this section, we apply the strategy presented in chapter 5 to detect the breaks and estimate their locations. In other words, we want to check if our method is able to detect and localize the breaks or not. We first treat the case where there is no break. Next, we study the case of multiple breaks in the data generated by many particular cases of model (6.1).

No break in the data

We start by calculating the asymptotic local power given by (3.13) in case where there is no break in the model (6.1), that means the case $k = 0$. We consider a sample of $n = 200$ observations generated from model (6.1), for $\gamma_0 = 0$ and f a standard Gaussian density. For $\rho_{0,1} = 0.5, \rho_{0,2} = \rho_{0,3} = \theta_2 = 0$ and $\theta_1 = 1$, the asymptotic local power of our test with different levels of significance is plotted on Figures 6.8 and 6.7. We can see there that the local power does not exceed 0.1012 when $\alpha = 10\%$, does not exceed 0.0507 when $\alpha = 5\%$, and does not exceed 0.01018 when $\alpha = 1\%$. Then, for any $\alpha = 1\%, 5\%$, and 10% , for the values $\zeta = 0.0002, 0.008$ and 0.002 , we keep the null hypothesis and we conclude that there is no break in the data.

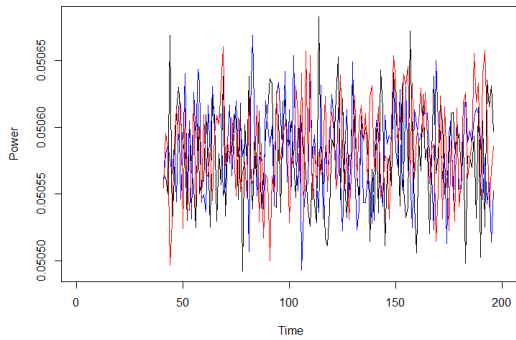
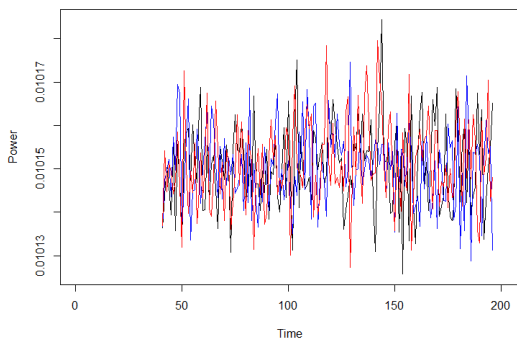
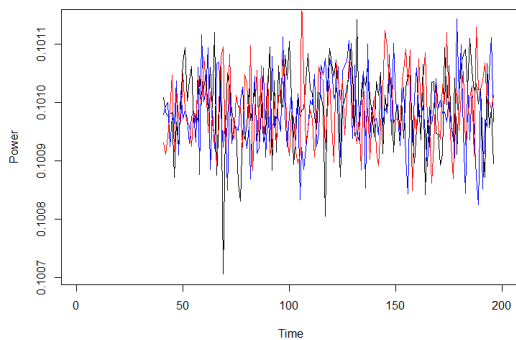


Figure 6.7: No break ($\alpha = 5\%$).



(a) No break ($\alpha = 1\%$).



(b) No break ($\alpha = 10\%$).

Figure 6.8: No break in the data.

One single break

We consider in this part the model (6.1) with $n = 200$ and $\gamma_0 = 0$ and we treat two situations of breaks. First, the break happens in the parameter $\rho_{0,1}$ which induces a change in the mean. Second, the break happens in $\rho_{0,2}$ inducing a change in the autocorrelation.

First case

We take $\rho_{0,2} = \rho_{0,3} = \theta_2 = 0$, $\theta_1 = 1$, $\beta_{1,1} = 1$ and f a standard Gaussian density. The corresponding model is the following shifted white noise:

$$X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \varepsilon_t. \quad (6.4)$$

The data are obtained for $\rho_{0,1} = 0$ from the following shifted white noise model

$$\begin{cases} X_{t-1} = \varepsilon_t, & t = 1, \dots, \tau_1 \\ X_{t-1} = \frac{\beta_{1,1}}{\sqrt{n}} + \varepsilon_t, & t = \tau_1 + 1, \dots, n. \end{cases}$$

The power computed at any location is significantly different from the level of significance α . The problem now is to estimate the location of τ_1 . The data are obtained for different times of breaks: $\tau_1 = 40, 60, 100, 150$. For $\alpha = 5\%$, the asymptotic local power is plotted on Figure (??); for $\alpha = 1\%$, on Figure (??) and for $\alpha = 10\%$ on Figure (??).

From Figure (??), we can see that the power of the test is maximal at the corresponding instant of break τ_1 ; that means, the power of the test considered is a good tool for the break detection in the mean. For $\rho_{0,1} = 0.5$ and $n = 200$, Table 6.1 shows different estimation of break times for different values of $\beta_{1,1}$.

$\tau_1 \backslash \beta_{1,1}$	1	1.5	2	2.5	4
60	61	60	59	60	60
80	80	82	80	80	79
100	99	101	100	100	100
140	138	141	140	140	140
185	183	183	185	184	185

Table 6.1: Estimation of break location.

Second case.

Here, we start by taking $\rho_{0,1} = \rho_{0,3} = \theta_2 = 0$, $\theta_1 = 1$ and f a standard Gaussian density. This corresponds to the following model

$$X_{t-1} = \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t.$$

We are thus studying the break in the autocorrelation. The data are obtained from the following model

$$\begin{cases} X_{t-1} = \rho_{0,2} X_{t-1} + \varepsilon_t, & t = 1, \dots, \tau_1 \\ X_{t-1} = \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, & t = \tau_1 + 1, \dots, n. \end{cases}$$

The power of the test is larger than the level of significance α , at any time index. The problem now is to estimate τ_1 , the location of the break. The data are obtained for different times of break. Here, we need sufficient observations at the beginning to achieve the stationarity of the first observations as explained in Chapter 5. For 1000 replications, $n = 200$, $\rho_{0,2} = 0.3$ and $\tau_1 = 100, 150, 170$, and for $\alpha = 5\%$, we obtain Figure (6.9).

