



**HAL**  
open science

# Testing a Class of Piece-Wise CHARN Models with Application to Change-Point Study

Youssef Salman, Joseph Ngatchou-Wandji, Zaher Khraibani

► **To cite this version:**

Youssef Salman, Joseph Ngatchou-Wandji, Zaher Khraibani. Testing a Class of Piece-Wise CHARN Models with Application to Change-Point Study. Mathematics , 2024, 12 (13), pp.2092. 10.3390/math12132092 . hal-04810496

**HAL Id: hal-04810496**

**<https://ehesp.hal.science/hal-04810496v1>**

Submitted on 10 Dec 2024

**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.



Distributed under a Creative Commons Attribution 4.0 International License

Article

# Testing a Class of Piece-Wise CHARN Models with Application to Change-Point Study

Youssef Salman <sup>1,2,†</sup> , Joseph Ngatchou-Wandji <sup>3,\*,†</sup> and Zaher Khraibani <sup>2</sup> 

<sup>1</sup> Mines Saint-Etienne, CNRS, UMR 6158 LIMOS, Institut Henri Fayol, University Clermont Auvergne, 42023 Saint-Etienne, France; youssef.salman@emse.fr

<sup>2</sup> Department of Applied Mathematics, Faculty of Sciences, Lebanese University, Beirut 2038 1003, Lebanon; zaher.khraibani@gmail.com

<sup>3</sup> EHESP Rennes and Institut Élie Cartan de Lorraine, CEDEX, 54506 Vandoeuvre-lès-Nancy, France

\* Correspondence: joseph.ngatchou-wandji@univ-lorraine.fr

† These authors contributed equally to this work.

**Abstract:** We study a likelihood ratio test for testing the conditional mean of a class of piece-wise stationary 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. We describe an algorithm for detecting change points and estimating their locations. The estimates are obtained as time indices, maximizing the estimate of the local power. The simulation study we conduct shows the good performance of our method on the examples considered. This method is also applied to a set of financial data.

**Keywords:** CHARN models; change points; LAN; likelihood ratio tests

**MSC:** 62M10; 62M02; 62M05; 62F03; 62F05



**Citation:** Salman, Y.; Ngatchou-Wandji, J.; Khraibani, Z. Testing a Class of Piece-Wise CHARN Models with Application to Change-Point Study. *Mathematics* **2024**, *12*, 2092. <https://doi.org/10.3390/math12132092>

Academic Editor: Manuel Alberto M. Ferreira

Received: 3 June 2024

Revised: 29 June 2024

Accepted: 1 July 2024

Published: 3 July 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Let  $d, p, k, n \in \mathbb{N}$  and  $k \ll n$ . Assume the observations  $X_1, \dots, X_n$  issued from the following piece-wise stationary CHARN model (see, e.g., [1])

$$X_t = T(\rho_0 + \gamma \odot \omega(t); \mathbf{X}_{t-1}) + V(\mathbf{X}_{t-1})\varepsilon_t, t \in \mathbb{Z}, \quad (1)$$

with

$$X_t = Y_{t,j} = T(\rho_0 + \gamma_j \omega_j(t); \mathbf{X}_{t-1,j}) + V(\mathbf{X}_{t-1,j})\varepsilon_t, \quad \tau_{j-1} \leq t < \tau_j, \\ j = 1, \dots, k + 1, \quad (2)$$

where for  $j = 1, \dots, k$ ,  $(Y_{t,j})_{t \in \mathbb{Z}}$  is a stationary and ergodic process;  $\rho_0 \in \mathbb{R}^p$ ,  $T(\rho_0, \cdot)$  and  $V(\cdot)$  are real-valued functions with  $\inf_{x \in \mathbb{R}^d} V(x) > 0$ ; the  $\tau_j$ ,  $j = 0, \dots, k + 1$ , are potential instants of changes with  $\tau_0 = 1$  and  $\tau_{k+1} = n + 1$ ; for  $j = 1, \dots, k$ ,  $\mathbf{X}_{t,j} = (Y_{t,j}, \dots, Y_{t-d+1,j})^\top$ ,  $\mathbf{X}_{\tau_{j-1}+\ell} = \mathbf{X}_{\tau_{j-1}+\ell,j}$ ,  $\ell = 0, \dots, d - 1$  and for  $t \in [\tau_{j-1} + d - 1, \tau_j)$ ,  $\mathbf{X}_t = (X_t, \dots, X_{t-d+1})^\top$ ; for  $j, \ell = 1, \dots, k$ ,  $j \neq \ell$ , the processes  $(Y_{t,j})_{t \in \mathbb{Z}}$  and  $(Y_{t,\ell})_{t \in \mathbb{Z}}$  are mutually independent;  $(\varepsilon_t)_{t \in \mathbb{Z}}$  is standard white noise with density  $f$ .  $\gamma = (\gamma_1^\top, \dots, \gamma_{k+1}^\top)^\top$ ,  $\gamma_j \in \mathbb{R}^p$ ,  $j = 1, \dots, k + 1$ ;  $\omega(t) = (\mathbf{1}_{[\tau_0, \tau_1)}(t), \mathbf{1}_{[\tau_1, \tau_2)}(t), \dots, \mathbf{1}_{[\tau_{k-1}, \tau_k)}(t), \mathbf{1}_{[\tau_k, \tau_{k+1})}(t))^\top = (\omega_1(t), \dots, \omega_{k+1}(t)) \in \{0, 1\}^{k+1}$ ; for  $\gamma = (\gamma_1^\top, \dots, \gamma_{k+1}^\top)^\top$  and  $\omega(t) = (\omega_1(t), \dots, \omega_{k+1}(t))^\top$ ,  $\gamma \odot \omega(t)$  stands for  $\gamma \odot \omega(t) = \gamma_1 \omega_1(t) + \dots + \gamma_{k+1} \omega_{k+1}(t) \in \mathbb{R}^p$ , and  $\gamma_i \omega_i = (\gamma_{i,1} \omega_i, \dots, \gamma_{i,p} \omega_i) \in \mathbb{R}^p$ . The class of models (2) is very large. It contains models such as AR( $p$ ), ARCH( $p$ ), EXPAR( $p$ ), and GEXPAR( $p$ ) whose statistical and probability properties are widely studied in the literature (see, e.g., [2] for a study of the ergodicity of GEXPAR models).

As noted in [3], the assumption that  $(\mathbf{X}_{t,j})_{t \in \mathbb{Z}}$  and  $(\mathbf{X}_{t,\ell})_{t \in \mathbb{Z}}$  are independent can be extended to some weak dependence assumption. In this paper, for  $\gamma_0 \in \mathbb{R}^{p(k+1)}$  and  $\beta \in \mathbb{R}^{p(k+1)}$  depending on the  $\tau_j$ s, we construct a likelihood ratio test for testing

$$H_0 : \gamma = \gamma_0 \quad \text{against} \quad H_\beta^{(n)} : \gamma = \gamma_n = \gamma_0 + \frac{\beta}{\sqrt{n}}. \quad (3)$$

A particular case of this work is studied in [4]. The literature on change points is extensive and varied. Some basic notions and theory are presented in [5], where one can find number of references on the first works on the subject. Most of the recent papers on change points are in time series or regression contexts. Various methods and techniques are used for the study. Ref. [6] proposes a test for parameter changes. The observations are assumed to follow an exponential distribution. The author presents a derivation using the method of [7]. Ref. [8] studies the problem of changes in the parameters of AR models and the variance in the white noise using the likelihood ratio statistic. Ref. [9] proposes test statistics for detecting a break in the trend function of a dynamic univariate time series. The tests are based on the mean and exponential statistics of [10] and the supremum statistic of [11]. Another method for detecting change points is introduced in [12]. The authors present a multiple-change-point analysis for which the Markov Chain Monte Carlo (MCMC) sampler plays a fundamental role. They propose an attractive methodology for the change-point problem in a Bayesian context. The reversible jump algorithm is presented. Ref. [13] also studies the problem of detecting change points in the mean of a signal corrupted by additive noise. The number of change points is estimated by a method based on a penalized least-square criterion. Ref. [14] uses the minimum description length for detecting change points for a non-stationary time series with an application to GARCH models, stochastic volatility models and generalized state-space models as the parametric model for the segments. Ref. [15] uses maximum likelihood to estimate the instant of the change. The authors study the asymptotic distribution of their test by contiguity. Ref. [16] investigates the regression function or its  $\nu$ th derivative in generalized linear models which may have a change (discontinuity) point at an unknown location. Ref. [17] studies change points in the mean of a sequence of independent normally distributed random vectors. The asymptotic distribution of the test statistic is studied by using results from [18]. Also, Ref. [19] studies this problem for independent normal means as a multiple testing problem. The authors consider two stepwise methods, the binary segmentation method of [20] and the maximum residual down method of [21]. They prove the consistency of these methods. Ref. [22] studies the existence of changes in the regression parameters in a linear model where the regressors and errors are weakly dependent. They study the asymptotic distribution under the null hypothesis and under contiguous alternatives. In [23], the authors develop a method for detecting and estimating change points in the tail of multiple time series data. They 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, based on multiple cross-sectional time series. Ref. [24] proposes a procedure based on the Bayesian information criterion (BIC) in combination with the binary segmentation algorithm to look for changes in the mean, autoregressive coefficients, and variance in the perturbation in piecewise autoregressive processes. The authors explain briefly the Auto-PARM and Auto-SLEX methods. They present different algorithms useful to the search of multiple change points. Ref. [25] proposes a likelihood ratio scan method for estimating multiple change points in piecewise stationary processes. Ref. [26] aims to estimate the instant of change in a regression model. The authors use a sequential Bayesian change-point algorithm that provides uncertainty bounds on both the number and location of the change. A class of change-point test statistics is proposed in [27] that utilizes a weighting and trimming scheme for the cumulative sum (CUSUM) process inspired by Renyi. Using an asymptotic analysis and simulations, the authors demonstrate that this 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 end of the sample. The

authors develop a generalization of these "Renyi" statistics for testing for changes in the parameters of linear and non-linear regression models, and in the generalized method of moment estimation.

In this paper, we are interest in weak change detection. A weak change is one with a too-small magnitude. Such a change may be a harbinger signaling a forthcoming critical behavior of the phenomenon studied. It can manifest in various domains including economics and finance, public health, bio-science, engineering, climatology, hydrology, linguistics, genomics, signal processing and many others.

Classical change detection methods can fail in detecting weak changes. Therefore, it may be of importance to develop new methods for their detection. In the context of time series, very few studies have tested no change against local alternatives of weak changes. Refs. [4,28] study this problem for the case of testing the mean of the model (1). As changes can happen elsewhere than the mean, it can be interesting to study more general models. Our main purpose in this paper is to extend these works to the conditional mean of (1). With this purpose, we proceed with the same techniques. We first construct a test based on the likelihood ratio, and we study its null distribution. Next, we establish the LAN property for the likelihood families under study. From this, we prove the contiguity of  $H_0$  and  $H_\beta^{(n)}$  and use it together with Le Cam's third lemma to find the asymptotic distribution of the test under  $H_\beta^{(n)}$ . Then, we prove the optimality of our test in the case in which the parameters are known. In the case that the parameters are assumed to be unknown, we prove the convergence of the estimated version of the central sequence based on the parameter estimators to its true version. Finally, we prove that the test remains optimal in the case of unknown parameters. Based on the explicit expression of the power, we construct an algorithm for detecting change points and estimating their locations. The simulation study shows the good performance of our method for detecting weak changes and estimating their locations in the examples considered.

In Section 2, we specify the notation and list some of the main assumptions. In Section 3, we state the theoretical results in the case that  $\rho_0$  is known and in the case that it is unknown. The results of this section are used in Section 4 to construct an algorithm for testing change points and estimating their locations. In Section 5, a simulation experiment is conducted for the application of our algorithm. Section 6 concludes our work, and the last section contains the proofs of the results stated in Section 3.

## 2. Notation and Assumptions

In this section, we specify the notation and list some of the main assumptions needed.

### 2.1. The Notation

In the sequence,  $\mathcal{M}_{m,n}(\mathbb{R})$  is the space of real  $m \times n$  matrices and  $\mathcal{M}_n(\mathbb{R}) = \mathcal{M}_{n,n}(\mathbb{R})$ .  $M^\top$  is the transpose of  $M \in \mathcal{M}_{m,n}(\mathbb{R})$ , and  $\|M\|_{m \times n}$  is its Euclidean matrix norm.  $\|\cdot\|_p$  is the Euclidean norm of  $\mathbb{R}^p$ .

Let  $U \in \mathbb{R}^{p(k+1)}$ ; we write  $U = (U_1^\top, \dots, U_{k+1}^\top)^\top$  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

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

where for  $i, j \in \{1, \dots, k + 1\}$ ,  $M_{i,j} \in \mathcal{M}_p(\mathbb{R})$ .

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

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

where  $\partial_{\gamma_i}[\mathfrak{S}(x)]$  is the gradient of  $\mathfrak{S}$  with respect to  $\gamma_i$  at  $x \in \mathbb{R}^d$ :

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

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

We also denote by  $H_\gamma[\mathfrak{S}(x)]$  the matrix

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

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

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

We denote any differentiable function  $g$  with derivative  $g'$  by

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

For any  $t \in \{1, \dots, n\}$ , let  $\mathcal{F}_t = \sigma(\mathbf{X}_1, \dots, \mathbf{X}_t)$  be the  $\sigma$ -algebra generated by  $\mathbf{X}_1, \dots, \mathbf{X}_t$  such that  $\varepsilon_t$  is independent of  $\mathcal{F}_{t-1}$ .

### 2.2. The Main Assumptions

In this section, we outline the key assumptions needed for our methodology. These are crucial for establishing our theoretical results. Following their enumeration, we include a remark that articulates their significance. So, we assume that

- (A1)  $\int_{\mathbb{R}} x f(x) dx = 0$  and  $\int_{\mathbb{R}} x^2 f(x) dx = 1$ .
- (A2)  $f$  is differentiable with derivative  $f'$ .
- (A3)  $\lim_{x \rightarrow +\infty} f(x) = \lim_{x \rightarrow -\infty} f(x) = 0 = \lim_{x \rightarrow +\infty} f'(x) = \lim_{x \rightarrow -\infty} f'(x)$ .
- (A4)  $\phi_f$  is differentiable with derivative  $\phi'_f$  and is  $c_\phi$ -Lipschitz where  $0 < c_\phi < +\infty$ .
- (A5)  $\max \left\{ \int_{\mathbb{R}} |\phi_f(x)|^3 f(x) dx, \int_{\mathbb{R}} |\phi'_f(x)| f(x) dx \right\} < \infty$ .
- (A6) For any  $j = 1, \dots, k + 1$ ,  $n_j(n)$  designates the number of observations between the instants  $\tau_j$  and  $\tau_{j-1}$ , such that  $n_j(n) \rightarrow +\infty$  and  $n_j(n)/n \rightarrow \alpha_j$ , as  $n$  tends to  $+\infty$ .
- (A7) For all  $j = 1, \dots, k + 1$ , the sequence  $(\mathbf{X}_t)_{t \in \mathbb{Z}}$  is stationary and ergodic on  $[\tau_{j-1}, \tau_j)$  with stationary cumulative distribution function  $F_j$ .

(A<sub>8</sub>) For any  $j = 1, \dots, k + 1, 1 \leq h \leq m \leq p$  and  $b \leq 3$ ,

$$\eta_{j,b}^{(h,m)}(\rho_0, \gamma_0) = I(f) \int_{\mathbb{R}} \left(\frac{1}{V(x)}\right)^b \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_0, \gamma_0, x) dF_j(x) < \infty.$$

(A<sub>9</sub>)  $\max\left\{\sup_{\gamma} |T(\rho_0, \gamma, x)|, \sup_{\gamma} \|\partial_{\gamma}[T(\rho_0, \gamma, x)]\|_{p(k+1)}, \sup_{\gamma} \|H_{\gamma}[T(\rho_0, \gamma, x)]\|_{p(k+1)}\right\} < \nu(x)$ , for some positive function  $\nu$  defined on  $\mathbb{R}^d$ .

(A<sub>10</sub>) For  $j = 1, \dots, k + 1, a, b \in \{1, 2, 3\}$ ,  $\int_{\mathbb{R}^d} \frac{\nu(x)^b}{V(x)^a} dF_j(x) < \infty$ .

(A<sub>11</sub>) The density function of the first  $d$  observations on each interval  $[\tau_{j-1}, \tau_j], j = 1, \dots, k + 1$ , under  $H_{\beta}^{(n)}$  converges to its density function under  $H_0$ .

**Remark 1.**

- Assumptions (A<sub>1</sub>)–(A<sub>5</sub>) are regularity properties required for the density  $f$ . They are satisfied at least by the standard Gaussian density function.
- Assumption (A<sub>6</sub>) allows for the application of the ergodic theorem on each segment  $[\tau_{j-1}, \tau_j]$ . This assumption is very usual in the literature.
- Assumption (A<sub>7</sub>) ensures the ergodicity and stationarity of the process on each segment  $[\tau_{j-1}, \tau_j]$ . It holds at least for piece-wise stationary and ergodic AR and ARCH models.
- Assumptions (A<sub>8</sub>)–(A<sub>10</sub>) are constraints on the function  $T$  and its derivatives. They are satisfied by usual models as parametric AR, ARCH, TAR, and EXPAR models with Gaussian noise.
- Assumption (A<sub>11</sub>) allows for the simplification of the forms of the likelihoods.

