Annals of the Institute of Statistical Mathematics (2022) 74:801–835
https://doi.org/10.1007/s10463-021-00815-1

Inference for nonstationary time series of counts with application to change-point problems

William Kengne[1] **· Isidore S. Ngongo**[2,3]

Received: 26 January 2021 / Revised: 10 November 2021 / Accepted: 15 November 2021 /
Published online: 18 January 2022

Abstract

We consider an integer-valued time series $(Y_t)_{t\in\mathbb{Z}}$ where the model after a time k^* is Poisson autoregressive with the conditional mean that depends on a parameter $\theta^* \in \Theta \subset \mathbb{R}^d$. The structure of the process before k^* is unknown; it could be any other integer-valued process, that is, $(Y_t)_{t\in\mathbb{Z}}$ could be nonstationary. It is established that the maximum likelihood estimator of θ^* computed on the nonstationary observations is consistent and asymptotically normal. Subsequently, we carry out the sequential change-point detection in a large class of Poisson autoregressive models, and propose a monitoring scheme for detecting change. The procedure is based on an updated estimator, which is computed without the historical observations. The above results of inference in a nonstationary setting are applied to prove the consistency of the proposed procedure. A simulation study as well as a real data application are provided.

Keywords Time series of counts · Poisson autoregression · likelihood estimation · Change-point · Sequential detection · Weak convergence

William Kengne was developed within the ANR BREAKRISK: ANR-17-CE26-0001-01 and the CY Initiative of Excellence (Grant "Investissements d'Avenir" ANR-16-IDEX-0008), Project "EcoDep" PSI-AAP2020-0000000013.

✉ William Kengne
william.kengne@u-cergy.fr

Isidore S. Ngongo
ngongoisidore@gmail.com

[1] THEMA, CY Cergy Paris Université, 33 Boulevard du Port, 95011 Cergy-Pontoise Cedex, France

[2] ENS, Université de Yaoundé 1, B.P. 47 Yaoundé, Cameroun

[3] S.A.M.M., Université Paris 1 Panthéon-Sorbonne, 90, rue de Tolbiac, 75634 Paris, France

1 Introduction

We consider a process $Y = (Y_t)_{t\in\mathbb{Z}}$ satisfying

$$Y_t|\mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \quad \text{with} \quad \lambda_t = \mathbb{E}(Y_t|\mathcal{F}_{t-1}); \tag{1}$$

where $\mathcal{F}_t = \sigma(Y_s, s \le t)$ is the σ-field generated by the whole past. A large literature on this model has recently been developed by assuming that $\lambda_t = \mathbb{E}(Y_t|\mathcal{F}_{t-1}) = f(Y_{t-1}, Y_{t-2}, \ldots)$ for all $t \in \mathbb{Z}$, where f is a measurable non-negative function, satisfying some Lipschitz-type conditions. This entails that the process $(Y_t, \lambda_t)_{t\in\mathbb{Z}}$ is strictly stationary with finite moment of any order. But such result does not hold in many practical situations. For instance, in the change-point problem, it often holds that

$$\lambda_t = \begin{cases} f_0(Y_{t-1}, Y_{t-2}, \ldots) \text{ for } t \le k^*, \\ f_1(Y_{t-1}, Y_{t-2}, \ldots) \text{ for } t > k^* \end{cases}$$

with $f_0 \neq f_1$ and $k^* \in \mathbb{Z}$. Thus, the process $(Y_t, \lambda_t)_{t\in\mathbb{Z}}$ is not stationary.

We consider a nonstationary autoregressive process $Y = (Y_t)_{t\in\mathbb{Z}}$ in a parametric framework; we assume that Y satisfies

$$Y_t|\mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \quad \text{with} \quad \lambda_t = \mathbb{E}(Y_t|\mathcal{F}_{t-1}) = f_{\theta^*}(Y_{t-1}, Y_{t-2}, \ldots) \quad \text{for all } t > k^* \tag{2}$$

with $k^* \in \mathbb{Z}$, where θ^* is the parameter belonging to a compact set $\Theta \subset \mathbb{R}^d$ ($d \in \mathbb{N}$) and f_θ is a measurable non-negative function, assumed to be known up to the parameter θ. If (2) holds for $t < k^*$, then with some Lipschitz-type conditions on f, there exists a strictly stationary and ergodic solution of (2), denoted by $(\widetilde{Y}_t, \widetilde{\lambda}_t)_{t\in\mathbb{Z}}$, with finite moment of any order (see for instance Doukhan et al. 2012). It is assumed here that the process $Y = (Y_t)_{t\in\mathbb{Z}}$ is a solution of (2) (with $t > k^*$) and we focus in a more general situation where the structure of the process $(Y_t)_{t\le k^*}$ is assumed to be unknown; it could be a Poisson autoregressive model depending on a parameter different from θ^* or could be any other integer-valued time series.