On Figure (6.9), we can see that the power of the test is maximal at the corresponding instant of break, meaning that our method is a good tool for break detection in the autocorrelation.

For $\rho_{0,2} = 0.3$ and $n = 200$, Table 6.2 shows different estimations of break detection for different values of $\beta_{1,2}$.

$\tau_1 \backslash \beta_{1,2}$	3	4	5	-6	8
60	62	61	59	60	60
80	78	81	79	80	80
100	97	103	101	101	100
140	142	141	138	140	140
185	183	184	187	186	185

Table 6.2: Estimation of the break location in the autocorrelation.

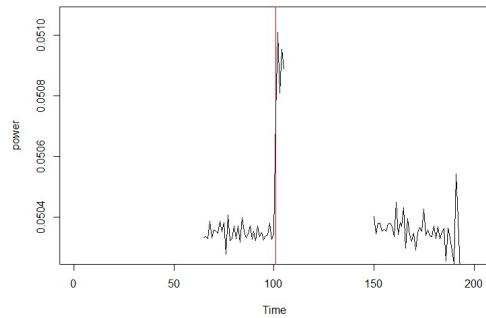


Figure 6.9: Estimation of the break location for $\tau_1 = 100$ in a class of AR(1) models (first case).

Third case

Here, we take $\rho_{0,3} = \theta_2 = 0$, $\theta_1 = 1$ and f is a standard Gaussian density. In this case, we are studying the break the in parameters of the following model

$$X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, \quad t = 1, \dots, n. \quad (6.5)$$

The data are obtained from the following model

$$\begin{cases} X_{t-1} = \rho_{0,1} + \rho_{0,2}X_{t-1} + \varepsilon_t, & t = 1, \dots, \tau_1, \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, & t = \tau_1 + 1, \dots, n, \end{cases}$$

where (ε_t) is a sequence of standard Gaussian white noise. We consider different values of τ_1 and we monitor the power of test to evaluate if it can detect the break. We treat many cases. First, we assume that $\beta_{1,2} = 0$, that means we verify if our test can detect breaks in a class of models (6.5) when the break happens only at $\rho_{0,1}$. That means, the data are obtained by the following model

$$\begin{cases} X_{t-1} = \rho_{0,1} + \rho_{0,2}X_{t-1} + \varepsilon_t, & t = 1, \dots, \tau_1 \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \rho_{0,2}X_{t-1} + \varepsilon_t, & t = \tau_1 + 1, \dots, n, \end{cases}$$

For $n = 200$, $\rho_{0,1} = \rho_{0,3} = \theta_2 = 0$, $\rho_{0,2} = 0.5$, $\theta_1 = 1$ and $\alpha = 5\%$, Figure 6.10 shows the power for different values of τ_1 . Table 6.4 shows the estimation of τ_1 obtained by our method.

From Figure 6.10, we see that the power of the test is maximal around the correct instant of break. This shows that our method is a good tool for breaks detection in case of AR(1) models when a break occurs in $\rho_{0,1}$.

$\tau_1 \backslash \beta_{1,1}$	1	-1	1.5	2	-2
100	101	100	99	101	99
130	129	129	130	131	131
155	156	155	155	156	155

Table 6.3: Estimation of the break locations for many different τ_1 and different values of $\beta_{1,1}$.

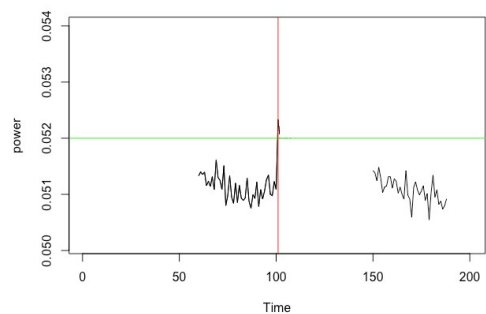


Figure 6.10: Estimation of the break location for $\tau_1 = 100$ in a class of AR(1) models (second case).

Now, we generalize the previous case and we study the break at any of the parameters of the following model

$$\begin{cases} X_{t-1} = \rho_{0,1} + \rho_{0,2}X_{t-1} + \varepsilon_t, & t = 1, \dots, \tau_1, \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, & t = \tau_1 + 1, \dots, n, \end{cases}$$

The results for $\tau_1 = 100, 140, 160, 185$, $\rho_{0,1} = 0.5$, $\rho_{0,2} = -0.2$, $\beta_{2,1} = -3$ and $\beta_{2,2} = 2$, are given by Figure 6.11. For $n = 200$, $\alpha = 5\%$, $\tau_1 = 80, 100, 120, 140, 160, 185$, and for different values of $\beta = (\beta_{1,1}, \beta_{1,2})$, the estimation of the break location is given in Table 6.4.

$\tau_1 \backslash \begin{pmatrix} \beta_{1,1} \\ \beta_{1,2} \end{pmatrix}$	$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 3 \end{pmatrix}$	$\begin{pmatrix} -3 \\ 2 \end{pmatrix}$
80	78	81	82	80	80
100	99	99	102	101	100
120	122	121	119	120	121
140	137	138	141	140	139
160	155	156	162	161	161
185	179	179	179	180	180

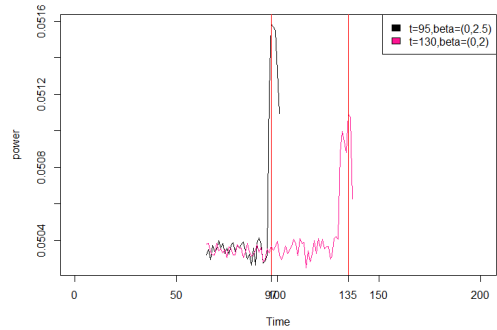


Table 6.4: Estimation of break location for $n = 200$, $\rho_{0,1} = 0.5$, $\rho_{0,2} = -0.2$, $\rho_{0,3} = \theta_2 = 0$, $\theta_1 = 1$ and $\alpha = 5\%$.

Figure 6.11: Estimation of break location for different τ_1 in a class of AR(1) model (third case).

Case of two breaks ($k = 2$).