**3. The Theoretical Results**

3.1. The Parameters Are Known

We first study the case where  $\rho_0$  and  $\gamma_0$  are assumed to be known. This will enlighten the case where they are unknown. We start by establishing a LAN and contiguity results. We denote by  $\Theta_n(\rho_0, \gamma_0, \beta)$  the log-likelihood ratio of  $H_0$  against  $H_{\beta}^{(n)}$ , and we define the sequence  $\Pi_n$  by

$$\Pi_n(\rho_0, \gamma_0, \beta) = \frac{1}{\sqrt{n}} \sum_{j=1}^{k+1} \sum_{t=\tau_{j-1}}^{\tau_j} \frac{1}{V(\mathbf{X}_{t-1})} \beta_j^{\top} N(\rho_0, \gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\rho_0, \gamma_0)], \tag{4}$$

where

$$N(\rho_0, \gamma_0, \mathbf{X}_{t-1}) = \omega(t) \odot D_{\gamma}[T(\rho_0, \gamma_0, \mathbf{X}_{t-1})] \in \mathbb{R}^p,$$

and for all  $\gamma = (\gamma_1, \dots, \gamma_{k+1}) \in \mathbb{R}^{p(k+1)}$ ,

$$\varepsilon_t(\rho_0, \gamma) = \frac{\mathbf{X}_t - T(\rho_0, \gamma, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})}, \quad t \in \mathbb{Z}. \tag{5}$$

**Theorem 1 (LAN).** Assume that (A<sub>1</sub>)–(A<sub>10</sub>) hold. Then, for any  $\beta \in \mathbb{R}^{p(k+1)}$ , under  $H_0$ , as  $n \rightarrow +\infty$ ,

$$\Theta_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_0, \gamma_0, \beta) - \frac{\eta(\rho_0, \gamma_0, \beta)}{2} + o_p(1),$$

$$\Pi_n(\rho_0, \gamma_0, \beta) \xrightarrow{D} \mathcal{N}(0, \eta(\rho_0, \gamma_0, \beta)),$$

with

$$\eta(\rho_0, \gamma_0, \beta) = \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0),$$

$$\eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_0, \gamma_0, x) dF_j(x).$$

**Proof.** See Appendix A.  $\square$

**Corollary 1.** Assume that (A<sub>1</sub>)–(A<sub>10</sub>) hold. Then, for any  $\beta \in \mathbb{R}^{p(k+1)}$ , the sequences  $\{H_\beta^{(n)} : n \geq 1\}$  and  $\{H_0^{(n)} = H_0 : n \geq 1\}$  are contiguous. Furthermore, under  $H_\beta^{(n)}$ , as  $n \rightarrow +\infty$ ,

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

**Proof.** See Appendix A.  $\square$

For known  $\rho_0 \in \mathbb{R}^p$  and  $\gamma_0 \in \mathbb{R}^{p(k+1)}$ , 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

$$\mathcal{T}_n(\rho_0, \gamma_0, \beta) = \frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\widehat{\vartheta}_n(\rho_0, \gamma_0, \beta)}, \tag{6}$$

where  $\widehat{\vartheta}_n(\rho_0, \gamma_0, \beta)$  is any consistent estimator of  $\vartheta(\rho_0, \gamma_0, \beta) = \eta^{\frac{1}{2}}(\rho_0, \gamma_0, \beta)$ .

At the level of significance of  $\alpha \in (0, 1)$ , we reject  $H_0$  whenever  $\mathcal{T}_n(\rho_0, \gamma_0, \beta) > Z_\alpha$ , where  $Z_\alpha$  is a  $(1 - \alpha)$ -quantile of the standard Gaussian distribution.

In practice,  $\widehat{\vartheta}_n(\rho_0, \gamma_0, \beta)$  can be taken as a natural estimator of  $\widehat{\eta}_n^{\frac{1}{2}}(\rho_0, \gamma_0, \beta)$  with  $\widehat{\eta}_n(\rho_0, \gamma_0, \beta) = \sum_{j=1}^{k+1} \widehat{\alpha}_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \widehat{\eta}_{j,2}^{(h,m)}(\rho_0, \gamma_0)$ , and for  $j = 1, \dots, k + 1$ ,  $\widehat{\alpha}_j$  is an estimator of  $\alpha_j = \lim_{n \rightarrow +\infty} n_j(n)/n$  and

$$\widehat{\eta}_{j,2}^{(h,m)}(\rho_0, \gamma_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_0, \gamma_0, x) d\widehat{F}_j(x),$$

where  $\widehat{F}_j$  is the empirical distribution function of the observations with indices in  $[\tau_{j-1}, \tau_j)$ . This can be written again as

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

**Theorem 2 (Optimality).** Assume that (A<sub>1</sub>)–(A<sub>10</sub>) hold. Then, for any given  $\beta \in \mathbb{R}^{p(k+1)}$ ,

- [i] Under  $H_0$ , as  $n \rightarrow +\infty$ ,  $\mathcal{T}_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$ .
- [ii] Under  $H_\beta^{(n)}$ , at the level of significance of  $\alpha \in [0, 1]$ , the asymptotic power of the test based on  $\mathcal{T}_n(\rho_0, \gamma_0, \beta)$  is

$$\mathcal{P}_{k,\tau^k} = 1 - \Phi(z_\alpha - \vartheta(\rho_0, \gamma_0, \beta)), \tag{7}$$

where  $z_\alpha$  is the  $(1 - \alpha)$ -quantile of a standard Gaussian distribution with cumulative distribution function  $\Phi$ .

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

**Proof.** See Appendix A.  $\square$

### 3.2. The Parameters Are Unknown

Here, we place ourselves in the framework of Model (1) with  $\rho_0$  unknown. We study the case in which  $\gamma_0$  is known and the case in which it is unknown. We previously studied the asymptotic normality of an estimator of  $\rho_0$  under  $H_0$  and under  $H_\beta^{(n)}$ . For any  $t = 1, \dots, n$ ,  $\rho \in \mathbb{R}^p$  and  $\gamma \in \mathbb{R}^{p(k+1)}$ , define

$$\varepsilon_t(\rho, \gamma, \mathbf{X}_{t-1}) = \frac{X_t - T(\rho + \gamma \odot \omega(t), \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})}. \tag{8}$$

We consider the following additional assumptions:

(B<sub>1</sub>) The model is identifiable, that is, for  $\gamma_1, \gamma_2 \in \mathbb{R}^{p(k+1)}$ ,  $\gamma_1 \neq \gamma_2 \implies T(\rho_0 + \gamma_1 \odot \omega, x) \neq T(\rho_0 + \gamma_2 \odot \omega, x)$ ,  $x \in \mathbb{R}^d$ ,  $\omega \in \{0, 1\}^{k+1}$ ,

(B<sub>2</sub>) The true parameter  $\rho_0$  has a consistent estimator  $\rho_n$  that satisfies the Bahadur representation (see, e.g., [29]), given by

$$n^{\frac{1}{2}}(\rho_n - \rho_0) = n^{-\frac{1}{2}} \sum_{t=1}^n Y(\rho_0, \mathbf{X}_{t-1}) \aleph(\varepsilon_t(\rho_0, \gamma_0)) + o_p(1), \tag{9}$$

where

- $Y(x, \rho_0) = (Y_1(x, \rho_0), \dots, Y_p(x, \rho_0))^T \in \mathbb{R}^p$ .
- For any  $j = 1, \dots, k + 1$ ,  $x \in \mathbb{R}^d$ ,  $\exists \varrho \geq 0$  such that  $\int_{\mathbb{R}^d} \|Y(\rho_0, x)\|^{2+\varrho} dF_j(x) < \infty$ .
- $\int_{\mathbb{R}} |\aleph(x)|^{2+\varrho} f(x) dx < \infty$  and  $\int_{\mathbb{R}} \aleph(x) f(x) dx = 0$ .

(B<sub>3</sub>) For any  $j \in \{1, \dots, k + 1\}$ ,  $h \in \{1, \dots, p\}$ ,

$$\max_{1 \leq \ell \leq p} \int_{\mathbb{R}^d} \left| \frac{Y_\ell(\rho_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, x) \right| dF_j(x) < \infty,$$

(B<sub>4</sub>) For any  $i = 1, \dots, k + 1$  and  $j = 1, \dots, p$ , there exists a ball  $B(r)$  of radius  $r$ , such that

$$\max \left\{ \sup_{\rho \in B(r)} \|\partial_\rho T(\rho, \gamma_0, x)\|_p, \sup_{\rho \in B(r)} \left\| \partial_\rho \left( \frac{\partial T}{\partial \gamma_{i,j}} \right) (\rho, \gamma_0, x) \right\|_p, \sup_{\rho \in B(r)} \left\| \left\| \partial_\rho^2 \left( \frac{\partial T}{\partial \gamma_{i,j}} \right) (\rho, \gamma_0, x) \right\|_p \right\|_p \right\} \leq \chi(x),$$

for some positive function  $\chi$  defined on  $\mathbb{R}^d$ ,

(B<sub>5</sub>) For  $j = 1, \dots, k + 1$ ,  $\ell = 1, 2, 3$  and  $a, b \in \{0, 1, 2, 3\}$ ,

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

#### Remark 2.

- In the literature, one can find numbers of models with functions  $T(\rho_0, \cdot)$  satisfying (B<sub>4</sub>) and (B<sub>5</sub>).
- Assumption (B<sub>1</sub>) is useful for the estimation of  $\rho_0$ , while (B<sub>2</sub>) helps for the study of the distribution of the test statistic. It has been used before in [4]. It is satisfied by least-squares and likelihood type estimators for some usual models within (1).



Recall that, for any  $\beta \in \mathbb{R}^{p(k+1)}$ , under  $H_0$ , the central sequence with the true parameter  $\rho_0$  is denoted by  $\Pi_n(\rho_0, \gamma_0, \beta)$  and its estimated version by  $\Pi_n(\rho_n, \gamma_0, \beta)$ .

**Proposition 1.** Under the assumptions  $(A_1)$ – $(A_{10})$  and  $(B_1)$ – $(B_2)$ , we have

[i] Under  $H_0$ :

$$\sqrt{n}(\rho_n - \rho_0) \xrightarrow{D} \mathcal{N}(0, \Sigma),$$

[ii] Under  $H_\beta^{(n)}$ :

$$\sqrt{n}(\rho_n - \rho_0) \xrightarrow{D} \mathcal{N}(C, \Sigma),$$

where

$$C = \int_{\mathbb{R}} \aleph(x) \phi_f(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \sum_{h=1}^p \beta_{j,h} \int_{\mathbb{R}^d} \frac{Y(\rho_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, x) dF_j(x) \in \mathbb{R}^p$$

and

$$\Sigma = \int_{\mathbb{R}} \aleph^2(x) f(x) dx \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} Y(\rho_0, x) Y^\top(\rho_0, x) dF_j(x) \in \mathcal{M}_p(\mathbb{R}).$$

**Proof.** See Appendix A.  $\square$

### 3.2.1. The Parameter $\gamma_0$ Is Known

As explained in [4], in practice, the case where the parameter  $\gamma_0$  is known may be encountered when there is no apparent change, and one wishes to test for possible weak changes. That is the situation where  $\gamma_0 = 0$ . This is what is usually tested in the literature. Recall that

$$\eta(\rho_0, \gamma_0, \beta) = \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0).$$

Note that, by our assumptions, the following real numbers

$$\eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_0, \gamma_0, x) dF_j(x)$$

are finite. Furthermore, since for any  $j = 1, \dots, k + 1$ ,  $\eta_{j,2}(\rho_0, \gamma_0)$  depends on  $\rho_0$  and on  $F_j$ , which itself depends on  $\rho_0$  (which is unknown) and on  $\gamma_0$ , we estimate it by  $\hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0)$ , given by

$$\hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0) = \frac{I(f)}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j-1} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_n, \gamma_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_n, \gamma_0, x). \tag{10}$$

Despite the fact that we consider any consistent estimators of  $\eta(\rho_0, \gamma_0, \beta)$  and  $\vartheta(\rho_0, \gamma_0) = \eta^{\frac{1}{2}}(\rho_0, \gamma_0, \beta)$ , we will take them here to be, respectively,

$$\hat{\eta}_n(\rho_n, \gamma_0, \beta) = \sum_{j=1}^{k+1} \hat{\alpha}_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \hat{\eta}_{j,2}^{(h,m)}(\rho_n, \gamma_0)$$

and

$$\widehat{\vartheta}_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \widehat{\eta}_n^{\frac{1}{2}}(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}),$$

where for all  $j = 1, \dots, k + 1$ ,  $\widehat{\alpha}_j$  is an estimator of  $\alpha_j$  which can be taken to be  $\widehat{\alpha}_j = n_j(n)/n$ .

**Proposition 2.** Under the assumptions  $(A_1)$ – $(A_{10})$  and  $(B_1)$ – $(B_5)$ , for any  $n \geq 0$ , we have, for any sequence of positive integers  $s(n)$  such that  $n/s(n) \rightarrow 0$  as  $n \rightarrow +\infty$ ,

[i]

$$\Pi_n(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) + o_P(1), \tag{11}$$

[ii]

$$\widehat{\vartheta}_n(\boldsymbol{\rho}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \longrightarrow \vartheta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}). \tag{12}$$

**Proof.** See Appendix A.  $\square$

In order to test  $H_0$  against  $H_{\boldsymbol{\beta}}^{(n)}$ , for any  $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$ , we consider the following statistic:

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

**Theorem 3 (Optimality).** Assume that  $(A_1)$ – $(A_{10})$  and  $(B_1)$ – $(B_5)$  hold. Then, for any given  $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$  and for any sequence  $s(n)$  of positive integers such that,  $n/s(n) \rightarrow 0$  as  $n \rightarrow +\infty$ , we have the following:

[i] Under  $H_0$ , as  $n \rightarrow +\infty$ ,  $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta}) \xrightarrow{D} \mathcal{N}(0, 1)$ .

[ii] Under  $H_{\boldsymbol{\beta}}^{(n)}$ , at the 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  $\mathcal{P}_{k, \tau^k} = 1 - \Phi(z_\alpha - \vartheta(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0, \boldsymbol{\beta}))$ , where  $z_\alpha$  is the  $(1 - \alpha)$ -quantile of the standard Gaussian distribution with cumulative distribution function  $\Phi$ .

[iii] The test based on the statistic  $\mathcal{T}_n(\boldsymbol{\rho}_{s(n)}, \boldsymbol{\gamma}_0, \boldsymbol{\beta})$  is locally asymptotically optimal.

**Proof.** See Appendix A.  $\square$

### 3.2.2. The Parameter $\boldsymbol{\gamma}_0$ Is Unknown

In practice,  $\boldsymbol{\gamma}_0$  is generally unknown and has to be estimated, as well as  $\boldsymbol{\gamma}_n$ , where, for any  $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$ ,  $\boldsymbol{\gamma}_n = \boldsymbol{\gamma}_0 + \boldsymbol{\beta}/\sqrt{n}$ . Many methods can be used to obtain consistent estimators of these parameters. Let  $\widehat{\boldsymbol{\gamma}}_{0,n}$  be the maximum likelihood estimator of  $\boldsymbol{\gamma}_0$  under  $H_0$  and let  $\widehat{\boldsymbol{\gamma}}_n$  be the maximum likelihood estimator of  $\boldsymbol{\gamma}_n$  under  $H_{\boldsymbol{\beta}}^{(n)}$ . Then, we easily have that in probability, asymptotically,

$$\widehat{\boldsymbol{\gamma}}_n = \widehat{\boldsymbol{\gamma}}_{0,n} + \boldsymbol{\beta}/\sqrt{n}.$$

The above equality allows for the study of the test statistics in the same lines as in the case where  $\boldsymbol{\gamma}_0$  is known.

We need the following assumptions:

$$(\mathcal{B}'_1) \quad \text{For any } m = 1, \dots, k + 1, h = 1, \dots, p, \text{ for } T^{\gamma_{m,h}}(\rho, \gamma, x) = \frac{\partial T}{\partial \gamma_{m,h}}(\rho, \gamma, x),$$

$$\max \left\{ \sup_{\gamma} |T^{\gamma_{m,h}}(\rho, \gamma, x)|, \sup_{\gamma} \|\partial_{\gamma} T^{\gamma_{m,h}}(\rho, \gamma, x)\|_{p(k+1)}, \sup_{\gamma} \|\partial_{\gamma}^2 T^{\gamma_{m,h}}(\rho, \gamma, x)\|_{p(k+1)} \right\} < \kappa(x),$$

for some positive function  $\kappa$  defined on  $\mathbb{R}^d$ ,

$$(\mathcal{B}'_2) \quad \text{For } j = 1, \dots, k + 1, u = 1, 2, 3 \text{ and } a, b \in \{0, 1, 2, 3\},$$

$$\delta_{a,b}^{(u)} = \int_{\mathbb{R}^d} \frac{v^a(x) \kappa^b(x)}{V^u(x)} dF_j(x) < \infty.$$

$$(\mathcal{B}'_3) \quad \sqrt{n}(\hat{\gamma}_{0,n} - \gamma_0) = O_P(1).$$

Let  $s(n)$  be any sequence of positive integers satisfying  $n/s(n) \rightarrow 0$  as  $n \rightarrow +\infty$ . For testing  $H_0$  against  $H_{\beta}^{(n)}$ ,  $\beta \in \mathbb{R}^{p(k+1)}$ , we use the test based on the statistic

$$\mathcal{T}_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta) = \frac{\Pi_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta)}{\hat{\vartheta}_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta)}.$$

**Proposition 3.** Assume that  $(A_1)–(A_{10})$ ,  $(\mathcal{B}_1)–(\mathcal{B}_5)$  and  $(\mathcal{B}'_1)–(\mathcal{B}'_3)$  hold. Then, for any sequence  $s(n)$  of positive integers satisfying  $n/s(n) \rightarrow 0$  as  $n \rightarrow +\infty$ , for any sequence of consistent and asymptotically normal estimators  $\{\hat{\gamma}_{0,n}\}_{n \geq 1}$  of  $\gamma_0$  and for any  $\beta \in \mathbb{R}^{p(k+1)}$ , we have, under  $H_0$  and as  $n \rightarrow +\infty$ ,

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

**Proof.** See Appendix A.  $\square$

**Theorem 4 (Optimality).** Assume that  $(A_1)–(A_{10})$ ,  $(\mathcal{B}_1)–(\mathcal{B}_5)$  and  $(\mathcal{B}'_1)–(\mathcal{B}'_3)$  hold. Then, for any given  $\beta \in \mathbb{R}^{p(k+1)}$ , we have

[i] Under  $H_0$ , as  $n \rightarrow +\infty$ ,  $\mathcal{T}_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, 1)$ .

[ii] Under  $H_{\beta}^{(n)}$ , at the level of significance  $\alpha \in ]0, 1[$ , the asymptotic power of the test based on the statistic

$\mathcal{T}_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta)$  is  $\mathcal{P}_{k,\tau^k} = 1 - \Phi(z_{\alpha} - \vartheta(\rho_0, \gamma_0, \beta))$ , where  $z_{\alpha}$  is the  $(1 - \alpha)$ -quantile of the standard Gaussian distribution with cumulative distribution function  $\Phi$ .

[iii] The test based on the statistic  $\mathcal{T}_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta)$  is locally asymptotically optimal.

**Proof.** See Appendix A.  $\square$

#### 4. Application to Detection of Change Points and Their Location Estimation

The time series at hand has jumps if the parameters of its distribution change at certain times. The current test, applied to the model (1) adjusted to this time series, for testing the null hypothesis of no change against at least one change, is conducted with a  $\gamma_0$  whose components are all equal to, say,  $\rho_1$ , the parameter of the stationary distribution on the first segment  $[\tau_0, \tau_1) : \gamma_0 = (\rho_1, \rho_1, \dots, \rho_1)$ .

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. This 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  $\gamma_0$  will no longer be the same, and some of the  $\tau_j$ 's in the model would be considered known. Thus, our test can be used for testing the null hypothesis of  $l$  changes against at least  $l + 1$  changes, for some given  $l \in \mathbb{N}$ .

For any  $k \geq 1$ , denote by  $\widehat{\mathcal{P}}_{k,\tau^k}$  any estimator of the local power  $\mathcal{P}_{k,\tau^k}$  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.

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:**

**(A1):** Take any  $t$  between 1 and  $m + j$ , so that there is a large number of indices before and after  $t$  (for example  $t = \lfloor (m + j)/2 \rfloor$ ).

Adjust Model (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  $|\widehat{\mathcal{P}}_{1,t} - \mathcal{P}_{0,\tau^0}| > \zeta$ ,

Put  $\tau_1 = m + j$  and go to **Location 2** (first change location estimated)

Else

Carry out  $j = j + 1$  and go to **(A1)**.

**Location 2:**

Consider the next  $h$  observations to  $X_{\tau_1}$ :  $X_{\tau_1+1}, \dots, X_{\tau_1+h}$

Put  $j = 1$  and conduct the following:

**(A2):** Take any  $t$  between  $\tau_1$  and  $\tau_1 + h + j$  so that there is a large number of indices before and after  $t$ .

Adjust Model (1) to  $X_{\tau_1}, \dots, X_{\tau_1+h+j}$  with a potential change located at the time index  $t$  and apply the testing procedure studied.

If  $|\widehat{\mathcal{P}}_{1,t} - \mathcal{P}_{0,\tau^0}| > \zeta$ ,

Put  $\tau_2 = \tau_1 + h + j$  and go to **Location 3** (second change location estimated)

Else

Carry out  $j = j + 1$  and go to **(A2)**.

**Location i:**

Let  $\tau_{i-1}$  be the change location at step  $i-1$

Put  $j = 1$  and perform

**(Ai):** Take any  $t$  between  $\tau_{i-1}$  and  $\tau_{i-1} + h + j$  so that there is a large number of indices before and after  $t$ .

Adjust Model (1) to  $X_{\tau_{i-1}}, \dots, X_{\tau_{i-1}+h+j}$  with a potential change located at the time index  $t$  and apply the testing procedure studied

If  $|\widehat{\mathcal{P}}_{1,t} - \mathcal{P}_{0,\tau^0}| > \zeta$ ,

Put  $\tau_i = \tau_{i-1} + h + j$ ,  $i = i + 1$  and go to **Location i** ( $i$ th change location estimated)

Else

Carry out  $j = j + 1$  and go to **(Ai)**.

Note that simple  $\widehat{\mathcal{P}}_{k,\tau^k}$  can be obtained by plugging estimators of the parameters into the expressions of the local power given in Theorems 2–4.

**5. Simulation Experiment**

In this section, the theoretical results are applied to simulated data, using the software R 4.2.0. We first study the power of the test as a function of the magnitude of the breaks when these are given. Next, we use the power for estimating the location of the breaks when they are no more assumed to be fixed. The results we present in the sequel are

obtained for the nominal levels  $\alpha = 1\%, 5\%, 10\%$ . Almost all the estimators in this section are computed from 5000 replications. We use the following particular CHARN model:

$$X_t = \rho_{0,1} + \gamma_{0j,1} + \frac{\beta_{j,1}}{\sqrt{n}} + \left( \rho_{0,2} + \gamma_{0j,2} + \frac{\beta_{j,2}}{\sqrt{n}} \right) X_{t-1} e^{\left( \rho_{0,3} + \gamma_{0j,3} + \frac{\beta_{j,3}}{\sqrt{n}} \right) X_{t-1}^2} + (\theta_1 + \theta_2 X_{t-1}^2)^{\frac{1}{2}} \varepsilon_t, \quad j = 1, \dots, k, \quad t \in \mathbb{Z}, \tag{13}$$

where  $n$  denotes the number of observations,  $(\varepsilon_t)_t$  is a standard white noise with a differentiable density  $f$ . Here, on  $[\tau_{j-1}, \tau_j]$ ,  $\rho_0 = (\rho_{0,1}, \rho_{0,2}, \rho_{0,3}) \in \mathbb{R}^3$ ,  $\gamma_{0j} = (\gamma_{0j,1}, \gamma_{0j,2}, \gamma_{0j,3})$ ,  $\beta_j = (\beta_{j,1}, \beta_{j,2}, \beta_{j,3}) \in \mathbb{R}^3$ ;  $\rho_0$ ,  $\theta_1$ ,  $\theta_2$  and  $\gamma_0$  are parameters to be specified in each particular model considered.

### 5.1. Simulation Methodology

Our methodology for conducting the simulation experiment works as follows. For some given and fixed values of  $\rho_0 = (\rho_{0,1}, \rho_{0,2}, \rho_{0,3})$  and number of changes  $k$ , for  $1 \leq j \leq k + 1$ , we consider different values of the triplet  $\beta_j = (\beta_{j,1}, \beta_{j,2}, \beta_{j,3})$  corresponding to the shift in the parameter on each interval  $[\tau_{j-1}, \tau_j]$ , with  $\beta_1 = (0, 0, 0)$  (indicating no changes in the first interval). This provides us with  $n_j(n)$ ,  $j = 1, \dots, k + 1$  observations in the  $j$ -th interval. Subsequently, we utilize the model (13) to simulate these observations, to which our algorithm (see Section 4) is then applied.

### 5.2. Power Study for Given Break Locations

#### 5.2.1. $\gamma_0$ and $f$ Are Known

Now, we treat a particular case of (13). We consider  $n = 100$ ,  $n_1(n) = 40$ ,  $n_2(n) = 60$ ,  $\gamma_0 = \mathbf{0}$ ,  $f$  is the standard Gaussian density,  $\theta_1 = 1$ ,  $\theta_2 = 0.02$ ,  $\rho_{0,1} = 0.8$ ,  $\rho_{0,2} = 0.2$ ,  $\rho_{0,3} = \beta_{j,3} = 0$ ,  $T(\rho_0 + \gamma_j \omega_j, x) = \rho_{0,1} + \gamma_{j,1} \omega_j + (\rho_{0,2} + \gamma_{j,2} \omega_j) x$ , and for  $i = 1, 2$ ,  $\gamma_{j,i} = \gamma_{0j,i} + \beta_{j,i} / \sqrt{n}$  with  $\beta_{j,i} \in [-10, 10]$ ,  $V(x) = \sqrt{1 + 0.02x^2}$ ,  $\partial T / \partial \gamma_{j,1}(\rho_0, \gamma_0, x) = 1$  and  $\partial T / \partial \gamma_{j,2}(\rho_0, \gamma_0, x) = x$ . For  $j = 1, 2$ , to study the behavior of the power as a function of magnitudes of the changes, we fix one component of each  $\beta_j$  and we compute the power of the test as a function of the other components. The results are plotted on Figure 1, where one can observe that the power grows quickly to one as the norm of the magnitude grows.

#### 5.2.2. $\gamma_0$ and $f$ Are Unknown

As we said before, in practice,  $\gamma_0$  and  $f$  are unknown and must be estimated. While we estimated  $\gamma_0$  by the least-squares method, we estimate  $f$  by the Parzen–Rosenbatt estimator (see [30]), defined 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

$$\hat{\varepsilon}_t = \frac{X_t - T(\hat{\rho}_n + \hat{\gamma}_n \odot \omega(t), \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})},$$

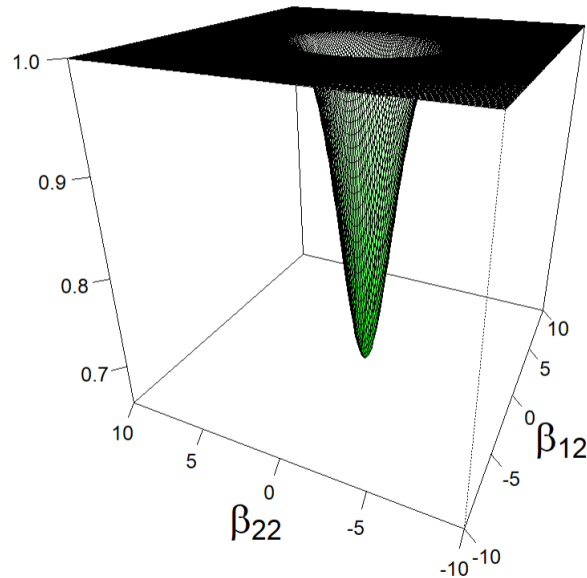
where  $\hat{\rho}_n$  and  $\hat{\gamma}_n$  denote the least-squares estimators of  $\rho_0$  and  $\gamma_0$ , respectively,  $h_n$  is the smoothing parameter, and  $K$  the kernel symmetric function having the following properties:

1.  $K(x) > 0$  for any  $x \in \mathbb{R}$  (positivity).
2.  $\int_{\mathbb{R}} K(x) dx = 1$  (density).
3.  $\int_{\mathbb{R}} xK(x) dx = 0$  (by symmetry).

Now, we calculate the power of the test using these estimators. We choose a Gaussian kernel  $K$  and  $h_n \sim \sigma_n n^{-\frac{1}{5}}$ , with  $\sigma_n$  being the sample standard deviation. We consider a

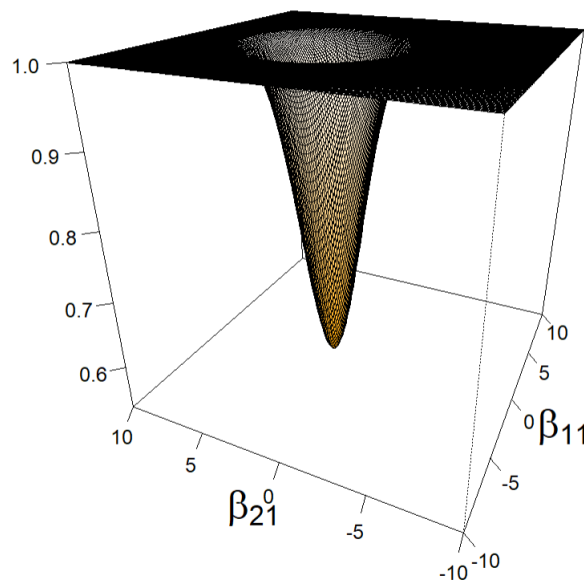
sample of  $n = 100$  observations,  $n_1(n) = 40$ ,  $n_2(n) = 60$  and we generate the observations from (13). The results for  $\rho_{0,1} = 0.8$ ,  $\rho_{0,2} = 0.2$ ,  $\rho_{0,3} = 0$ ,  $\theta_1 = 1$  and  $\theta_2 = 0.02$  at the level of significance 5% are given in Figure 2. It is clear that the local power of the test has approximately the same behavior for the standard Gaussian and the standard Student densities. The results do not change significantly for Epanechnikov, uniform kernel, quadratic kernel, etc.

### Power



(a)  $\beta_{1,1} = 2$  and  $\beta_{2,1} = 2.5$

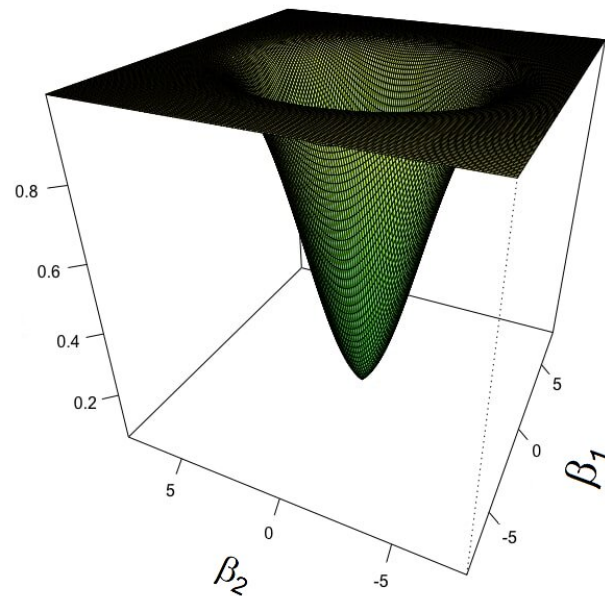
### Power



(b)  $\beta_{1,2} = 1$  and  $\beta_{2,2} = 3.5$

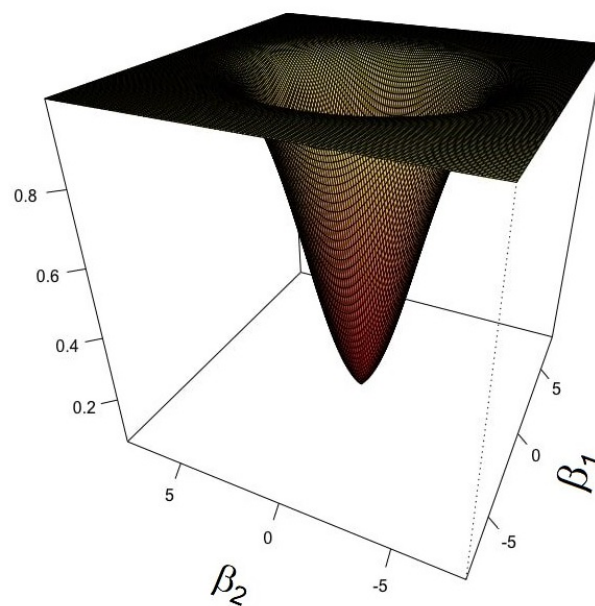
**Figure 1.** Power of the test with respect to  $\beta$  in a class of AR(1) models when  $f$  is a standard Gaussian density.

## Power (Student case)



(a) Student case

## Power (Normal Case)



(b) Normal case

**Figure 2.** Power of the test with respect to  $\beta$  in a class of AR(1) model when  $f$  is the kernel estimated.

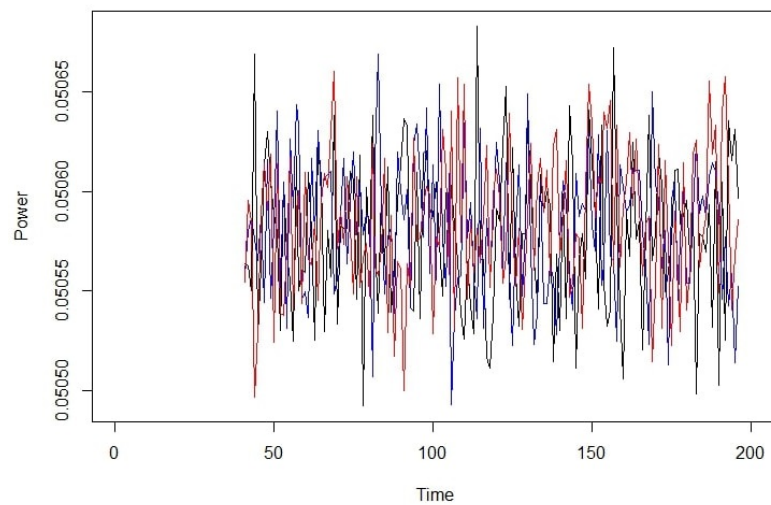
### 5.3. Detection of Change Points and Estimation of Their Locations

In this subsection, we detect change points and we estimate their locations in simulated data. Ref. [4] studied the case of changes only in  $\rho_{0,1}$ . We start by evaluating the power of the test in case of no break in the data. Next, we study changes in  $\rho_{0,2}$ . Finally, we study changes in  $\rho_{0,1}$  and  $\rho_{0,2}$  simultaneously.