The data to be studied are obtained from the following model

$$\begin{cases} X_t = \rho_{0,1} + \varepsilon_t, & t = 1, \dots, \tau_1, \\ X_t = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \varepsilon_t, & t = \tau_1 + 1, \dots, \tau_2, \\ X_t = \rho_{0,1} + \frac{\beta_{2,1}}{\sqrt{n}} + \varepsilon_t, & t = \tau_2 + 1, \dots, n, \end{cases} \quad (6.6)$$

where (ε_t) is a sequence of standard Gaussian white noise. Our aim is to detect and locate the changes for which the number and the locations are assumed to be unknown. We consider different instants of break $(\tau_1, \tau_2) = (120, 210)$ and $(140, 240)$. For 5000 re-olocations, $n = 300$, $\rho_{0,1} = 0.2$, for different values of $\beta_{1,1}$ and $\beta_{2,1}$, and different thresholds corresponding to the choice of ζ , the results of detecting the breaks and estimating their

locations are given on Table 6.5. It is easy to see that the more ζ is high, the more the estimation of the breaks are far from the exact breaks locations. While, whenever ζ is very small, many false detections can appear.

Figure 6.12 shows the local power for $\beta_{1,1} = 2$, $\beta_{2,1} = -1$ and $(\tau_1, \tau_2) = (110, 220)$. When the local power crosses the threshold considered, which is the nominal level plus ζ (0.051 in this figure), we consider that we are around the zone of a break and we calculate the local power at the next few observations. The observation which has the maximum local power will be considered a change-point. Next, we skip some observations (as assumed in Chapter 5), and after that, we repeat the same technique (sequential method).

The results shown in Table 6.5 demonstrate that the algorithm presented in Chapter 5 is efficient for detecting more than one break in the parameters of CHARN (1.1) models. Compared to the other methods in literature, it is more accurate for detecting the weak changes and estimating their locations. Our method is sensitive to the choice of ζ .

For a class of models (6.4), for $n = 300$, $\alpha = 5\%$, $(\tau_1, \tau_2) = (110, 220)$, $\rho_{0,1} = 0.3$, $\beta_{1,1} = 2$, $\beta_{2,1} = -1$ and for $\zeta = 0.06\%$, 0.09% , 0.1% , 0.13% and 0.15% , Table 6.6 shows the results of breaks locations estimation corresponding to different values of ζ .

$(\tau_1, \tau_2) \backslash (\beta_{1,1}, \beta_{2,1})$	$(2, -1)$	$(3, -1)$	$(-1, 3)$	$(2, 3)$
$(110, 220)$	$(113, 221)$	$(115, 221)$	$(117, 222)$	$(112, 222)$
ζ	0.1%	0.1%	0.1%	0.1%
$(130, 230)$	$(131, 234)$	$(131, 231)$	$(132, 231)$	$(131, 234)$
ζ	0.1%	0.1%	0.1%	0.1%

Table 6.5: Estimation of breaks locations ($k = 2$).

$(\tau_1, \tau_2) \backslash (\beta_{1,1}, \beta_{2,1})$	$(2, -1)$	$(2, -1)$	$(2, -1)$	$(2, -1)$	$(2, -1)$
$(110, 220)$	$(65, 110, 165, 220)$	$(111, 221)$	$(113, 221)$	$(117, 223)$	$(118, 225)$
ζ	0.06%	0.09%	0.1%	0.13%	0.15%

Table 6.6: Breaks locations for the same magnitude of change and for different thresholds.

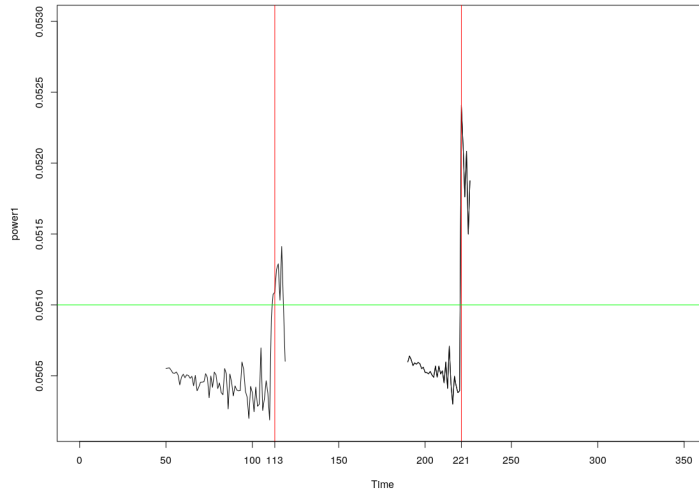


Figure 6.12: Breaks locations estimation in the case of two breaks.

Case of three breaks ($k = 3$).

The data to be studied are obtained from the following model

$$\left\{ \begin{array}{ll} X_t = \rho_{0,1} + \varepsilon_t, & t = 1, \dots, \tau_1, \\ X_t = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \varepsilon_t, & t = \tau_1 + 1, \dots, \tau_2, \\ X_t = \rho_{0,1} + \frac{\beta_{2,1}}{\sqrt{n}} + \varepsilon_t, & t = \tau_2 + 1, \dots, \tau_3, \\ X_t = \rho_{0,1} + \frac{\beta_{3,1}}{\sqrt{n}} + \varepsilon_t, & t = \tau_3 + 1, \dots, n, \end{array} \right. \quad (6.7)$$

where (ε_t) is a sequence of standard Gaussian white noise. As before, we aim to detect and locate the changes from which their number and their locations are assumed to be unknown.

The results for 5000 replications, $n = 400$, $\rho_{0,1} = 0.2$, $(\tau_1, \tau_2, \tau_3) = (90, 195, 310)$, and for different values of $\beta_{1,1}, \beta_{2,1}, \beta_{3,1}$ and for the same threshold corresponding to $\zeta = 0.1\%$, are shown in Table 6.7.

(τ_1, τ_2, τ_3) \diagdown $(\beta_{1,1}, \beta_{2,1}, \beta_{3,1})$	$(2, -1, 0.5)$	$(3, 1, -0.5)$	$(-1, 2, 4)$	$(1, 2, 3)$
$(90, 195, 310)$	$(91, 196, 316)$	$(91, 197, 317)$	$(93, 196, 315)$	$(93, 198, 316)$
ζ	0.1%	0.1%	0.1%	0.1%

Table 6.7: Breaks locations estimation in the case of three breaks.