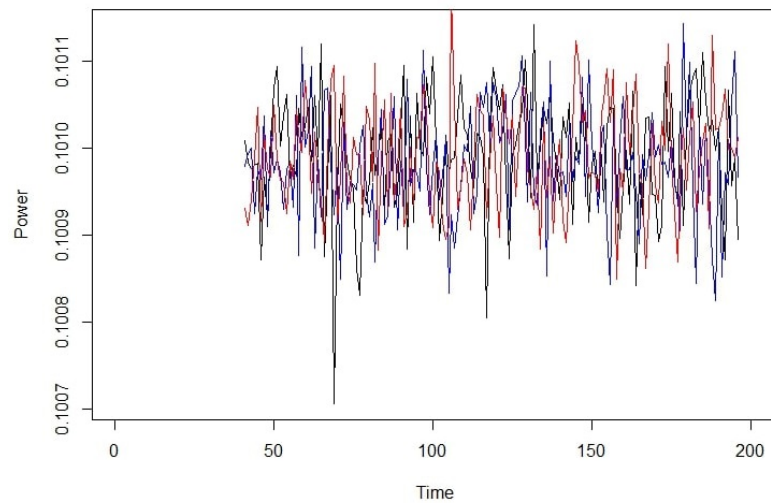
### 5.3.1. No Break

Following the algorithm in Section 4, we start by calculating the asymptotic local power given by (A8) in the case where there is no break, that is, for  $k = 0$ . We consider a sample of  $n = 200$  observations generated from Model (13), 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 Figure 3. We can see there that the local power does not exceed 0.1012 when  $\alpha = 10\%$  and does not exceed 0.0507 when  $\alpha = 5\%$ . Then, for any  $\alpha = 5\%$ , or  $10\%$ , for the thresholds corresponding to  $\zeta = 0.002$  and  $0.0008$ , respectively, we keep the null hypothesis and conclude that there is no break in the data.



(a) No break ( $\alpha = 5\%$ )



(b) No break ( $\alpha = 10\%$ )

**Figure 3.** No break in the data.

### 5.3.2. Case of One Single Break

Here, we consider the problem of detecting one single break when it happens jointly in  $\rho_{0,1}$  and  $\rho_{0,2}$ . For  $\alpha = 5\%$ ,  $n = 200$ ,  $\rho_{0,1} = 0.5$ ,  $\rho_{0,2} = -0.2$  and  $\rho_{0,3} = \beta_{1,3} = 0$ , for different values of  $\tau_1$  and  $\beta_1 = (\beta_{1,1}, \beta_{1,2})$ ; the estimation of the break location, as well as the root mean square error (RMSE), is presented in Table 1. One can see from this table that the estimation is accurate and that the RMSE is large for smaller  $\|\beta\|$ .



**Table 1.** Break location in an AR(1) model with the corresponding RMSE.

$\tau_1$	$\begin{pmatrix} \beta_{1,1} \\ \beta_{1,2} \end{pmatrix}$				
	$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 3 \end{pmatrix}$	$\begin{pmatrix} -3 \\ 2 \end{pmatrix}$
80	78 (4.32)	81 (1.41)	82 (4.01)	80 (0.654)	80 (0.34)
100	99 (6.96)	99 (4.67)	102 (3.34)	101 (1.23)	100 (0.54)
120	122 (5.12)	121 (2.32)	119 (2.45)	120 (0.23)	121 (1.23)
140	143 (10.24)	142 (2.22)	141 (2.35)	140 (0.22)	141 (1.12)
160	164 (11.65)	164 (4.43)	162 (2.23)	161 (1.32)	161 (1.12)
185	188 (8.21)	189 (5.32)	190 (6.33)	190 (7.43)	189 (8.23)

5.4. Case of Three Breaks ( $k = 3$ )

Now, we study the case of three breaks when piece-wise models AR(1) and AR(1)-ARCH(1) are adjusted to the data. Note that these models are sub-classes of CHARN(1,1) models.

5.4.1. AR(1) Models

We start with AR(1) models. For  $\tau = (\tau_1, \tau_2, \tau_3)$ , the data are obtained from (13); for  $j = 1, 2, 3, \rho_{0,3} = \beta_{j,3} = 0, \theta_2 = 0, \theta_1 = 1$ , and  $(\varepsilon_t)$  is a sequence of standard Gaussian white noise. The number of change points is assumed to be unknown, and we aim to detect them and estimate their locations using our theoretical results and following our algorithm. For 5000 replications,  $n = 400, \rho_0 = (\rho_{0,1}, \rho_{0,2}) = (0.2, 0.3), \tau = (\tau_1, \tau_2, \tau_3) = (90, 190, 275)$ , for different values of the magnitude of change  $\beta_j = (\beta_{j,1}, \beta_{j,2})^\top, j = 1, 2, 3$ , and for the same threshold  $\zeta = 0.1\%$ , the estimations obtained are displayed in Table 2 together with their associated RMSE (in brackets). These results seem to show that our method tends to estimate the correct number of changes but overestimates their locations with a relatively large RMSE when the jumps in the parameters of the AR(1) models are too small. Again for an AR(1) model, for  $j = 1, 2, 3$ , we fix  $\beta_j = ((3, 2), (1, 3), (-1, 1))^\top$ , and the instants of breaks  $\tau = (90, 190, 275)$ , and we monitor the corresponding break estimates with respect to the variation in the threshold corresponding to different  $\zeta$ . For  $\zeta = 0.07\%$ , our method overestimates the number of changes (six instants detected). For  $\zeta = 0.1\%$  and  $0.15\%$ , it estimates the correct number of changes but overestimates their locations. For  $\zeta = 0.25\%$ , it underestimates the number of changes and overestimates their locations. The overestimation of the break locations may be explained by the weakness of the magnitude of changes. If we consider the same study for  $\beta = ((5, 3), (-1 - 2), (4, 6))^\top$ , we obtain the same results for the number of changes, but with more accurate change location estimations.

5.4.2. AR(1)-ARCH(1) Models

Here, we consider Model (13) for  $\rho_{0,1} = 0.2, \rho_{0,2} = 0.3, \beta_1 = (3, -2)^\top, \beta_2 = (1, 1)^\top, \beta_3 = (-2, 4)^\top, \theta_1 = 1, \theta_2 = 0.02$  and  $\rho_{0,3} = \beta_{j,3} = 0, j = 1, 2, 3$ , which leads to an AR(1)-ARCH(1) model. For  $n = 350$ , Table 3 shows the estimation of change locations corresponding to the same  $\zeta$  and different magnitudes of change. We can see that our method estimates the correct number of changes but overestimates their locations with a relatively large RMSE.

**Table 2.** Break location estimation in a class of AR(1) models for a fixed  $\zeta$ .

$(\beta_1, \beta_2, \beta_3)$	$\tau = (\tau_1, \tau_2, \tau_3) = (90, 190, 275)$ and $\zeta = 0.1$
$\left( \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$	(93,193,277) (5.78)
$\left( \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right)$	(93,194,278) (9.87)
$\left( \begin{bmatrix} -2 \\ 1.5 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -0.5 \\ -1 \end{bmatrix} \right)$	(92,193,277) (8.99)

**Table 3.** Break location estimation in AR-ARCH models.

$(\beta_1, \beta_2, \beta_3)$	$\tau = (\tau_1, \tau_2, \tau_3) = (90, 190, 275)$ and $\zeta = 0.1$
$\left( \begin{bmatrix} 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right)$	(93,193,277) (9.43)
$\left( \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \end{bmatrix} \right)$	(93,194,278) (10.64)
$\left( \begin{bmatrix} -2 \\ 1.5 \end{bmatrix}, \begin{bmatrix} 1 \\ 3 \end{bmatrix}, \begin{bmatrix} -0.5 \\ -1 \end{bmatrix} \right)$	(92,193,277) (6.10)

### 5.4.3. Conclusions

Based on the previous simulation results, we can conclude that our method is sensitive to the choice of  $\zeta$ , and is efficient in detecting weak changes and estimating their locations in an AR(1)-ARCH(1) model we have considered when the magnitudes of the changes are not too small.

### 5.5. Comparison with [27]

In a class of shifted models, Ref. [28] performed a comparison between her method, which is a different case from ours, and other methods including the one of [27]. She concluded that her method is more efficient for estimating weak break locations.

In this section, we compare our method to that of [27], denoted by SCUSUM, for a class of more general models. Recalling Model (13), we consider many cases of one single break corresponding to different instants  $\tau_1$ , and we take  $\rho_{0,3} = \beta_{1,3} = \theta_2 = 0$  and  $\theta_1 = 1$ . For  $n = 200$ ,  $\alpha = 5\%$ ,  $\rho_{0,1} = 0.5$ ,  $\rho_{0,2} = -0.2$  and different values of  $\beta_1 = (\beta_{1,1}, \beta_{1,2})$  we perform 1000 replications, and at each replication, the change location is estimated by SCUSUM and by our method. Table 4 shows the results obtained.

For most of the 1000 replications, SCUSUM was not able to detect any change. For that reason, we kept only the cases where it detected a change, and we calculated the mean of the change locations estimated. The results are displayed in Table 4, from which it is obvious that our method is more accurate than SCUSUM for the detection of weak changes in the parameters of the AR(1) model studied.

**Table 4.** Break location estimation obtained by our method and SCUSUM for different instants of break  $\tau_1$  and different magnitudes of change.

$\tau_1 = 80$	$\begin{pmatrix} \beta_{1,1} \\ \beta_{1,2} \end{pmatrix}$				
	$\begin{pmatrix} 2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -2 \\ 3 \end{pmatrix}$	$\begin{pmatrix} -3 \\ 2 \end{pmatrix}$
Our method	83	81	82	80	80
SCUSUM	95	99	86	122	99
$\tau_1 = 120$					
Our method	122	121	121	120	121
SCUSUM	110	114	105	125	122
$\tau_1 = 185$					
Our method	189	188	186	186	185
SCUSUM	105	103	122	114	101

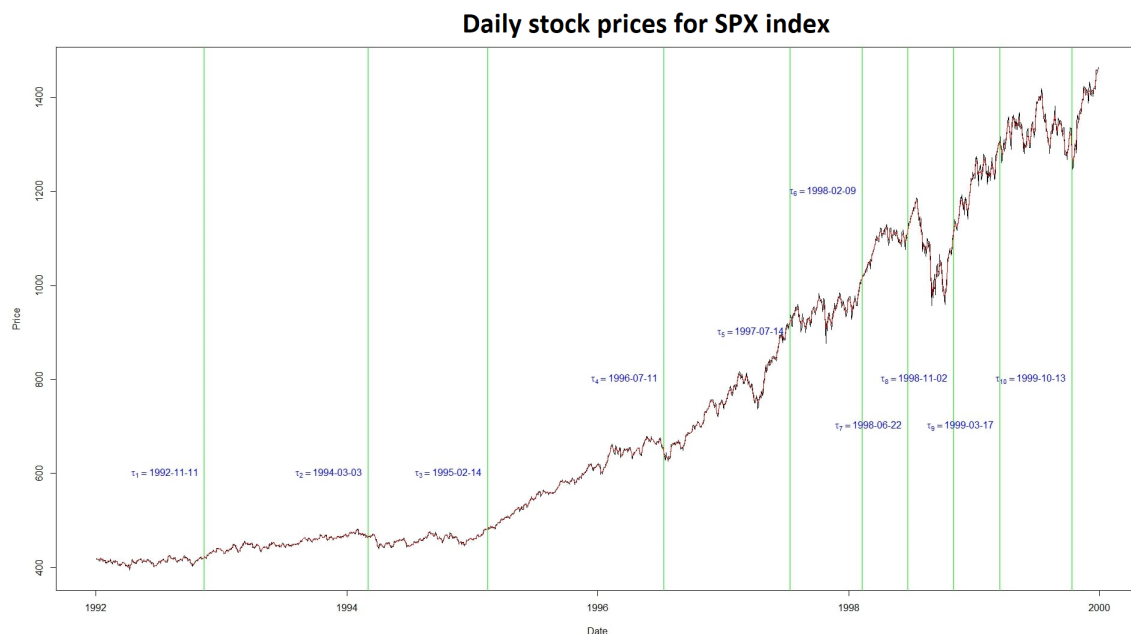
5.6. Application to Real Data

Here, we applied our methodology to detecting changes in the log S&P stock price data obtained from the website <https://finance.yahoo.com/quote/>, accessed on 1 May 2024. These daily data cover the period from January 1992 to December 2000 and represent one of the most closely followed stock market indices worldwide, serving as a significant indicator of the U.S. economy. The raw data exhibit a trend, which shows that the S&P 500 index is non-stationary (see Figure 4). With this, our methodology can not be directly applied to this series.

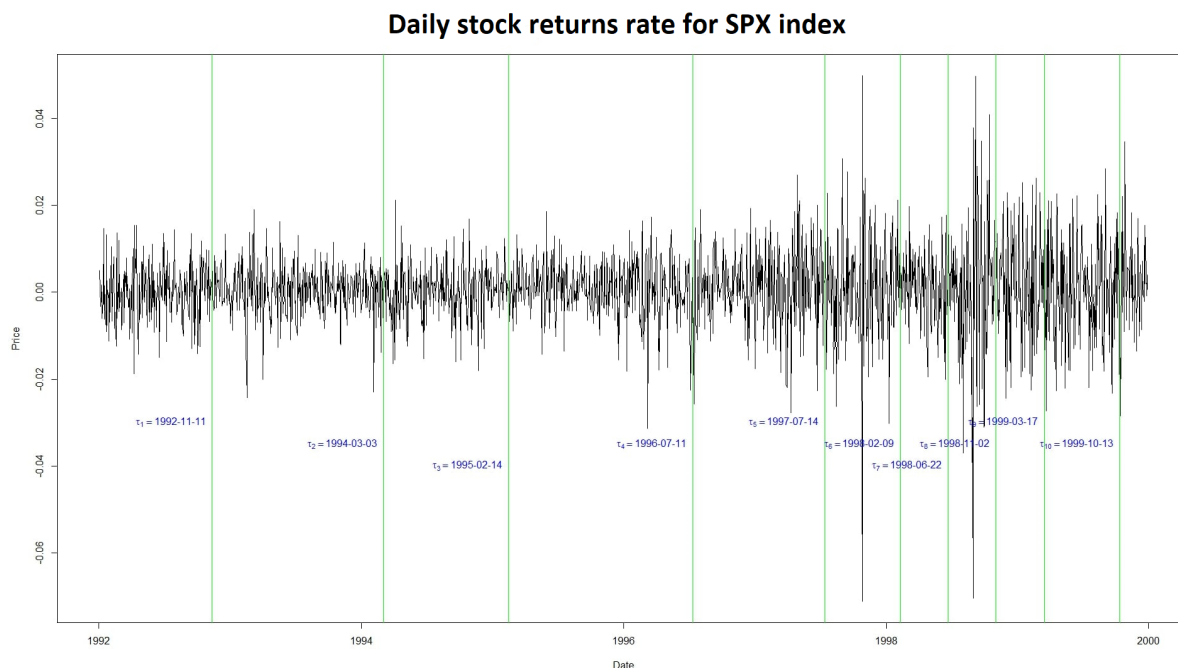
Let  $P_t$  denote the S&P 500 stock price index on day  $t$ , and define  $X_t$  as

$$X_t = \log\left(\frac{P_t}{P_{t-1}}\right)$$

The function  $\log$  being monotonic implies that the change-point locations in  $(P_t)$ ,  $(\log(P_t))$ , and  $(X_t)$  are identical. Graphically (refer to Figure 5),  $X_t$  appears to be approximately piece-wise stationary over a finite number of segments, which aligns with the requirements for applying our methodology to study changes in the raw data.



**Figure 4.** Estimated change points in the S&P 500 indices.



**Figure 5.** Estimated change points in the residual series of S&P 500 indices.

To accommodate these characteristics, we adjust the CHARN model  $X_t = \beta_j/\sqrt{n} + \theta_j \varepsilon_t$  within each segment  $[\tau_j, \tau_{j+1})$ , where  $\varepsilon_t \sim \mathcal{N}(0, 1)$ . The Gaussian assumption is validated by applying the Shapiro–Wilk test.

Then, applying our procedure to this model, we obtained the following break location dates: 1992-11-11, 1994-03-03, 1995-02-14, 1996-07-11, 1997-07-14, 1998-02-09, 1998-06-22, 1998-11-02, 1999-03-17, and 1999-10-13. The changes occurring in 1992 can be linked to the damage caused by the hurricane Andrew or by the European Monetary System crisis. The one in 1994 can be associated with the U.S. lifting of the trade embargo on Vietnam. Those in 1995 can be due to the bankruptcy of the Barings bank. That in 1997 may be associated with the Asian crisis. Those in 1998 may be connected to the rescue organized by the New York Federal Reserve Bank. Finally, those of 1999 can be associated with the cancellation of the 1933 Glass–Steagall Act by the so-called Grammi–Leachi–Bliley Act.

### 6. Conclusions

We generalized the work of [4] to a class of more general CHARN models. We studied weak breaks in the parameters of the function  $T$  when the function  $V$  and the parameters  $\rho_0$  and  $\gamma_0$  are known. We established a LAN and contiguity results. We given an explicit expression of the local power of the test.

Next, we studied the case where  $\rho_0$  is unknown and  $\gamma_0$  is known or unknown. We estimated these parameters, and proved the convergence of the central sequence based on the estimated parameters to the one based on the true parameters. In this case, we proved that the test remains optimal if we replace the parameters with their estimators. From these results, we used the theoretical power for detecting weak breaks and estimating their locations in time series through an algorithm that we constructed.

The simulation experiment conducted shows that our method can detect weak breaks in the parameters of the linear AR(1) and the non-linear AR(1)-ARCH(1) models considered. Also, the location of the breaks as well as their number can be accurately estimated when the magnitudes are not too small.

Compared to [27], it seems to be more efficient for estimating weak break locations. Sometimes, the method in [27] detects breaks in data simulated with no break. This did not happen with our method when we chose a suitable  $\zeta$ . Our method was also applied to a set of financial data.

**Author Contributions:** Methodology, Y.S. and J.N.-W.; Software, Y.S.; Validation, J.N.-W. and Z.K.; Investigation, Y.S.; Writing—original draft, Y.S., J.N.-W. and Z.K.; Writing—review & editing, J.N.-W.; Visualization, Z.K.; Supervision, J.N.-W. and Z.K.; Project administration, J.N.-W. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

**Conflicts of Interest:** The authors have no conflicts of interest to declare.

### Appendix A. Proofs

This section provides the proofs of the results stated in the preceding sections.

#### Appendix A.1. Proof of Theorem 1

For any  $\beta \in \mathbb{R}^{p(k+1)}$ , the log-likelihood ratio of  $H_0$  against  $H_\beta^{(n)}$  is given by

$$\Theta_n(\rho_0, \gamma_0, \beta) = \sum_{t=1}^n \{ \log[f(\varepsilon_t(\rho_0, \gamma_n))] - \log[f(\varepsilon_t(\rho_0, \gamma_0))] \}. \tag{A1}$$

First, we show that as  $n \rightarrow +\infty$ ,  $\Theta_n(\rho_0, \gamma_0, \beta)$  decomposes into

$$\Theta_n(\rho_0, \gamma_0, \beta) = \Pi_n - \Delta_n + o_p(1),$$

where

$$\Delta_n = \frac{1}{2n} \sum_{t=1}^n \left\{ \frac{1}{V^2(\mathbf{X}_{t-1})} \beta^\top M(\gamma_0, \mathbf{X}_{t-1}) \beta \phi'_f[\varepsilon_t(\rho_0, \gamma_0)] - \frac{1}{V(\mathbf{X}_{t-1})} \beta^\top \mathcal{H}(\gamma_0, \mathbf{X}_{t-1}) \beta \phi_f[\varepsilon_t(\rho_0, \gamma_0)] \right\}, \tag{A2}$$

$$\Pi_n = \frac{1}{\sqrt{n}} \sum_{t=1}^n \left\{ \frac{1}{V(\mathbf{X}_{t-1})} \beta^\top N(\gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\rho_0, \gamma_0)] \right\}, \tag{A3}$$

and

- $N(\gamma, \mathbf{X}_{t-1}) = \left( \omega_1(t) \partial_{\gamma_1} [T(\rho_0, \gamma, \mathbf{X}_{t-1})], \dots, \omega_{k+1}(t) \partial_{\gamma_{k+1}} [T(\rho_0, \gamma, \mathbf{X}_{t-1})] \right)^\top \in \mathbb{R}^{p(k+1)}$ , and for  $i = 1, \dots, k+1$ ,  $\omega_i \in \{0, 1\}$ ,
  - $M(\gamma, \mathbf{X}_{t-1}) = \begin{pmatrix} M_1(\gamma, \mathbf{X}_{t-1}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & M_2(\gamma, \mathbf{X}_{t-1}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & M_{k+1}(\gamma, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_{p(k+1)}(\mathbb{R})$ ,
- where  $\mathbf{0} \in \mathcal{M}_p(\mathbb{R})$  is a null matrix and for any  $i = 1, \dots, k+1$ ,

$$M_i(\gamma, \mathbf{X}_{t-1}) = \omega_i^2(t) \begin{pmatrix} \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \gamma, \mathbf{X}_{t-1}) & \dots & \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \gamma, \mathbf{X}_{t-1}) & \dots & \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}),$$

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

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

$$\mathcal{H}_i(\gamma, \mathbf{X}_{t-1}) = \omega_i^2(t) \begin{pmatrix} \frac{\partial^2 T}{\partial \gamma_{i,1}^2}(\rho_0, \gamma, \mathbf{X}_{t-1}) & \dots & \frac{\partial^2 T}{\partial \gamma_{i,p} \partial \gamma_{i,1}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 T}{\partial \gamma_{i,1} \partial \gamma_{i,p}}(\rho_0, \gamma, \mathbf{X}_{t-1}) & \dots & \frac{\partial^2 T}{\partial \gamma_{i,p}^2}(\rho_0, \gamma, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}).$$

Applying a first-order Taylor expansion on  $\log \{f[\varepsilon_t(\rho_0, \gamma)]\}$  in a neighborhood of  $\gamma_0$ , we obtain, for some  $\tilde{\gamma}$  lying between  $\gamma_0$  and  $\gamma_n$ ,

$$\begin{aligned} \log \{f[\varepsilon_t(\rho_0, \gamma_n)]\} - \log \{f[\varepsilon_t(\rho_0, \gamma_0)]\} &= (\gamma_n - \gamma_0)^\top (D_\gamma [\log \{f[\varepsilon_t(\gamma)]\}])|_{\gamma=\gamma_0} \\ &+ \frac{1}{2} (\gamma_n - \gamma_0)^\top H_\gamma [\log \{f[\varepsilon_t(\gamma)]\}]|_{\gamma=\tilde{\gamma}} (\gamma_n - \gamma_0). \end{aligned}$$

To simplify the study, we calculate all the expressions we need.

- $D_\gamma [\varepsilon_t(\rho_0, \gamma)] = -\frac{1}{V(\mathbf{X}_{t-1})} N(\gamma, \mathbf{X}_{t-1}),$
- $H_\gamma [\varepsilon_t(\rho_0, \gamma)] = -\frac{1}{V(\mathbf{X}_{t-1})} H_\gamma [T(\rho_0, \gamma, \mathbf{X}_{t-1})],$
- $D_\gamma [\log \{f[\varepsilon_t(\rho_0, \gamma)]\}] = \frac{1}{V(\mathbf{X}_{t-1})} N(\gamma, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\rho_0, \gamma)]$   
 $= -D_\gamma [\varepsilon_t(\rho_0, \gamma)] \phi_f[\varepsilon_t(\rho_0, \gamma)],$
- $H_\gamma (\log \{f[\varepsilon_t(\rho_0, \gamma)]\}) = -\frac{1}{V^2(\mathbf{X}_{t-1})} M(\gamma, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\rho_0, \gamma)]$   
 $+ \frac{1}{V(\mathbf{X}_{t-1})} H_\gamma [T(\rho_0, \gamma, \mathbf{X}_{t-1})] \phi_f[\varepsilon_t(\rho_0, \gamma)]$   
 $= -\frac{1}{V^2(\mathbf{X}_{t-1})} M(\gamma, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\rho_0, \gamma)]$   
 $- H_\gamma [\varepsilon_t(\rho_0, \gamma)] \phi_f[\varepsilon_t(\rho_0, \gamma)].$

Then,

$$\begin{aligned} \log \{f[\varepsilon_t(\rho_0, \gamma_n)]\} - \log \{f[\varepsilon_t(\rho_0, \gamma_0)]\} &= \frac{1}{\sqrt{n}} \frac{\beta^\top N(\gamma_0, \mathbf{X}_{t-1})}{V(\mathbf{X}_{t-1})} \phi_f[\varepsilon_t(\rho_0, \gamma_0)] \\ &- \frac{1}{2n} \left[ \frac{1}{V^2(\mathbf{X}_{t-1})} \beta^\top M(\tilde{\gamma}, \mathbf{X}_{t-1}) \beta \phi'_f[\varepsilon_t(\rho_0, \tilde{\gamma})] \right] \\ &- \frac{1}{2n} \left[ \frac{1}{V(\mathbf{X}_{t-1})} \beta^\top H_\gamma [T(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1})] \beta \phi_f[\varepsilon_t(\rho_0, \tilde{\gamma})] \right]. \end{aligned}$$

Now,

$$\begin{aligned} \Theta_n(\rho_0, \gamma_0, \beta) &= \frac{1}{\sqrt{n}} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \beta^\top N(\gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\rho_0, \gamma_0)] \\ &- \frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \beta^\top M(\tilde{\gamma}, \mathbf{X}_{t-1}) \beta \phi'_f[\varepsilon_t(\rho_0, \tilde{\gamma})] \\ &+ \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \beta^\top \mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1}) \beta \phi_f[\varepsilon_t(\rho_0, \tilde{\gamma})], \end{aligned}$$

where, observing that for any  $t$  and for any  $i \neq j$ ,  $\omega_i(t)\omega_j(t) = 0$ ,

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

with  $\mathbf{0} \in \mathcal{M}_p(\mathbb{R})$  standing for a null matrix and

$$M_i(\tilde{\gamma}, \mathbf{X}_{t-1}) = \omega_i^2(t) \begin{pmatrix} \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \frac{\partial T}{\partial \gamma_{i,1}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \frac{\partial T}{\partial \gamma_{i,p}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}),$$

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

with

$$\mathcal{H}_i(\tilde{\gamma}, \mathbf{X}_{t-1}) = \omega_i^2(t) \begin{pmatrix} \frac{\partial^2 T}{\partial \gamma_{i,1}^2}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \frac{\partial^2 T}{\partial \gamma_{i,p} \partial \gamma_{i,1}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 T}{\partial \gamma_{i,1} \partial \gamma_{i,p}}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) & \dots & \frac{\partial^2 T}{\partial \gamma_{i,p}^2}(\rho_0, \tilde{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_p(\mathbb{R}).$$

Let

$$\begin{aligned} \chi(\tilde{\gamma}, \mathbf{X}_{t-1}) &= -\frac{1}{2n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top M(\tilde{\gamma}, \mathbf{X}_{t-1}) \boldsymbol{\beta} \phi'_f[\varepsilon_t(\rho_0, \tilde{\gamma})] \\ &\quad + \frac{1}{2n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \boldsymbol{\beta}^\top (\mathcal{H}(\tilde{\gamma}, \mathbf{X}_{t-1})) \boldsymbol{\beta} \phi_f[\varepsilon_t(\rho_0, \tilde{\gamma})]. \end{aligned}$$

Using (A3), we have  $\mathbb{E}\{\phi_f[\varepsilon_t(\rho_0, \gamma)]\} = 0$ ,  $\mathbb{E}\{\phi'_f[\varepsilon_t(\rho_0, \gamma)]\} = I(f)$  and  $\mathbb{E}\{|\phi_f[\varepsilon_t(\rho_0, \gamma)]|\} < \infty$ . Now, using (A4) and the ergodic theorem, from a simple calculation, we can prove that

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

It results from above that

$$\Theta_n(\rho_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n - \Delta_n + o_p(1),$$

where  $\Delta_n$  and  $\Pi_n$  are defined by (A2) and (A3), respectively. Now, we study of the asymptotic behavior of  $\Delta_n$  under  $H_0$ .

By the piecewise stationarity and ergodicity, for any  $j = 1, \dots, k + 1$ , we can write, almost surely,

$$\lim_{n \rightarrow +\infty} \Delta_n = \frac{1}{2} \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) = \frac{\eta(\rho_0, \gamma_0, \beta)}{2},$$

with

$$\eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) = I(f) \int_{\mathbb{R}^d} \frac{1}{V^2(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, \gamma_0, x) \frac{\partial T}{\partial \gamma_{j,m}}(\rho_0, \gamma_0, x) dF_j(x).$$

Thus, we can write

$$\Theta_n = \Pi_n - \frac{\eta(\rho_0, \gamma_0, \beta)}{2} + o_P(1).$$

Now, we prove that under  $H_0$ ,

$$\Pi_n \xrightarrow{\mathcal{D}} \mathcal{N}(0, \eta(\rho_0, \gamma_0, \beta)).$$

We consider the sequence

$$\Theta_{n,j} = \frac{1}{\sqrt{n}} \sum_{t=1}^j \frac{1}{V(\mathbf{X}_{t-1})} \beta^\top N(\gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\rho_0, \gamma_0)], \quad j = 1, \dots, n,$$

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

$$\psi_{n,t} = \frac{1}{\sqrt{n}} \frac{1}{V(\mathbf{X}_{t-1})} \beta^\top N(\gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\rho_0, \gamma_0)].$$

We use Corollary 3.1 of [31] to study the asymptotic behavior of  $\Theta_{n,j}$ .

It is easy to prove that  $\{(\Theta_{n,j}, \mathcal{F}_j), j = 1, \dots, n\}$  is a martingale sequence.

Using the fact that  $\varepsilon_t$  is independent of  $\mathcal{F}_{t-1}$  for  $t = 1, \dots, n$  and using the ergodic theorem, we can show that, almost surely,

$$\lim_{n \rightarrow +\infty} \sum_{t=1}^n \mathbb{E}(\psi_{n,t}^2 | \mathcal{F}_{t-1}) = \sum_{j=1}^{k+1} \alpha_j \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) = \eta(\rho_0, \gamma_0, \beta) < \infty,$$

which shows that the first condition of Corollary 3.1 of [31] is verified. It remains to check the Linderberg condition. Let  $\epsilon > 0$ ; by the Hölder inequality, using Markov inequality and the ergodic theorem, we can write that as  $n \rightarrow \infty$ ,

$$\sum_{t=1}^n \mathbb{E}(\psi_{n,t}^2 \mathbf{1}_{|\psi_{n,t}| > \epsilon} | \mathcal{F}_{t-1}) \xrightarrow{a.s.} 0.$$

Then, the conditions of Corollary 3.1 of [31] are completely verified, so that under  $H_0$ , we have

$$\Pi_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \eta(\rho_0, \gamma_0, \beta)). \tag{A4}$$

Consequently, under  $H_0$ , we have

$$\Theta_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{\eta(\rho_0, \gamma_0, \beta)}{2}, \eta(\rho_0, \gamma_0, \beta)\right). \tag{A5}$$



Collecting the above results, the LAN property is established with the central sequence  $\Pi_n(\rho_0, \gamma_0, \beta)$ .

Appendix A.2. Proof of Corollary 1

For any  $\beta \in \mathbb{R}^{p(k+1)}$ , from Theorem 1, under  $H_0$ , as  $n \rightarrow +\infty$ ,

$$\Pi_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \eta(\rho_0, \gamma_0, \beta)).$$

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

$$\Theta_n(\rho_0, \gamma_0, \beta) \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{\eta(\rho_0, \gamma_0, \beta)}{2}, \eta(\rho_0, \gamma_0, \beta)\right).$$

Then, it is easy to see that under  $H_0$ , as  $n \rightarrow +\infty$ ,

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

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

Since  $\mathbb{E}(\Pi_n) = 0$  and

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

under  $H_0$ , we have

$$\begin{pmatrix} \Pi_n(\rho_0, \gamma_0, \beta) \\ \Theta_n \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix} 0 \\ -\frac{\eta(\rho_0, \gamma_0, \beta)}{2} \end{pmatrix}, \begin{pmatrix} \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \\ \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \end{pmatrix}\right). \tag{A6}$$

Using [32] or [33], 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  $n \rightarrow +\infty$ ,

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

Appendix A.3. Proof of Theorem 2

From Theorem 1 and Corollary 1, we can conclude immediately that, under  $H_0$ , as  $n \rightarrow +\infty$ ,

$$\begin{pmatrix} \Pi_n(\rho_0, \gamma_0, \beta) \\ \Theta_n \end{pmatrix} \xrightarrow{\mathcal{D}} \mathcal{N}\left(\begin{pmatrix} 0 \\ -\frac{\eta(\rho_0, \gamma_0, \beta)}{2} \end{pmatrix}, \begin{pmatrix} \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \\ \eta(\rho_0, \gamma_0, \beta) & \eta(\rho_0, \gamma_0, \beta) \end{pmatrix}\right). \tag{A7}$$

Part [i] is a direct consequence of Theorem 2 and is briefly explained in the proof of Corollary 1.