(τ_1, τ_2, τ_3)	$(\beta_{1,1}, \beta_{2,1}, \beta_{3,1})$	(2, -1, 1)	(2, -1, 1)	(2, -1, 1)	(2, -1, 1)
(90,195,310)		(50,90,140, 195,245,310)	(91,196,314)	(91,196,315)	(94,198,319)
ζ		0.06%	0.1%	0.13%	0.2%

Table 6.8: Breaks locations estimation in the case of three breaks for the same magnitude of change and for different values of ζ .

The behavior of the local power corresponding to the case of $\beta_{1,1} = 3, \beta_{2,1} = -1, \beta_{3,1} = -0.5$ and $\zeta = 0.1\%$ is presented in Figure 6.13.

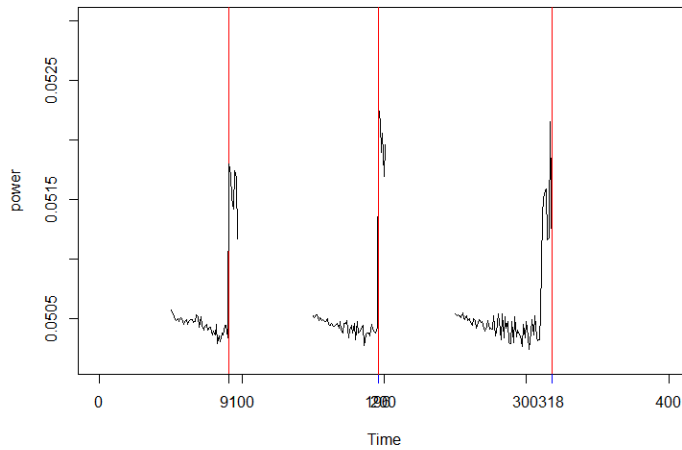


Figure 6.13: Behavior of the local power of the test in case of three breaks in a class of shifted models.

Now, we generalize the situation of detecting multiple change-points to be in a class of models (6.5). The data under study are obtained from the following model

$$\left\{ \begin{array}{l} X_{t-1} = \rho_{0,1} + \rho_{0,2}X_{t-1} + \varepsilon_t, \quad t = 1, \dots, \tau_1 \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, \quad t = \tau_1 + 1, \dots, \tau_2, \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{2,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{2,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, \quad t = \tau_2 + 1, \dots, \tau_3, \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{3,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{3,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, \quad t = \tau_3 + 1, \dots, n, \end{array} \right.$$

where (ε_t) is a sequence of standard Gaussian white noise. The number of change-points are assumed to be unknown and we aim to detect these weak breaks and estimate their locations.

Note that, in this case, the code in "R" program took much more time than the above situations which is related to the replications of the fitted AR(1) model on a large number of instants checked. For 5000 replications, $n = 300$, $\rho_0 = (\rho_{0,1}, \rho_{0,2}) = (0.2, 0.3)$ $\tau = (\tau_1, \tau_2, \tau_3) = (90, 190, 275)$ and for different values of the components of $\beta_j = (\beta_{j,1}, \beta_{j,2})$, $j = 1, 2, 3$ and for the same threshold corresponding to $\zeta = 0.1\%$, we obtain Table 6.11.

$(\beta_1, \beta_2, \beta_3)$	(τ_1, τ_2, τ_3)	$(90, 190, 275)$ $\zeta = 0.1\%$
$\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$		(93,195,280)
$\left(\begin{bmatrix} 1 \\ -0.5 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right)$		(96,196,278)
$\left(\begin{bmatrix} -2 \\ 1.5 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -0.5 \\ -1 \end{bmatrix} \right)$		(94,193,277)

Table 6.9: Estimation of the breaks locations in a class of AR(1) model.

The results presented in Table 6.11 show that our method is efficient to detect the weak breaks in the parameters of AR(1) models. Such breaks contain those in the autocorrelation which is a very important result.

Again for AR(1) models, the data to be studied are obtained for a fixed values of β_j , $j = 1, 2, 3$, and the same instants of break $\tau = (\tau_1, \tau_2, \tau_3)$. Table 6.10 shows the results of the corresponding estimates of change-points locations for different values of ζ . When the threshold is too small, the local power crosses it early and then, a change-point is detected which may be false alarm. This is what happens for $\zeta = 0.7\%$. However, when the threshold is high enough, the values of the local power doesn't exceed the threshold considered, especially when the values of the magnitudes of the change β_j/\sqrt{n} is too

small. This is the case for $\zeta = 2.5\%$.

$(\beta_1, \beta_2, \beta_3)$ \diagdown (τ_1, τ_2, τ_3)	(90, 190, 275)
$\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$	(56, 93, 142, 195, 245, 280) $\zeta = 0.7\%$
$\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$	(93, 195, 280) $\zeta = 0.1\%$
$\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$	(94, 197, 282) $\zeta = 0.15\%$
$\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$	(96, 283) $\zeta = 0.25\%$

Table 6.10: Estimation of the breaks locations in a class of AR(1) model for different threshold.

AR-ARCH models.

Now, we treat the case of multiple change-points corresponding to the data obtained by taken $\theta_1 = 1$, $\theta_2 = 0.02$ and $\rho_{0,3} = 0$. Then, the data under study are generated by the following AR-ARCH models

$$\left\{ \begin{array}{l} X_{t-1} = \rho_{0,1} + \rho_{0,2}X_{t-1} + \sqrt{1 + 0.02X_{t-1}^2}\varepsilon_t, \quad t = 1, \dots, \tau_1 \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \sqrt{1 + 0.02X_{t-1}^2}\varepsilon_t, \quad t = \tau_1 + 1, \dots, \tau_2, \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{2,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{2,2}}{\sqrt{n}} \right) X_{t-1} + \sqrt{1 + 0.02X_{t-1}^2}\varepsilon_t, \quad t = \tau_2 + 1, \dots, \tau_3, \\ X_{t-1} = \rho_{0,1} + \frac{\beta_{3,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{3,2}}{\sqrt{n}} \right) X_{t-1} + \sqrt{1 + 0.02X_{t-1}^2}\varepsilon_t, \quad t = \tau_3 + 1, \dots, n, \end{array} \right.$$

For $n = 350$, $\rho_{0,1} = 0.2$, $\rho_{0,2} = 0.3$, $\beta_1 = (3, -2)$, $\beta_2 = (1, 1)$, $\beta_3 = (-2, 4)$, the detection of the changes are illustrated with Figure.

$(\beta_1, \beta_2, \beta_3)$ \ / \ (τ_1, τ_2, τ_3)	$(90, 190, 275)$ $\zeta = 0.1\%$
$\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$	(93,195,280)
$\left(\begin{bmatrix} 1 \\ -0.5 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right)$	(96,196,278)
$\left(\begin{bmatrix} -2 \\ 1.5 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -0.5 \\ -1 \end{bmatrix} \right)$	(94,193,277)

Table 6.11: Estimation of the breaks locations in a class of AR(1) model.

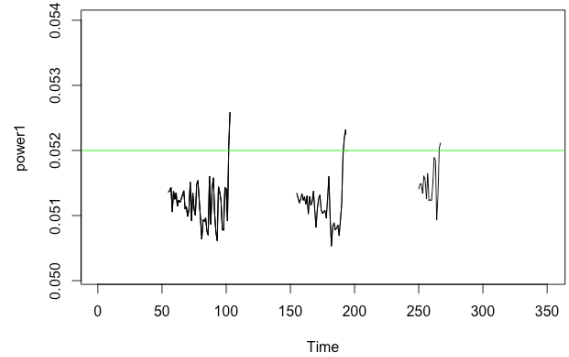


Figure 6.14: Detection of multiple change-points in a class of ARCH models

Collecting the results obtained above, in a class of AR-ARCH models, we can see that our method is efficient and accurate for detecting multiple breaks in a class of non-linear models.

6.3 Comparison with Horváth et al. (2020)

In a class of shifted model (6.4), Ltaifa (2021) did a comparison between her method, which is a particular case from ours, and other methods including the one of Horváth et al. (2020). She concluded that her method is more powerful for estimating the localization of weak breaks.

In this section, we compare our method to that of Horváth et al. (2020) denoted by SCUSUM, for a class of more general model. Recall the model (6.5),

$$X_{t-1} = \rho_{0,1} + \frac{\beta_{1,1}}{\sqrt{n}} + \left(\rho_{0,2} + \frac{\beta_{1,2}}{\sqrt{n}} \right) X_{t-1} + \varepsilon_t, \quad t = 1, \dots, n.$$

For $n = 200$, in the case of one single break at $\tau_1 = 80, 120, 185$, we do 1000 replications and at each replication, we obtain a data generated by the same model (6.5) with same parameter values and we investigate the estimation of change-point by SCUSUM. For $\alpha = 5\%$, $\rho_{0,1} = 0.5$, $\rho_{0,2} = -0.2$ and different values of $\beta = (\beta_{1,1}, \beta_{1,2})^\top$ we obtain Table 6.12.

During these 1000 replications, for most of them, SCUSUM has not been able to detect any change-points. For that, we keep only the cases where this method detects a change, and we calculate the mean of the change location estimated among these 1000 replications, and we display it in Table 6.12.

It is obvious that our method is more accurate than SCUSUM for the detection of weak changes in the parameters of model (6.5).

$\tau_1 = 80$	$\begin{pmatrix} \beta_{1,1} \\ \beta_{1,2} \end{pmatrix}$	$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 3 \end{pmatrix}$	$\begin{pmatrix} -3 \\ 2 \end{pmatrix}$
Our method		78	81	82	80	80
SCUSUM		95	99	86	122	99
$\tau_1 = 120$						
Our method		122	121	119	120	121
SCUSUM		110	114	105	125	122
$\tau_1 = 185$						
Our method		179	179	179	180	180
SCUSUM		105	103	122	114	101

Table 6.12: Breaks locations estimation obtained by our method and SCUSUM for different instants of break τ_1 and different magnitudes of change.

6.4 Conclusion

In this chapter, we have evaluated the performance of a new method for change-points detection and for estimating the localization. The results obtained show that our method is efficient on several models considered.

Compared to Horváth et al. (2020), it seems to be more efficient for detecting weak changes and estimating their locations. It seems also more performant for detecting changes in the earliest data and in the latest ones. This fact was already pointed out by Ltaifa (2021) and Ngatchou-Wandji and Ltaifa (2021).

Chapter 7

General conclusion and outlook

7.1 Conclusion

Ngatchou-Wandji and Ltaifa (2021) studied weak break in the mean of a class of CHARN models. We have generalized their work to a class of more general CHARN models.

In our work, we studied weak breaks in the parameters of the function T . We started our work by assuming that the function T , V and the parameters ρ_0 and γ_0 are known. We established a LAN and a contiguity results. Also, we derived an explicit expression of the local power of the likelihood test that constructed.

Next, we assumed that the function T and V have known forms, but the parameters ρ_0 and γ_0 are unknown. First, we estimated these parameters, then we proved the convergence between the central sequence based on the estimated parameters and the one based on the true parameters. In this case, we derived an explicit expression of the local power, and we proved that our test remains optimal if we replace the parameters with their estimators. From these results, we used the theoretical power for detecting weak breaks in time series through an algorithm that we constructed.

The simulation experiment that we constructed shows that our method can detect the weak breaks in the mean of shifted white noise models. It can also detect weak breaks in the autocorrelation, and in all the parameters of some AR models considered.

We can see that, in all the examples illustrated in the simulation part, the detection of the break instants was exact or very close (differ for one or two instants) to the exact instants of breaks. This is a good indicator of the efficiency of our method.

Compared to other method, like Horváth et al. (2020), it seems to be more efficient for detecting weak changes and estimating their locations. Sometimes, Horváth et al. (2020) method detects breaks in data having no break. This has not happened with our method

when we choose a suitable ζ .