As explained there, the sequences of hypotheses are contiguous, and under  $H_\beta^{(n)}$ , as  $n \rightarrow +\infty$ , we have

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

By (A6) and the Le Cam's third lemma (Proposition 4.2 in [32]), 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 recall that, under  $H_0$ , as  $n \rightarrow +\infty$ ,

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

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

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  $n \rightarrow +\infty$ ,

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

Indeed, for  $n \geq 1$ , we can write

$$\frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\widehat{\vartheta}_n(\rho_0, \gamma_0, \beta)} = \frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\vartheta(\rho_0, \gamma_0, \beta)} \times \frac{\vartheta(\rho_0, \gamma_0, \beta)}{\widehat{\vartheta}_n(\rho_0, \gamma_0, \beta)}.$$

From which it results that, under  $H_\beta^{(n)}$  and as  $n \rightarrow +\infty$ ,

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

For parts [ii] and [iii], to calculate the asymptotic power of our test statistic, we calculate the asymptotic cumulative distribution of  $\frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\widehat{\vartheta}_n(\rho_0, \gamma_0, \beta)}$  under  $H_\beta^{(n)}$ . We have

$$\begin{aligned} \lim_{n \rightarrow +\infty} P\left(\frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\widehat{\vartheta}_n(\rho_0, \gamma_0, \beta)} > z_\alpha \mid H_\beta^{(n)}\right) &= \lim_{n \rightarrow +\infty} P\left(\frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\vartheta(\rho_0, \gamma_0, \beta)} > z_\alpha \mid H_\beta^{(n)}\right) \\ &= 1 - \Phi(z_\alpha - \vartheta(\rho_0, \gamma_0, \beta)) \\ &= \mathcal{P}_{k, \tau^k}, \end{aligned} \tag{A8}$$

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

By Section 4.4.3 of [33], the test based on  $\mathcal{T}_n(\rho_0, \gamma_0, \beta)$  is locally asymptotically optimal.

#### Appendix A.4. Proof of Proposition 1

##### Appendix A.4.1. Proof of [i]

From the Bahadur representation (9), as in [34,35], we consider the following sequence:

$$R_{n,j} = n^{-\frac{1}{2}} \sum_{t=1}^j Y(\rho_0, \mathbf{X}_{t-1}) \aleph[\varepsilon_t(\rho_0, \gamma_0)], \quad j = 1, \dots, n.$$

and

$$u^\top R_{n,j} = \frac{1}{\sqrt{n}} \sum_{t=1}^j y_t(u) \in \mathbb{R}, \quad \text{where } y_t(u) = u^\top Y(\rho_0, \mathbf{X}_{t-1}) \aleph[\varepsilon_t(\rho_0, \gamma_0)],$$

$$u \in \mathbb{R}^p.$$

It is easy to see that  $u^\top R_{n,j}$  is centered for every  $j = 1, \dots, n$ . Since  $\mathbb{E}\{\aleph[\varepsilon_t(\rho_0, \gamma_0)]\} = 0$  for any  $t \in \mathbb{N}$ , from a simple calculation, we prove that  $\{(u^\top R_{n,j}, \mathcal{F}_j), j = 1, \dots, n\}$  is a martingale sequence.

We check now the first condition of Corollary 3.1 of [31]. Since  $\varepsilon_t$  is independent of  $\mathcal{F}_{t-1}$  for  $t = 1, \dots, n$ , we can write

$$\begin{aligned} & \sum_{t=1}^n \mathbb{E} \left\{ \left[ n^{-\frac{1}{2}} y_t(u) \right]^2 \middle| \mathcal{F}_{t-1} \right\} \\ &= \sum_{t=1}^n n^{-1} \mathbb{E} \left\{ \left[ u^\top Y(\rho_0, \mathbf{X}_{t-1}) \aleph[\varepsilon_t(\rho_0, \gamma_0)] \right]^2 \middle| \mathcal{F}_{t-1} \right\} \\ &= \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left[ u^\top Y(\rho_0, \mathbf{X}_{t-1}) \right]^2 \mathbb{E} \left\{ \aleph^2[\varepsilon_t(\rho_0, \gamma_0)] \right\}. \end{aligned}$$

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

$$\begin{aligned} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \left[ u^\top Y(\rho_0, \mathbf{X}_{t-1}) \right]^2 \mathbb{E} \left\{ \aleph^2[\varepsilon_t(\rho_0, \gamma_0)] \right\} &\xrightarrow{a.s.} \int_{\mathbb{R}^d} \left[ u^\top Y(\rho_0, x) \right]^2 dF_j(x) \\ &\times \int_{\mathbb{R}} \aleph^2(x) f(x) dx < \infty. \end{aligned}$$

Then,

$$\begin{aligned} \sum_{t=1}^n \mathbb{E} \left\{ \left[ \frac{1}{\sqrt{n}} y_t(u) \right]^2 \middle| \mathcal{F}_{t-1} \right\} &\xrightarrow[n \rightarrow +\infty]{a.s.} s = \int_{\mathbb{R}} \aleph^2(x) f(x) dx \\ &\times \sum_{j=1}^{k+1} \alpha_j \int_{\mathbb{R}^d} \left[ u^\top Y(\rho_0, x) \right]^2 dF_j(x). \end{aligned}$$

Finally, we check the Linderberg condition, that is, the second condition of Corollary 3.1 of [31]. In this purpose, we prove that, as  $n \rightarrow +\infty$ ,

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

Let  $\epsilon > 0$ . By the Hölder inequality, and Markov inequalities, we can write

$$\begin{aligned} & \sum_{t=1}^n \mathbb{E} \left\{ \left[ n^{-\frac{1}{2}} y_t(u) \right]^2 \mathbf{1}_{\left| n^{-\frac{1}{2}} y_t(u) \right| > \epsilon} \middle| \mathcal{F}_{t-1} \right\} \\ &\leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ \left[ n^{-\frac{1}{2}} y_t(u) \right]^3 \middle| \mathcal{F}_{t-1} \right\} \mathbb{E}^{\frac{1}{3}} \left\{ \mathbf{1}_{\left| n^{-\frac{1}{2}} y_t(u) \right| > \epsilon} \middle| \mathcal{F}_{t-1} \right\} \\ &\leq \sum_{t=1}^n \mathbb{E}^{\frac{2}{3}} \left\{ \left[ n^{-\frac{1}{2}} y_t(u) \right]^3 \middle| \mathcal{F}_{t-1} \right\} \frac{\mathbb{E}^{\frac{1}{3}} \left\{ \left| n^{-\frac{1}{2}} y_t(u) \right|^3 > \epsilon \middle| \mathcal{F}_{t-1} \right\}}{\sqrt[3]{\epsilon}} \\ &\leq \frac{1}{\sqrt[3]{\epsilon} \sqrt{n}} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \|u^\top Y(\rho_0, \mathbf{X}_{t-1})\|^3 \mathbb{E} \left\{ |\aleph[\varepsilon_t(\rho_0, \gamma_0)]|^3 \right\}. \end{aligned}$$

By the piece-wise stationarity and the ergodic theorem, for  $j = 1, \dots, k + 1$ , we obtain almost surely that

$$\lim_{n \rightarrow +\infty} \frac{1}{\sqrt{n}} \sum_{j=1}^{k+1} \frac{n_j(n)}{n} \left( \frac{1}{n_j(n)} \sum_{t=\tau_{j-1}}^{\tau_j} \|u^\top Y(\rho_0, \mathbf{X}_{t-1})\|^3 \mathbb{E} \left\{ |\aleph[\varepsilon_t(\rho_0, \gamma_0)]|^3 \right\} \right) = 0$$

Then, using Corollary 3.1 of [31], we conclude that, under  $H_0$ , we have

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

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  $\Sigma$  is the covariance matrix defined as

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

#### Appendix A.4.2. 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

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

with

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

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

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

Since  $\mathbb{E}\{\phi_f[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} = 0$ ,  $\lim_{n \rightarrow +\infty} \mathbb{E}(Q_n) = 0$ ,  $\varepsilon_t$  is independent of  $\mathcal{F}_{t-1}$ ,  $\mathbb{E}\{\aleph[\varepsilon_t(\boldsymbol{\rho}_0, \boldsymbol{\gamma}_0)]\} = 0$ , and using the stationarity and the ergodic theorem, we can easily see that

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

where

$$C = \int_{\mathbb{R}} \aleph(x)\phi_f(x)f(x)dx \sum_{j=1}^n \alpha_j \omega_j \sum_{h=1}^p \beta_{j,h} \int_{\mathbb{R}^d} \frac{Y(\rho_0, x)}{V(x)} \frac{\partial T}{\partial \gamma_{j,h}}(\rho_0, x) dF_j(x).$$

Then, under  $H_0$ , we have

$$\left( \begin{matrix} Q_n \\ \Theta_n(\rho_0, \gamma_0, \beta) \end{matrix} \right) \xrightarrow{\mathcal{D}} \mathcal{N} \left( \left( \begin{matrix} 0 \\ -\frac{\eta(\rho_0, \gamma_0, \beta)}{2} \end{matrix} \right), \left( \begin{matrix} \Sigma & C^\top \\ C & \eta(\rho_0, \gamma_0, \beta) \end{matrix} \right) \right).$$

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

$$\sqrt{n}(\rho_n - \rho_0) = Q_n \xrightarrow{\mathcal{D}} \mathcal{N}(C, \Sigma).$$

Appendix A.5. Proof of Proposition 2

Appendix A.5.1. Proof of [i]

We prove the convergence of the central sequence (4) to its estimated version in order to verify that the test still be optimal when we replace the parameter by its estimator. For any  $\rho \in \mathbb{R}^p$  and  $\gamma, \beta \in \mathbb{R}^{p(k+1)}$ , we define

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

Then the log-likelihood ratio of  $H_0$  against  $H_\beta^{(n)}$  is  $\Theta(\rho_0, \gamma_0, \beta)$ . For  $\tilde{\rho}_n$  lying between  $\rho_n$  and  $\rho_0$ , we write a second-order Taylor expansion of  $\Pi_n(\rho_0, \gamma_0, \beta)$  around  $\rho_n$  and obtain

$$\begin{aligned} \Pi_n(\rho_0, \gamma_0, \beta) &= \Pi_n(\rho_n, \gamma_0, \beta) + (\rho_0 - \rho_n)^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) \\ &\quad + \frac{1}{2} (\rho_0 - \rho_n)^\top \partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta) (\rho_0 - \rho_n). \end{aligned} \tag{A9}$$

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

$$(\rho_0 - \rho_n)^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) = o_P(1) \tag{A10}$$

$$\frac{1}{2} (\rho_0 - \rho_n)^\top \partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta) (\rho_0 - \rho_n) = o_P(1). \tag{A11}$$

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

$$\begin{aligned} &\frac{1}{\sqrt{n}} \partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta) \\ &= \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\rho^2 (T^{\gamma_{m,h}})(\tilde{\rho}_n, \gamma_0, \beta) \phi_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)] \\ &\quad - \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\rho (T^{\gamma_{m,h}})(\tilde{\rho}_n, \gamma_0, \beta) \partial_\rho^\top T(\tilde{\rho}_n, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)] \\ &\quad - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\tilde{\rho}_n, \gamma_0, \beta) \partial_\rho^2 T(\tilde{\rho}_n, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)] \\ &\quad + \frac{1}{n} \sum_{t=1}^n \frac{1}{V^3(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\tilde{\rho}_n, \gamma_0, \beta) \partial_\rho T(\tilde{\rho}_n, \mathbf{X}_{t-1}) \partial_\rho^\top T(\tilde{\rho}_n, \mathbf{X}_{t-1}) \\ &\quad \times \phi''_f[\varepsilon_t(\tilde{\rho}_n, \gamma_0)] \\ &= \Delta_{1,n}(\tilde{\rho}_n, \gamma_0, \beta) + \Delta_{2,n}(\tilde{\rho}_n, \gamma_0, \beta) + \Delta_{3,n}(\tilde{\rho}_n, \gamma_0, \beta) + \Delta_{4,n}(\tilde{\rho}_n, \gamma_0, \beta). \end{aligned}$$

By a multiple use of a Taylor expansion, by the assumptions  $(B_4)$ ,  $(A_9)$ , by the ergodic theorem, from a simple calculation, we prove that, for  $i = 1, \dots, 4$ ,  $\|\Delta_{i,n}(\tilde{\rho}_n, \gamma_0, \beta)\|_p$  is bounded.

Thus, 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 (A11) that, as  $n \rightarrow +\infty$ , we have

$$\begin{aligned} & \frac{1}{2} \left| (\rho_0 - \rho_n)^\top \partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta) (\rho_0 - \rho_n) \right| \\ & \leq \frac{1}{2} \|\sqrt{n}(\rho_0 - \rho_n)\|_p \left\| \frac{1}{\sqrt{n}} \partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta) \right\|_p \|\rho_0 - \rho_n\|_p \\ & \leq \frac{1}{2} \|\sqrt{n}(\rho_0 - \rho_n)\|_p \|\rho_0 - \rho_n\|_p \times b. \end{aligned}$$

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

$$\sqrt{n}(\rho_0 - \rho_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \Sigma),$$

it results that

$$\frac{1}{2} (\rho_0 - \rho_n)^\top \partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta) (\rho_0 - \rho_n) = o_P(1). \tag{A12}$$

From (A12), we can write

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_n, \gamma_0, \beta) + (\rho_0 - \rho_n)^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) + o_P(1). \tag{A13}$$

Now, we prove that

$$(\rho_0 - \rho_n)^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) = o_P(1).$$

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

$$\begin{aligned} & \Pi_n(\rho_0, \gamma_0, \beta) \\ & = \Pi_n(\rho_n, \gamma_0, \beta) + (\rho_0 - \rho_n + \rho_{s(n)} - \rho_{s(n)})^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) + o_P(1) \\ & = \Pi_n(\rho_n, \gamma_0, \beta) + (\rho_0 - \rho_{s(n)})^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) + (\rho_{s(n)} - \rho_n)^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) \\ & \quad + o_P(1), \end{aligned}$$

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  $n \rightarrow +\infty$ ,

$$\sqrt{n}(\rho_0 - \rho_{s(n)})^\top = \sqrt{s(n)}(\rho_0 - \rho_{s(n)})^\top \times \sqrt{\frac{n}{s(n)}} = o_P(1),$$

it is easy to see that,

$$(\rho_0 - \rho_{s(n)})^\top \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) = \sqrt{n}(\rho_0 - \rho_{s(n)})^\top \frac{1}{\sqrt{n}} \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta).$$

Then, it suffices to show that  $\partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) / \sqrt{n}$  converges in probability to a random vector. For this, we can write the following decomposition:

$$\begin{aligned} & \frac{1}{\sqrt{n}} \partial_\rho \Pi_n(\rho_n, \gamma_0, \beta) \\ & = \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\rho (T^{\gamma_{m,h}})(\rho_n, \gamma_0, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\rho_n, \gamma_0)] \end{aligned}$$

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

Using the assumptions (B<sub>4</sub>) and (A<sub>5</sub>), and the ergodic theorem, since  $\varepsilon_t$  is independent of  $\mathcal{F}_{t-1}$ , the study of the asymptotic behavior of  $\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})$  and  $\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})$  shows that, as  $n \rightarrow +\infty$ ,

$$\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) \xrightarrow{\mathbb{P}} \mathbf{0} \in \mathbb{R}^p.$$

Thus,  $\|\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})\|_p$  tends in probability to a finite positive real number. Consequently,

$$\left\| \frac{1}{\sqrt{n}} \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) \right\|_p \leq \|\mathcal{U}_{1,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})\|_p + \|\mathcal{U}_{2,n}(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta})\|_p < \infty.$$

It results that

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

Hence,

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + (\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + o_P(1). \tag{A14}$$

In order to treat the above equation (A14), we need the following lemma.

**Lemma A1.** Assume that (B<sub>2</sub>) holds. 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  $\gamma_0, \boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$ ,  $\boldsymbol{\rho}_{s(n)}$  is asymptotically in the tangent space  $\mathbb{T}_n$  to the curve of  $\Pi_n(\boldsymbol{\rho}, \gamma, \boldsymbol{\beta})$  at  $\boldsymbol{\rho}_n$ , defined as follows:

$$\mathbb{T}_n = \left\{ z \in \mathbb{R}^p / \Pi_n(z, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + (z - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) \right\}.$$

**Proof.** Writing a second-order Taylor expansion of  $\Pi_n(\boldsymbol{\rho}_{s(n)}, \gamma_0, \boldsymbol{\beta})$  in a neighborhood of  $\boldsymbol{\rho}_n$ , for some  $\tilde{\boldsymbol{\rho}}_{s(n)}$  lying between  $\boldsymbol{\rho}_{s(n)}$  and  $\boldsymbol{\rho}_n$ , we obtain

$$\begin{aligned}
 \Pi_n(\boldsymbol{\rho}_{s(n)}, \gamma_0, \boldsymbol{\beta}) & = \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) + (\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}} \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) \\
 & \quad + \frac{1}{2} (\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_{s(n)}, \gamma_0, \boldsymbol{\beta}) (\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n).
 \end{aligned}$$

To prove that, as  $n \rightarrow +\infty$ ,  $\boldsymbol{\rho}_{s(n)}$  belongs to  $\mathbb{T}_n$ , it suffices to show that  $(\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)^\top \partial_{\boldsymbol{\rho}}^2 \Pi_n(\tilde{\boldsymbol{\rho}}_{s(n)}, \gamma_0, \boldsymbol{\beta}) (\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n) = o_P(1)$ .