7.2 Outlook

With regard to the simulation results in the preceding chapter, our work seems interesting and needs some extensions in many directions. In the sequel, we outline some perspectives.

- i. At a theoretical point of view, our work is done with the assumption that the density function f of ε_t is known. It can be interesting to study the case where f is unknown.
- ii. In our study the function V is assumed to be given. Our next purpose is to study weak breaks in its parameters. These breaks induce changes in the conditional variance.
- iii. We also plan to try the case where both the functions T and V are jointly parametric.
- iv. At a practical point of view, the creation of an R package that uses our method to detect the weak breaks in the data is something to be investigated.
- v. It will be interesting to update the "code" and make it able to check if a future observation is a break, in other words, we think to adapt the "R" program to the on-line detection.

Bibliography

- Adak, S. (1998). Time-dependent spectral analysis of nonstationary time series. *Journal of the American Statistical Association*, 93(444):1488–1501.
- An, H., Chen, M., and Huang, F. (1997). The geometric ergodicity and existence of moments for a class of non-linear time series model. *Statistics & probability letters*, 31(3):213–224.
- Anastasiou, A. and Fryzlewicz, P. (2022). Detecting multiple generalized change-points by isolating single ones. *Metrika*, 85(2):141–174.
- Anderson, T. W. and Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The annals of mathematical statistics*, pages 193–212.
- Andrews, D. W. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica: Journal of the Econometric Society*, pages 821–856.
- Andrews, D. W. and Ploberger, W. (1994). Optimal tests when a nuisance parameter is present only under the alternative. *Econometrica: Journal of the Econometric Society*, pages 1383–1414.
- Badagián, A. L., Kaiser, R., and Peña, D. (2015). Time series segmentation procedures to detect, locate and estimate change-points. In *Empirical Economic and Financial Research*, pages 45–59. Springer.
- Bahadur, R. R., Rao, R. R., et al. (1960). On deviations of the sample mean. *Ann. Math. Statist.*, 31(4):1015–1027.
- Baranowski, R., Chen, Y., and Fryzlewicz, P. (2019). not: Narrowest-over-threshold change-point detection. *R package version*, 1.

- Bardet, J.-M., Kengne, W. C., and Wintenberger, O. (2010). Detecting multiple change-points in general causal time series using penalized quasi-likelihood. *arXiv preprint arXiv:1008.0054*.
- Bardet, J.-M. and Wintenberger, O. (2009). Asymptotic normality of the quasi-maximum likelihood estimator for multidimensional causal processes. *The Annals of Statistics*, 37(5B):2730–2759.
- Basawa, I. and Hwang, S. Y. (2003). Estimation for nonlinear autoregressive models generated by beta-arch processes. *Sankhyā: The Indian Journal of Statistics*, pages 744–762.
- Bernoulli, J. (1713). *Ars conjectandi*. Impensis Thurnisiorum, fratrum.
- Bhattacharya, P. and Zhou, H. (2017). Nonparametric stopping rules for detecting small changes in location and scale families. In *From Statistics to Mathematical Finance*, pages 251–271. Springer.
- Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660.
- Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *The review of economics and statistics*, pages 542–547.
- Bollerslev, T., Engle, R. F., and Nelson, D. B. (1994). Arch models. *Handbook of econometrics*, 4:2959–3038.
- Brockwell, P., Davis, R., and Salehi, H. (1990). A state-space approach to transfer-function modeling. *Statistical Inference in Stochastic Processes*, 6:233.
- Brockwell, P. J. and Davis, R. A. (2002). *Introduction to time series and forecasting*. Springer.
- Brockwell, P. J., Davis, R. A., and Fienberg, S. E. (1991). *Time series: theory and methods: theory and methods*. Springer Science & Business Media.
- Brown, R. L., Durbin, J., and Evans, J. M. (1975). Techniques for testing the constancy of regression relationships over time. *Journal of the Royal Statistical Society: Series B (Methodological)*, 37(2):149–163.
- Cassart, D., Hallin, M., and Paindaveine, D. (2008). Optimal detection of fechner-asymmetry. *Journal of Statistical Planning and Inference*, 138(8):2499–2525.

- Chatterjee, S. and Das, S. (2003). Parameter estimation in conditional heteroscedastic models. *Communications in Statistics-Theory and Methods*, 32(6):1135–1153.
- Chen, G.-y., Gan, M., and Chen, G.-l. (2018). Generalized exponential autoregressive models for nonlinear time series: stationarity, estimation and applications. *Information Sciences*, 438:46–57.
- Chen, K.-m., Cohen, A., and Sackrowitz, H. (2011). Consistent multiple testing for change points. *Journal of multivariate analysis*, 102(10):1339–1343.
- Chen, M. and An, H. Z. (1998). A note on the stationarity and the existence of moments of the garch model. *Statistica Sinica*, pages 505–510.
- Chernoff, H. and Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to changes in time. *The Annals of Mathematical Statistics*, 35(3):999–1018.
- Cho, H. and Fryzlewicz, P. (2020). Multiple change point detection under serial dependence: Wild energy maximisation and gappy schwarz criterion. *arXiv preprint arXiv:2011.13884*.
- Ciuperca, G. (2011). A general criterion to determine the number of change-points. *Statistics & probability letters*, 81(8):1267–1275.
- Cohen, A., Sackrowitz, H. B., Xu, M., et al. (2009). A new multiple testing method in the dependent case. *The Annals of Statistics*, 37(3):1518–1544.
- Csörgő, M., Horváth, L., and Szyszkowicz, B. (1997). Integral tests for suprema of kiefer processes with application. *Statistics & Risk Modeling*, 15(4):365–378.
- Dahlhaus, R. (1997). Fitting time series models to nonstationary processes. *The annals of Statistics*, 25(1):1–37.
- Davis, R. A., Huang, D., and Yao, Y.-C. (1995). Testing for a change in the parameter values and order of an autoregressive model. *The Annals of Statistics*, pages 282–304.
- Davis, R. A., Lee, T. C., and Rodriguez-Yam, G. A. (2008). Break detection for a class of nonlinear time series models. *Journal of Time Series Analysis*, 29(5):834–867.
- Davis, R. A., Lee, T. C. M., and Rodriguez-Yam, G. A. (2006). Structural break estimation for nonstationary time series models. *Journal of the American Statistical Association*, 101(473):223–239.

- Dehling, H., Rooch, A., and Taqqu, M. S. (2013). Non-parametric change-point tests for long-range dependent data. *Scandinavian Journal of Statistics*, 40(1):153–173.
- Dehling, H., Rooch, A., and Taqqu, M. S. (2017). Power of change-point tests for long-range dependent data. *Electronic Journal of Statistics*, 11(1):2168–2198.
- der Vaart, V. and Statistics, A. A. (1998). Cambridge university press: New york. NY, USA.
- Donoho, D. L., Mallat, S. G., von Sachs, R., et al. (1998). Estimating covariances of locally stationary processes: rates of convergence of best basis methods. Technical report, Citeseer.
- Döring, M. (2011). Convergence in distribution of multiple change point estimators. *Journal of statistical planning and inference*, 141(7):2238–2248.
- Droesbeke and Fine, J.-J. (1996). *Inférence non paramétrique: Les statistiques de rangs*. Ed. de l'Université de Bruxelles; Ed. Ellipses.
- Dupuis, D., Sun, Y., and Wang, H. J. (2015). Detecting change-points in extremes.
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the econometric society*, pages 987–1007.
- Fotopoulos, S. B., Jandhyala, V. K., and Tan, L. (2009). Asymptotic study of the change-point mle in multivariate gaussian families under contiguous alternatives. *Journal of statistical planning and inference*, 139(3):1190–1202.
- Francq, C. and Zakoian, J.-M. (2004). Maximum likelihood estimation of pure garch and arma-garch processes. *Bernoulli*, 10(4):605–637.
- Francq, C. and Zakoian, J.-M. (2007). Quasi-maximum likelihood estimation in garch processes when some coefficients are equal to zero. *Stochastic Processes and their Applications*, 117(9):1265–1284.
- Frick, K., Munk, A., and Sieling, H. (2014). Multiscale change point inference. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 76(3):495–580.
- Fryzlewicz, P. (2014). Wild binary segmentation for multiple change-point detection. *The Annals of Statistics*, 42(6):2243–2281.

- Ghosh, H., Gurung, B., and Gupta, P. (2015). Fitting expar models through the extended kalman filter. *Sankhya B*, 77(1):27–44.
- Giraitis, L. and Robinson, P. M. (2001). Whittle estimation of arch models. *Econometric Theory*, 17(3):608–631.
- Gombay, E. and Serban, D. (2009). Monitoring parameter change in ar (p) time series models. *Journal of Multivariate Analysis*, 100(4):715–725.
- Granger, C. and Andersen, A. (1978). An introduction to bilinear time series models. vandenhoek & ruprecht, göttingen. *Recent Advances in ARCH Modelling*, 35:36.
- Guégan, D. et al. (1994). Séries chronologiques non linéaires à temps discret. Technical report.
- Hájek, J. (1961). Some extensions of the wald-wolfowitz-noether theorem. *The Annals of Mathematical Statistics*, pages 506–523.
- Hall, P. and Heyde, C. C. (2014). *Martingale limit theory and its application*. Academic press.
- Hall, W. and Mathiason, D. J. (1990). On large-sample estimation and testing in parametric models. *International Statistical Review/Revue Internationale de Statistique*, pages 77–97.
- Härdle, W., Tsybakov, A., and Yang, L. (1998). Nonparametric vector autoregression. *Journal of Statistical Planning and Inference*, 68(2):221–245.
- Horváth, L. and Kokoszka, P. (1997). The effect of long-range dependence on change-point estimators. *Journal of Statistical Planning and Inference*, 64(1):57–81.
- Horváth, L., Miller, C., and Rice, G. (2020). A new class of change point test statistics of rényi type. *Journal of Business & Economic Statistics*, 38(3):570–579.
- Huh, J. (2010). Detection of a change point based on local-likelihood. *Journal of multivariate analysis*, 101(7):1681–1700.
- Hwang, S. Y. and Basawa, I. (2001). Nonlinear time series contiguous to ar (1) processes and a related efficient test for linearity. *Statistics & probability letters*, 52(4):381–390.
- Jarušková, D. (2010). Asymptotic behaviour of a test statistic for detection of change in mean of vectors. *Journal of statistical planning and inference*, 140(3):616–625.

- Johnson, M. L. and Faunt, L. M. (1992). [1] parameter estimation by least-squares methods. In *Methods in enzymology*, volume 210, pages 1–37. Elsevier.
- Kim, M., Lee, S., et al. (2009). Test for tail index change in stationary time series with pareto-type marginal distribution. *Bernoulli*, 15(2):325–356.
- Koutmos, G. (1997). Do emerging and developed stock markets behave alike? evidence from six pacific basin stock markets. *Journal of International Financial Markets, Institutions and Money*, 7(3):221–234.
- Krämer, W., Ploberger, W., and Alt, R. (1988). Testing for structural change in dynamic models. *Econometrica: Journal of the Econometric Society*, pages 1355–1369.
- Krämer, W. and Schotman, P. (1992). Range vs. maximum in the ols-based version of the cusum test. *Economics letters*, 40(4):379–381.
- Kreiss, J.-P. (1987). On adaptive estimation in stationary arma processes. *The Annals of Statistics*, pages 112–133.
- Lavielle, M. and Lebarbier, E. (2001). An application of mcmc methods for the multiple change-points problem. *Signal processing*, 81(1):39–53.
- Le Cam, L. (1960). Locally asymptotically normal families of distributions. *Univ. California Publ. Statist.*, 3:37–98.
- Le Cam, L. (1986). The central limit theorem around 1935. *Statistical science*, pages 78–91.
- Lebarbier, É. (2005). Detecting multiple change-points in the mean of gaussian process by model selection. *Signal processing*, 85(4):717–736.
- LeBaron, B. (1992). Some relations between volatility and serial correlations in stock market returns. *Journal of Business*, pages 199–219.
- Lehmann, E. L., Romano, J. P., and Casella, G. (2005). *Testing statistical hypotheses*, volume 3. Springer.
- Li, H., Munk, A., and Sieling, H. (2016). Fdr-control in multiscale change-point segmentation. *Electronic Journal of Statistics*, 10(1):918–959.
- Linton, O. (1993). Adaptive estimation in arch models. *Econometric Theory*, pages 539–569.