To find the asymptotic distribution of  $\sqrt{n}(\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)$ , we add and subtract appropriate terms and we obtain

$$\begin{aligned}
 \sqrt{n}(\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n) & = \sqrt{s(n)}(\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_0) \sqrt{\frac{n}{s(n)}} + \sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n) \\
 & = o_P(1) + \sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n).
 \end{aligned}$$

Then, asymptotically,  $\sqrt{n}(\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)$  has the same distribution as  $\sqrt{n}(\boldsymbol{\rho}_0 - \boldsymbol{\rho}_n)$ . This means that  $\sqrt{n}(\boldsymbol{\rho}_{s(n)} - \boldsymbol{\rho}_n)$  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 the sequence  $\partial_\rho^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta) / \sqrt{n}$  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  $n \rightarrow +\infty$ ,  $\rho_n$  is given by  $(\mathcal{B}_2)$ , and  $\tilde{\rho}_{s(n)}$  lies between  $\rho_{s(n)}$  and  $\rho_0$ .

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  $\rho_n$  and  $\rho_0$ , we proved previously that  $\|\partial_\rho^2 \Pi_n(\tilde{\rho}_n, \gamma_0, \beta)\|_p / \sqrt{n}$  converges in probability, as  $n \rightarrow +\infty$ , to a finite positive number. By following the same lines, we can prove that  $\|\partial_\rho^2 \Pi_n(\tilde{\rho}_{s(n)}, \gamma_0, \beta)\|_p / \sqrt{n}$  converges in probability to a positive finite number, where  $\tilde{\rho}_{s(n)}$  lies between  $\rho_{s(n)}$  and  $\rho_0$ . 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 A1 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 (A14), we obtain

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_{s(n)}, \gamma_0, \beta) + o_P(1).$$

### Appendix A.5.2. Proof of [ii]

To prove (12), it suffices to show that, as  $n \rightarrow +\infty$ ,  $\hat{\eta}_n(\hat{\rho}_n, \gamma_0, \beta) \rightarrow \eta(\rho_0, \gamma_0, \beta)$ . For any  $\beta \in \mathbb{R}^{p(k+1)}$ , we have

$$\begin{aligned} & \hat{\eta}_n(\hat{\rho}_n, \gamma_0, \beta) - \eta(\rho_0, \gamma_0, \beta) \\ &= \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \left[ \hat{\alpha}_j \hat{\eta}_{j,2}^{(h,m)}(\hat{\rho}_n, \gamma_0) - \alpha_j \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right]. \end{aligned}$$

We add and subtract appropriate terms and we obtain

$$\begin{aligned} & \hat{\eta}_n(\hat{\rho}_n, \gamma_0, \beta) - \eta(\rho_0, \gamma_0, \beta) \\ &= \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} (\hat{\alpha}_j - \alpha_j) [\hat{\eta}_{j,2}^{(h,m)}(\hat{\rho}_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) + \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0)] \\ &+ \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \alpha_j [\hat{\eta}_{j,2}^{(h,m)}(\hat{\rho}_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0)] \tag{A15} \\ &= \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \left[ (\hat{\alpha}_j - \alpha_j) [\hat{\eta}_{j,2}^{(h,m)}(\hat{\rho}_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0)] \right. \\ &+ \left. (\hat{\alpha}_j - \alpha_j) \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0) \right] + \sum_{j=1}^{k+1} \sum_{1 \leq h \leq m \leq p} \beta_{j,h} \beta_{j,m} \alpha_j [\hat{\eta}_{j,2}^{(h,m)}(\hat{\rho}_n, \gamma_0) - \eta_{j,2}^{(h,m)}(\rho_0, \gamma_0)], \end{aligned}$$



where, for  $h = 1, \dots, p$ ,

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

For all  $j = 1, \dots, k + 1$ , we have

$$\hat{\alpha}_j - \alpha_j \xrightarrow{n \rightarrow +\infty} 0.$$

For  $j = 1, \dots, k + 1$  and  $h = 1, \dots, p$ , using the assumption (A9) 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 (A15) tends to 0.

Appendix A.6. Proof of Theorem 3

Appendix A.6.1. Proof of [i]

The statistic that we study is

$$\mathcal{T}_n(\rho_{s(n)}, \gamma_0, \beta) = \frac{\Pi_n(\rho_{s(n)}, \gamma_0, \beta)}{\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta)}.$$

We proved in Proposition (2) that

$$\Pi_n(\rho_{s(n)}, \gamma_0, \beta) = \Pi_n(\rho_0, \gamma_0, \beta) + o_P(1).$$

Also, we proved that, as  $n \rightarrow +\infty$ ,

$$\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta) \xrightarrow{\mathbb{P}} \vartheta(\rho_0, \gamma_0, \beta).$$

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

$$\lim_{n \rightarrow +\infty} \frac{\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta)}{\vartheta(\rho_0, \gamma_0, \beta)} = 1,$$

and

$$\begin{aligned} \mathcal{T}_n(\rho_{s(n)}, \gamma_0, \beta) &= \frac{\Pi_n(\rho_0, \gamma_0, \beta) + o_P(1)}{\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta)} \\ &= \frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\vartheta(\rho_0, \gamma_0, \beta)} \times \frac{\vartheta(\rho_0, \gamma_0, \beta)}{\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta)} + \frac{1}{\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta)} \times o_P(1) \\ &\simeq \mathcal{T}_n(\rho_0, \gamma_0, \beta). \end{aligned}$$

Appendix A.6.2. Proof of [ii] and [iii]

The convergence of  $\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta)$  to  $\vartheta(\rho_0, \gamma_0, \beta)$  remains true under the local alternatives  $H_\beta^{(n)}$  by contiguity. Then, under  $H_\beta^{(n)}$  and as  $n \rightarrow +\infty$ , using Le Cam’s third lemma, we have

$$\mathcal{T}_n(\rho_{s(n)}, \gamma_0, \beta) = \frac{\Pi_n(\rho_{s(n)}, \gamma_0, \beta)}{\hat{\vartheta}_n(\rho_{s(n)}, \gamma_0, \beta)} \xrightarrow{\mathcal{D}} \mathcal{N}(\vartheta(\rho_0, \gamma_0, \beta), 1),$$

which shows that the power of the test remains the same.

Appendix A.7. Proof of Proposition 3

The proof relies on a number of lemmas that we state and prove.

**Lemma A2.** Assume that  $(A_1)$ – $(A_{10})$ ,  $(B_1)$ – $(B_5)$ , and  $(B'_1)$ – $(B'_3)$  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  $\{\hat{\gamma}_{0,n}\}_{n \geq 1}$  of  $\gamma_0$  and for any  $\beta \in \mathbb{R}^{p(k+1)}$ , under  $H_0$  and as  $n \rightarrow +\infty$ ,

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_0, \hat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

**Proof.** For any  $\rho \in \mathbb{R}^p$  and  $(\gamma, \beta) \in \mathbb{R}^{p(k+1)} \times \mathbb{R}^{p(k+1)}$ , for  $\hat{\gamma}_{0,n}$  the maximum likelihood estimator of  $\gamma_0$ , we write a first-order Taylor expansion of  $\Pi_n(\rho_0, \gamma_0, \beta)$  in a neighborhood of  $\hat{\gamma}_{0,n}$  and we obtain, for some  $\tilde{\gamma}_{0,n}$  lying between  $\gamma_0$  and  $\hat{\gamma}_{0,n}$ ,

$$\begin{aligned} \Pi_n(\rho_0, \gamma_0, \beta) &= \Pi_n(\rho_0, \hat{\gamma}_{0,n}, \beta) - (\hat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Pi_n(\rho_0, \hat{\gamma}_{0,n}, \beta) \\ &\quad + (\hat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) (\hat{\gamma}_{0,n} - \gamma_0), \end{aligned}$$

where

$$\partial_\gamma^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) = \begin{pmatrix} \partial_{\gamma_1}^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) & \dots & \partial_{\gamma_1 \gamma_p}^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) \\ \vdots & \ddots & \vdots \\ \partial_{\gamma_p \gamma_1}^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) & \dots & \partial_{\gamma_p}^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) \end{pmatrix} \in \mathcal{M}_{p(k+1)}.$$

Our aim is to prove that, under  $H_0$ , as  $n \rightarrow +\infty$ ,

$$(\hat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Pi_n(\rho_0, \hat{\gamma}_{0,n}, \beta) = o_P(1), \tag{A16}$$

$$(\hat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) (\hat{\gamma}_{0,n} - \gamma_0) = o_P(1). \tag{A17}$$

Starting with (A17), by multiplying and dividing by  $\sqrt{n}$ , we observe that

$$\begin{aligned} & \left| (\hat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) (\hat{\gamma}_{0,n} - \gamma_0) \right| \\ & \leq \|\sqrt{n}(\hat{\gamma}_{0,n} - \gamma_0)\|_{p(k+1)} \times \frac{1}{\sqrt{n}} \left\| \left\| \partial_\gamma^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) \right\| \right\|_{p(k+1)} \times \|\hat{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}. \end{aligned}$$

For  $\hat{\gamma}_{0,n}$ , an estimator of  $\gamma_0$ , by following the same techniques as in Proposition 1, under  $H_0$ , as  $n \rightarrow +\infty$ ,  $\sqrt{n}(\hat{\gamma}_{0,n} - \gamma_0)$  converges in distribution to a normal distribution and  $\hat{\gamma}_{0,n} - \gamma_0$  tends to 0 in probability as  $n \rightarrow +\infty$ .

Then, to prove (A17), it suffices to show that  $\left\| \left\| \partial_\gamma^2 \Pi_n(\rho_0, \tilde{\gamma}_{0,n}, \beta) \right\| \right\|_{p(k+1)} / \sqrt{n}$  tends in probability, as  $n \rightarrow +\infty$ , to some positive random variable. Recalling (4), we have

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

Then,

$$\begin{aligned} & \partial_\gamma T^{\gamma_{m,h}}(\rho_0, \gamma, \mathbf{X}_{t-1}) \\ & = (\omega_1 \partial_{\gamma_1} T^{\gamma_{m,h}}(\rho_0, \gamma, \mathbf{X}_{t-1}), \dots, \omega_{k+1} \partial_{\gamma_{k+1}} T^{\gamma_{m,h}}(\rho_0, \gamma, \mathbf{X}_{t-1}))^\top \in \mathbb{R}^{p(k+1)}, \end{aligned}$$

and

$$\partial_\gamma^2 T^{\gamma_{m,h}}(\rho_0, \gamma, \mathbf{X}_{t-1}) =$$

$$\begin{pmatrix} \omega_1^2(t)H_{\gamma_1}T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \omega_2^2(t)H_{\gamma_2}T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \omega_{k+1}^2(t)H_{\gamma_{k+1}}T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \boldsymbol{\gamma}, \mathbf{X}_{t-1}) \end{pmatrix} \in \mathcal{M}_{p(k+1)}(\mathbb{R}),$$

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  $\gamma_i$ .

Recall that we wish to bound  $\frac{1}{\sqrt{n}}\|\|\partial_\gamma^2\Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)}$ . For any  $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$ , we have

$$\begin{aligned} & \frac{1}{\sqrt{n}}\partial_\gamma^2\Pi_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_\gamma^2T^{\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})] \\ & - \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}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}))(\partial_\gamma T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}))^\top \\ & \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\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, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}) \\ & \times \partial_\gamma^2T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1})\phi'_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\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} \\ & \times T^{\gamma_{m,h}}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1})(\partial_\gamma T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}))(\partial_\gamma T(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \mathbf{X}_{t-1}))^\top \\ & \times \phi''_f[\varepsilon_t(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n})] \\ & = \chi_{1,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) - \chi_{2,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) - \chi_{3,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}) + \chi_{4,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta}). \end{aligned}$$

By assumptions  $(B'_1)$ ,  $(A_9)$ ,  $(A_4)$  and  $(A_5)$ , using multiple Taylor expansion and the ergodic theorem, since  $\varepsilon_t$  is independent of  $\mathcal{F}_{t-1}$ , the study of the asymptotic behavior of  $\chi_{i,n}$ ,  $i = 1, \dots, 4$ , shows that

$$\|\|\chi_{i,n}(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)} < \infty.$$

Then, for  $i = 1, \dots, 4$ , we proved that  $\|\|\chi_{i,n}(\tilde{\boldsymbol{\rho}}_n, \boldsymbol{\gamma}_0, \boldsymbol{\beta})\|\|_{p(k+1)}$  converges to a finite positive number. From this, we find that  $\|\|\partial_\gamma^2\Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)}/\sqrt{n}$  converges to a finite positive number.

For proving (A17), 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 1, we can see that, under  $H_0$ ,  $\|\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0\|_{p(k+1)}/\sqrt{n}$  converges to a finite positive number. Since  $\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0$  tends to  $\mathbf{0} \in \mathbb{R}^{p(k+1)}$  in probability, and  $\|\|\partial_\gamma^2\Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})\|\|_{p(k+1)}/\sqrt{n}$  converges under  $H_0$  to some finite positive number, it follows that

$$(\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0)^\top\partial_\gamma^2\Pi_n(\boldsymbol{\rho}_0, \tilde{\boldsymbol{\gamma}}_{0,n}, \boldsymbol{\beta})(\hat{\boldsymbol{\gamma}}_{0,n} - \boldsymbol{\gamma}_0) = o_p(1). \tag{A18}$$

Now, we prove that

$$(\gamma_0 - \hat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) = o_P(1).$$

By adding and subtracting appropriate terms, we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\gamma_0 - \hat{\gamma}_{0,s(n)}) \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) \\ &\quad + (\hat{\gamma}_{0,s(n)} - \hat{\gamma}_{0,n}) \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1), \end{aligned}$$

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  $n \rightarrow +\infty$ ,

$$\sqrt{n}(\gamma_0 - \hat{\gamma}_{0,s(n)})^\top = \sqrt{s(n)}(\gamma_0 - \hat{\gamma}_{0,s(n)})^\top \times \sqrt{\frac{n}{s(n)}} = o_P(1), \tag{A19}$$

it is easy to see that,

$$(\gamma_0 - \hat{\gamma}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) = \sqrt{n}(\gamma_0 - \hat{\gamma}_{0,s(n)})^\top \frac{1}{\sqrt{n}} \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}).$$

By assumptions  $(\mathcal{B}'_1)$ ,  $(A_9)$ ,  $(\mathcal{B}_5)$  and  $(A_5)$ , using a suitable application of the ergodic theorem, since  $\varepsilon_t$  is independent of  $\mathcal{F}_{t-1}$ , we can show that

$$(\gamma_0 - \hat{\gamma}_{0,s(n)})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) = o_P(1).$$

Thus,

$$\Pi_n(\boldsymbol{\rho}_0, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\hat{\gamma}_{0,s(n)} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1). \tag{A20}$$

□

To treat the above equation, we need the following lemma.

**Lemma A3.** Let  $\hat{\gamma}_{0,n}$  be a consistent and asymptotically normal estimator of  $\gamma_0 \in \mathbb{R}^{p(k+1)}$ . 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  $\boldsymbol{\beta} \in \mathbb{R}^{p(k+1)}$ ,  $\hat{\gamma}_{0,s(n)}$  is asymptotically in the tangent space  $\tilde{\mathbb{T}}_n$  to the curve  $\Pi_n(\boldsymbol{\rho}, \gamma, \boldsymbol{\beta})$  at  $\hat{\gamma}_{0,n}$ , defined as follows:

$$\tilde{\mathbb{T}}_n = \left\{ \boldsymbol{y} \in \mathbb{R}^{p(k+1)} \mid \Pi_n(\boldsymbol{\rho}_0, \boldsymbol{y}, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\boldsymbol{y} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) \right\}.$$

**Proof.** Writing a second-order Taylor expansion of  $\Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta})$  in a neighborhood of  $\hat{\gamma}_{0,n}$ , for some  $\tilde{\gamma}_{0,s(n)}$  lying between  $\hat{\gamma}_{0,s(n)}$  and  $\hat{\gamma}_{0,n}$ , we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) &= \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\hat{\gamma}_{0,s(n)} - \hat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_0, \hat{\gamma}_{0,n}, \boldsymbol{\beta}) \\ &\quad + \frac{1}{2} (\hat{\gamma}_{0,s(n)} - \hat{\gamma}_{0,n}) \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) (\hat{\gamma}_{0,s(n)} - \hat{\gamma}_{0,n}). \end{aligned}$$

To prove that, as  $n \rightarrow +\infty$ ,  $\hat{\gamma}_{0,n}$  belongs to  $\tilde{\mathbb{T}}_n$ , it suffices to show that

$$(\hat{\gamma}_{0,s(n)} - \hat{\gamma}_{0,n})^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_0, \tilde{\gamma}_{0,s(n)}, \boldsymbol{\beta}) (\hat{\gamma}_{0,s(n)} - \hat{\gamma}_{0,n}) = o_P(1).$$

Then, we study the asymptotic distribution of  $\sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})$ . By adding and subtracting appropriate terms, we obtain

$$\begin{aligned} \sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) &= \sqrt{s(n)}(\widehat{\gamma}_{0,s(n)} - \gamma_0) \sqrt{\frac{n}{s(n)}} + \sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n}) \\ &= o_P(1) + \sqrt{n}(\gamma_0 - \widehat{\gamma}_{0,n}). \end{aligned}$$

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,  $\sqrt{n}(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})$  converges in distribution to a normal random vector.