- Lounis, T. (2011). Asymptotically optimal tests when parameters are estimated. *arXiv preprint arXiv:1110.0205*.
- Lounis, T. (2012). Local asymptotically optimal test in arch model. *arXiv preprint arXiv:1206.1393*.
- Ltaifa, M. (2021). *Tests optimaux pour détecter les signaux faibles dans les séries chronologiques*. Theses, Université de Lorraine ; Université de Sousse (Tunisie).
- MacNeill, I. B. (1974). Tests for change of parameter at unknown times and distributions of some related functionals on brownian motion. *The Annals of Statistics*, pages 950–962.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441.
- Million, E. (2007). The hadamard product. *Course Notes*, 3(6).
- Monfort, A. (1982). *Cours de statistique mathématique*. Economica.
- Ngatchou-Wandji, J. (2002). Weak convergence of some marked empirical processes: Application to testing heteroscedasticity. *Journal of Nonparametric Statistics*, 14(3):325–339.
- Ngatchou-Wandji, J. (2005). Checking nonlinear heteroscedastic time series models. *Journal of statistical planning and inference*, 133(1):33–68.
- Ngatchou-Wandji, J. (2008). Estimation in a class of nonlinear heteroscedastic time series models. *Electronic Journal of Statistics*, 2:40–62.
- Ngatchou-Wandji, J. and Harel, M. (2013). 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.
- Ngatchou-Wandji, J. and Ltaifa, M. (2021). On detecting weak changes in the mean of charn models. *arXiv preprint arXiv:2101.08597*.
- Ngatchou-Wandji, J., Puri, M. L., Harel, M., and Elharfaoui, E. (2019). Testing nonstationary and absolutely regular nonlinear time series models. *Statistical Inference for Stochastic Processes*, 22(3):557–593.
- Olshen, A. B., Venkatraman, E. S., Lucito, R., and Wigler, M. (2004). Circular binary segmentation for the analysis of array-based dna copy number data. *Biostatistics*, 5(4):557–572.

- Ombao, H., Raz, J., Von Sachs, R., and Guo, W. (2002). The slex model of a non-stationary random process. *Annals of the Institute of Statistical Mathematics*, 54(1):171–200.
- Ozaki, T. (1980). Non-linear time series models for non-linear random vibrations. *Journal of Applied Probability*, 17(1):84–93.
- Ozaki, T. (1982). The statistical analysis of perturbed limit cycle processes using nonlinear time series models. *Journal of Time Series Analysis*, 3(1):29–41.
- Ozaki, T. (1985). 2 non-linear time series models and dynamical systems. *Handbook of statistics*, 5:25–83.
- Ozaki, T. and Oda, H. (1977). Non-linear time series model identification by akaike's information criterion. *IFAC Proceedings Volumes*, 10(12):83–91.
- Page, E. S. (1954). Continuous inspection schemes. *Biometrika*, 41(1/2):100–115.
- Parzen, E. (1962). On estimation of a probability density function and mode. *The annals of mathematical statistics*, 33(3):1065–1076.
- Pit (1994). High Derivations for Multidimensional Stationary Gaussian Process with Independent Components . *Stability Problems for Stochastic Models*, pages 197–210.
- Prášková, Z. and Chochola, O. (2014). M-procedures for detection of a change under weak dependence. *Journal of Statistical Planning and Inference*, 149:60–76.
- Rényi, A. (1953). On the theory of order statistics. *Acta Math. Acad. Sci. Hung*, 4(2).
- Rissanen, J. (1998). *Stochastic complexity in statistical inquiry*, volume 15. World scientific.
- Ruggieri, E. and Antonellis, M. (2016). An exact approach to bayesian sequential change point detection. *Computational Statistics & Data Analysis*, 97:71–86.
- Schwarz, G. (1978). Estimating the dimension of a model. *The annals of statistics*, pages 461–464.
- Schwarz, G. et al. (1978). Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464.
- Shao, X. and Zhang, X. (2010). Testing for change points in time series. *Journal of the American Statistical Association*, 105(491):1228–1240.

- Strejc, V. (1980). Least squares parameter estimation. *Automatica*, 16(5):535–550.
- Taniguchi, M. and Kakizawa, Y. (2012). *Asymptotic theory of statistical inference for time series*. Springer Science & Business Media.
- Tjøstheim, D. (1990). Non-linear time series and markov chains. *Advances in applied probability*, 22(3):587–611.
- Tong, H. (1990). *Non-linear time series: a dynamical system approach*. Oxford university press.
- Vogelsang, T. J. (1997). Wald-type tests for detecting breaks in the trend function of a dynamic time series. *Econometric Theory*, 13(6):818–848.
- Vostrikova, L. Y. (1981). Detecting “disorder” in multidimensional random processes. In *Doklady akademii nauk*, volume 259, pages 270–274. Russian Academy of Sciences.
- Whittle, P. (1953). Estimation and information in stationary time series. *Arkiv för matematik*, 2(5):423–434.
- Wong, H., Ip, W., and Li, Y. (2001). Detection of jumps by wavelets in a heteroscedastic autoregressive model. *Statistics & probability letters*, 52(4):365–372.
- Yang, Y. and Song, Q. (2014). Jump detection in time series nonparametric regression models: a polynomial spline approach. *Annals of the Institute of Statistical Mathematics*, 66(2):325–344.
- Yau, C. Y. and Zhao, Z. (2016). Inference for multiple change points in time series via likelihood ratio scan statistics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(4):895–916.
- Zaffaroni, P. and d’Italia, B. (2003). Gaussian inference on certain long-range dependent volatility models. *Journal of econometrics*, 115(2):199–258.
- Zeileis, A. (2001). p values and alternative boundaries for cusum tests. Technical report, Technical Report.
- Zeileis, A. (2004). Alternative boundaries for cusum tests. *Statistical Papers*, 45(1):123–131.