To prove that  $(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta) (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) = o_P(1)$ , it suffices to show that  $\partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta) / \sqrt{n}$  converges in probability to a random vector.

Now, we write

$$\begin{aligned} \frac{1}{\sqrt{n}} \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta) &= \chi_{1,n}(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta) - \chi_{2,n}(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta) \\ &\quad - \chi_{3,n}(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta) + \chi_{4,n}(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta), \end{aligned}$$

where  $\{s(n)\}_{n \geq 1}$  stands for a sequence of positive integers satisfying  $n/s(n) \rightarrow 0$  as  $n \rightarrow +\infty$ ,  $\widehat{\gamma}_{0,n}$  is an asymptotically normal estimator of  $\gamma_0$  and  $\widetilde{\gamma}_{0,s(n)}$  lies between  $\widehat{\gamma}_{0,s(n)}$  and  $\gamma_0$ .

Previously, we proved that  $\|\partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,n}, \beta)\|_{p(k+1)} / \sqrt{n}$  converges in probability to a positive random variable where  $\widetilde{\gamma}_{0,n}$  lies between  $\widehat{\gamma}_{0,n}$  and  $\gamma_0$ . Following the same techniques, we can prove that the sequence  $\|\partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta)\|_{p(k+1)} / \sqrt{n}$  converges in probability to a positive random variable, where  $\widetilde{\gamma}_{0,s(n)}$  lies between  $\widehat{\gamma}_{0,s(n)}$  and  $\gamma_0$ .

It results that

$$(\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \partial_\gamma^2 \Pi_n(\rho_0, \widetilde{\gamma}_{0,s(n)}, \beta) (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n}) = o_P(1).$$

It follows from Lemma A3 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

$$\Pi_n(\rho_0, \widehat{\gamma}_{0,s(n)}, \beta) = \Pi_n(\rho_0, \widehat{\gamma}_{0,n}, \beta) + (\widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\rho_0, \widehat{\gamma}_{0,n}, \beta) + o_P(1)$$

Recalling (A20), we obtain

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_0, \widehat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

□

Now we need the following lemma.

**Lemma A4.** Assume that  $(A_1)–(A_{10})$ ,  $(B_1)–(B_5)$  and  $(B'_1)–(B'_3)$  hold. Let  $\{\widehat{\gamma}_{0,n}\}_{n \geq 1}$  be a sequence of consistent and asymptotically normal estimators of  $\gamma_0$ . Let  $s(n)$  be any sequence of positive integers such that  $n/s(n) \rightarrow 0$  as  $n \rightarrow +\infty$ . Then, for any  $\beta \in \mathbb{R}^{p(k+1)}$ , under  $H_0$ , as  $n \rightarrow +\infty$ , we have

$$\Pi_n(\rho_n, \gamma_0, \beta) = \Pi_n(\rho_n, \widehat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

**Proof.** Following the same techniques as above and by applying Taylor expansion of  $\Pi_n(\rho_n, \gamma_0, \beta)$  in a neighborhood of  $\widehat{\gamma}_{0,n}$ , for some  $\widetilde{\gamma}_{0,n}$  lying between  $\gamma_0$  and  $\widehat{\gamma}_{0,n}$ , we obtain

$$\Pi_n(\rho_n, \gamma_0, \beta) = \Pi_n(\rho_n, \widehat{\gamma}_{0,n}, \beta) - (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma \Pi_n(\rho_n, \widehat{\gamma}_{0,n}, \beta)$$

$$+ \frac{1}{2}(\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta})(\widehat{\gamma}_{0,n} - \gamma_0).$$

We have

$$\begin{aligned} & \left| (\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta})(\widehat{\gamma}_{0,n} - \gamma_0) \right| \\ & \leq \|\sqrt{n}(\widehat{\gamma}_{0,n} - \gamma_0)\|_{p(k+1)} \times \frac{1}{\sqrt{n}} \left\| \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)} \|\widehat{\gamma}_{0,n} - \gamma_0\|_{p(k+1)}. \end{aligned}$$

Recall that, under  $H_0$ ,  $\sqrt{n}(\widehat{\gamma}_{0,n} - \gamma_0)$  converges to a finite Gaussian random vector and, almost surely, as  $n \rightarrow +\infty$ ,  $\widetilde{\gamma}_{0,n} - \gamma_0$  tends to  $\mathbf{0} \in \mathbb{R}^{p(k+1)}$ . We study the convergence of  $\partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta}) / \sqrt{n}$ .

Based on the proof of Lemma A2, we have

$$\begin{aligned} & \frac{1}{\sqrt{n}} \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta}) \\ & = \frac{1}{n} \sum_{t=1}^n \frac{1}{V(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \partial_\gamma^2 T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \mathbf{X}_{t-1}) \phi_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n})] \\ & - \frac{2}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} (\partial_\gamma T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \mathbf{X}_{t-1})) (\partial_\gamma T(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}))^\top \\ & \times \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n})] - \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \mathbf{X}_{t-1}) \\ & \times \partial_\gamma^2 T(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \mathbf{X}_{t-1}) \phi'_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n})] + \frac{1}{n} \sum_{t=1}^n \frac{1}{V^2(\mathbf{X}_{t-1})} \sum_{m=1}^{k+1} \sum_{h=1}^p \beta_{m,h} \\ & \times T^{\gamma_{m,h}}(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \mathbf{X}_{t-1}) \left( \partial_\gamma T(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \mathbf{X}_{t-1}) \right) \left( \partial_\gamma T(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \mathbf{X}_{t-1}) \right)^\top \\ & \times \phi''_f[\varepsilon_t(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n})]. \end{aligned}$$

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  $(1/\sqrt{n}) \times \left\| \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)}$  converges to a finite positive number. Thus,  $(1/\sqrt{n}) \times \left\| \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta}) \right\|_{p(k+1)}$  converges almost surely to a finite random positive number.

From these results, we can conclude that

$$(\widehat{\gamma}_{0,n} - \gamma_0)^\top \partial_\gamma^2 \Pi_n(\boldsymbol{\rho}_n, \widetilde{\gamma}_{0,n}, \boldsymbol{\beta})(\widehat{\gamma}_{0,n} - \gamma_0) = o_P(1).$$

Then, we can write

$$\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + (\gamma_0 - \widehat{\gamma}_{0,n})^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1).$$

Adding and subtracting appropriate terms, we obtain

$$\begin{aligned} \Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) & = \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + \left( \gamma_0 - \widehat{\gamma}_{0,s(n)} \right)^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) \\ & + \left( \widehat{\gamma}_{0,s(n)} - \widehat{\gamma}_{0,n} \right)^\top \partial_\gamma \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,n}, \boldsymbol{\beta}) + o_P(1). \end{aligned}$$

By the assumptions  $(\mathcal{B}_1)–(\mathcal{B}_5)$  and by the ergodic theorem, we can prove that

$$\Pi_n(\boldsymbol{\rho}_n, \gamma_0, \boldsymbol{\beta}) = \Pi_n(\boldsymbol{\rho}_n, \widehat{\gamma}_{0,s(n)}, \boldsymbol{\beta}) + o_P(1),$$

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 3, and using Lemma A2, we have

$$\Pi_n(\rho_n, \gamma_0, \beta) = \Pi_n(\rho_n, \widehat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

We can write

$$\Pi_n(\rho_{s(n)}, \gamma_0, \beta) = \Pi_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

From Proposition 2, we have

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_{s(n)}, \gamma_0, \beta) + o_P(1),$$

and

$$\Pi_n(\rho_0, \gamma_0, \beta) - \Pi_n(\rho_{s(n)}, \gamma_{s(n)}, \beta) = \Pi_n(\rho_0, \gamma_0, \beta) - \Pi_n(\rho_{s(n)}, \gamma_0, \beta) + o_P(1).$$

Finally, using the above results, we can write

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta) + o_P(1)$$

□

#### Appendix A.8. Proof of Theorem 4

##### Appendix A.8.1. Proof of [i]

The test statistic is

$$\mathcal{T}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta) = \frac{\Pi_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta)}{\widehat{\vartheta}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta)}.$$

We proved in Proposition 3 that

$$\Pi_n(\rho_0, \gamma_0, \beta) = \Pi_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta) + o_P(1).$$

We also proved that, as  $n \rightarrow +\infty$ ,

$$\widehat{\vartheta}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta) \xrightarrow{\mathbb{P}} \vartheta(\rho_0, \gamma_0, \beta).$$

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

$$\lim_{n \rightarrow +\infty} \frac{\widehat{\vartheta}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta)}{\vartheta(\rho_0, \gamma_0, \beta)} = 1,$$

and

$$\begin{aligned} & \mathcal{T}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta) \\ &= \frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\widehat{\vartheta}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta)} + \frac{1}{\widehat{\vartheta}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta)} \times o_P(1) \\ &= \frac{\Pi_n(\rho_0, \gamma_0, \beta)}{\vartheta(\rho_0, \gamma_0, \beta)} \times \frac{\vartheta(\rho_0, \gamma_0, \beta)}{\widehat{\vartheta}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta)} + \frac{1}{\widehat{\vartheta}_n(\rho_{s(n)}, \widehat{\gamma}_{0,s(n)}, \beta)} \times o_P(1) \\ &\xrightarrow{\mathbb{P}} \mathcal{T}_n(\rho_0, \gamma_0, \beta). \end{aligned}$$

### Appendix A.8.2. Proof of [ii] and [iii]

The convergence of  $\hat{\vartheta}_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta)$  to  $\vartheta(\rho_0, \gamma_0, \beta)$  remains true under the local alternative  $H_\beta^{(n)}$  by contiguity. Then, under  $H_\beta^{(n)}$  and as  $n \rightarrow +\infty$ , using Le Cam's third lemma, we have

$$\mathcal{T}_n(\rho_{s(n)}, \gamma_0, \beta) = \frac{\Pi_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta)}{\hat{\vartheta}_n(\rho_{s(n)}, \hat{\gamma}_{0,s(n)}, \beta)} \xrightarrow{\mathcal{D}} \mathcal{N}(\vartheta(\rho_0, \gamma_0, \beta), 1).$$

Then, we can say that the power of the test remains the same.

For more proof details, see [36].

## References

- Härdle, W.; Tsybakov, A.; Yang, L. Nonparametric vector autoregression. *J. Stat. Plan. Inference* **1998**, *68*, 221–245. [\[CrossRef\]](#)
- Chen, G.; Gan, M.; Chen, G. Generalized exponential autoregressive models for nonlinear time series: Stationarity, estimation and applications. *Inf. Sci.* **2018**, *438*, 46–57. [\[CrossRef\]](#)
- Yau, Y.C.; Zhao, Z. The asymptotic behavior of the likelihood ratio statistic for testing a shift in mean in a sequence of independent normal variates. *J. R. Statist. Soc.* **2016**, *48*, 895–916. [\[CrossRef\]](#)
- Ngatchou-Wandji, J.; Ltaifa, M. Detecting weak changes in the mean of a class of nonlinear heteroscedastic models. *Commun. Stat. Simul. Comput.* **2023**, 1–33. [\[CrossRef\]](#)
- Csörgő, M.; Horváth, L.; Szyszkowicz, B. Integral tests for suprema of kiefer processes with application. *Stat. Risk Model.* **1997**, *15*, 365–378. [\[CrossRef\]](#)
- MacNeill, I.B. Tests for change of parameter at unknown times and distributions of some related functionals on brownian motion. *Ann. Stat.* **1974**, *2*, 950–962. [\[CrossRef\]](#)
- Chernoff, H.; Zacks, S. Estimating the current mean of a normal distribution which is subjected to changes in time. *Ann. Math. Stat.* **1964**, *35*, 999–1018. [\[CrossRef\]](#)
- Davis, R.A.; Huang, D.; Yao, Y.-C. Testing for a change in the parameter values and order of an autoregressive model. *Ann. Stat.* **1995**, *23*, 282–304. [\[CrossRef\]](#)
- Vogelsang, T.J. Wald-type tests for detecting breaks in the trend function of a dynamic time series. *Econom. Theory* **1997**, *13*, 818–848. [\[CrossRef\]](#)
- Andrews, D.W.; Ploberger, W. Optimal tests when a nuisance parameter is present only under the alternative. *Econom. J. Econom. Soc.* **1994**, *62*, 1383–1414. [\[CrossRef\]](#)
- Andrews, D.W. Tests for parameter instability and structural change with unknown change point. *Econom. J. Econom. Soc.* **1993**, *61*, 821–856. [\[CrossRef\]](#)
- Lavielle, M.; Lebarbier, E. An application of MCMC methods for the multiple change-points problem. *Signal Process.* **2001**, *81*, 39–53. [\[CrossRef\]](#)
- Lebarbier, É. Detecting multiple change-points in the mean of Gaussian process by model selection. *Signal Process.* **2005**, *85*, 717–736. [\[CrossRef\]](#)
- Davis, R.A.; Lee, T.C.; Rodriguez-Yam, G.A. Break detection for a class of nonlinear time series models. *J. Time Ser. Anal.* **2008**, *29*, 834–867. [\[CrossRef\]](#)
- Fotopoulos, S.B.; Jandhyala, V.K.; Tan, L. Asymptotic study of the change-point mle in multivariate gaussian families under contiguous alternatives. *J. Stat. Plan. Inference* **2009**, *139*, 1190–1202. [\[CrossRef\]](#)
- Huh, J. Detection of a change point based on local-likelihood. *J. Multivar. Anal.* **2010**, *101*, 1681–1700. [\[CrossRef\]](#)
- Jarušková, D. Asymptotic behaviour of a test statistic for detection of change in mean of vectors. *J. Stat. Plan. Inference* **2010**, *140*, 616–625. [\[CrossRef\]](#)
- Piterbarg, V.I. High Derivations for Multidimensional Stationary Gaussian Process with Independent Components. In *Stability Problems for Stochastic Models*; De Gruyter: Berlin, Germany; Boston, MA, USA, 1994; pp. 197–210.
- Chen, K.; Cohen, A.; Sackrowitz, H. Consistent multiple testing for change points. *J. Multivar. Anal.* **2011**, *102*, 1339–1343. [\[CrossRef\]](#)
- Vostrikova, L.Y. Detecting “disorder” in multidimensional random processes. In *Doklady Akademii Nauk*; Russian Academy of Sciences: Moscow, Russia, 1981; Volume 259, pp. 270–274.
- Cohen, A.; Sackrowitz, H.B.; Xu, M. A new multiple testing method in the dependent case. *Ann. Stat.* **2009**, *37*, 1518–1544. [\[CrossRef\]](#)
- Prášková, Z.; Chochola, O. M-procedures for detection of a change under weak dependence. *J. Stat. Plan. Inference* **2014**, *149*, 60–76. [\[CrossRef\]](#)
- Dupuis, D.; Sun, Y.; Wang, H.J. *Detecting Change-Points in Extremes*; International Press of Boston: Boston, MA, USA, 2015.
- Badagián, A.L.; Kaiser, R.; Peña, D. Time series segmentation procedures to detect, locate and estimate change-points. In *Empirical Economic and Financial Research*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 45–59.



25. Yau, C.Y.; Zhao, Z. Inference for multiple change points in time series via likelihood ratio scan statistics. *J. R. Stat. Soc. Ser. (Stat. Methodol.)* **2016**, *78*, 895–916. [[CrossRef](#)]
26. Ruggieri, E.; Antonellis, M. An exact approach to bayesian sequential change-point detection. *Comput. Stat. Data Anal.* **2016**, *97*, 71–86. [[CrossRef](#)]
27. Horváth, L.; Miller, C.; Rice, G. A new class of change point test statistics of Rényi type. *J. Bus. Econ. Stat.* **2020**, *38*, 570–579. [[CrossRef](#)]
28. Ltaifa, M. Tests Optimaux pour Détecter les Signaux Faibles dans les Séries Chronologiques. Ph.D. Thesis, Université de Lorraine, Nancy, France, Université de Sousse, Sousse, Tunisia, 2021.
29. Bahadur, R.R.; Rao, R.R. On deviations of the sample mean. *Ann. Math. Stat.* **1960**, *31*, 1015–1027. [[CrossRef](#)]
30. Parzen, E. On estimation of a probability density function and mode. *Ann. Math. Stat.* **1962**, *33*, 1065–1076. [[CrossRef](#)]
31. Hall, P.; Heyde, C.C. *Martingale Limit Theory and Its Application*; Academic Press: Cambridge, MA, USA, 2014.
32. Le Cam, L. The central limit theorem around 1935. *Stat. Sci.* **1986**, *1*, 78–91.
33. Dreesbeke and Fine, J.-J. *Inférence Non Paramétrique: Les Statistiques de Rangs*; Éditions de l'Université de Bruxelles: Bruxelles, Belgium; Éditions Ellipses: Paris, France, 1996.
34. Ngatchou-Wandji, J. Estimation in a class of nonlinear heteroscedastic time series models. *Electron. J. Stat.* **2008**, *2*, 40–62. [[CrossRef](#)]
35. Ngatchou-Wandji, J. Checking nonlinear heteroscedastic time series models. *J. Stat. Plan. Inference* **2005**, *133*, 33–68. [[CrossRef](#)]
36. Salman, Y. Testing a Class of Time-Varying Coefficients CHARN Models with Application to Change-Point Study. Ph.D. Thesis, Lorraine University, Nancy, France; Lebanese University, Beirut, Lebanese, 2022.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.