In this work, we firstly study the inference of the parameter θ^* in the model (2). This task has been considered by several authors; see among others (Fokianos et al. 2009; Fokianos and Tjøstheim 2012; Doukhan and Kengne 2015). These works (and many other) have been developed under the assumption that the process $(Y_t)_{t\in\mathbb{Z}}$ is strictly stationary, which restricts the application area of such results. To deal with the model (2), we conduct some preliminary works for the approximation of the nonstationary process with its stationary regime. Under some classical Lipschitz-type condition on the function f, there exists (see Doukhan et al. 2012, 2013) a strictly stationary process $\tilde{Y} = (\tilde{Y}_t)_{t\in\mathbb{Z}}$ with finite moments of any order, satisfying:

$$\tilde{Y}_t|\tilde{\mathcal{F}}_{t-1} \sim \text{ Poisson } (\tilde{\lambda}_t) \text{ with } \tilde{\lambda}_t = f_{\theta^*}(\tilde{Y}_{t-1}, \tilde{Y}_{t-2}, \ldots) \quad \text{for } t \in \mathbb{Z} \tag{3}$$

where $\tilde{\mathcal{F}}_t = \sigma(\tilde{Y}_s, s \le t)$ is the σ-field generated by the whole past of $\tilde{Y}$.

Let us remark that, models (1), (2) and (3) can be represented in terms of Poisson processes. Let $\{N_t(\cdot); t \in \mathbb{Z}\}$ be a sequence of independent Poisson processes of unit intensity. Y_t and $\tilde{Y}_t$ can respectively be seen as the number (say $N_t(\lambda_t)$) of events of $N_t(\cdot)$ that occur in the time interval $[0, \lambda_t]$ and $[0, \tilde{\lambda}_t]$. Therefore, we can also write

$$\begin{aligned} Y_t = N_t(\lambda_t),\ \tilde{Y}_t = N_t(\tilde{\lambda}_t) \text{ with } \lambda_t \\ = f_{\theta^*}(Y_{t-1}, Y_{t-2} \dots) \quad \text{and} \quad \tilde{\lambda}_t = f_{\theta^*}(\tilde{Y}_{t-1}, \tilde{Y}_{t-2}, \dots) \quad \text{for all } t > k^*. \end{aligned} \tag{4}$$

This representation is useful to approximate the processes $(Y_t)_{t\geq k^*}$ and $(\tilde{Y}_t)_{t\geq k^*}$. The question of this approximation has been addressed by Doukhan and Kengne (2015) (see Remark 4.1). In this work, we provide a detailed proof of this problem. In particular, we show that the expectation $\mathbb{E}|Y_{k^*+\ell} - \tilde{Y}_{k^*+\ell}|$ (for $\ell \geq 1$) can be controlled and tends to zero when ℓ goes to infinity, see Lemma 1. These approximation results are applied to establish that the conditional maximum likelihood estimator (MLE) of θ^*, based on the nonstationary observations, is consistent and asymptotically normal. Also, let us stress that numerous papers on change-point problem assume that the process is stationary after the change-point; see for instance (Doukhan and Kengne 2015; Diop and Kengne 2017; Franke et al. 2012; Kirch and Kamgaing 2015). This paper provides tools to avoid such condition, which is quite restrictive in practice.

As a second contribution, we consider the structural change-point problem in the Poisson autoregressive models. In the retrospective (or off-line) framework, this issue has already been addressed. See for instance (Franke et al. 2012; Kang and Lee 2014; Doukhan and Kengne 2015; Diop and Kengne 2017). But these works suffer from a drawback: the (asymptotic) study under the presence of change is either missing or done with the stationarity assumption on the observations after change-points, which is unrealistic in many practical problems. For the procedure proposed by these authors, the stationarity assumption after the change-point can be relaxed by applying Theorem 1 (see below), which establishes the consistency of the conditional MLE of the parameter of the nonstationary model after the change-point. In the sequel, we focus on the sequential (or on-line) framework.

Assume that the process $Y = (Y_t)_{t\in\mathbb{Z}}$ satisfies

$$Y_t|\mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \text{ with } \lambda_t = \begin{cases} f_{\theta_0^*}(Y_{t-1}, Y_{t-2}, \dots) \text{ for } t \leq k^*, \\ f_{\theta_1^*}(Y_{t-1}, Y_{t-2}, \dots) \text{ for } t > k^* \end{cases} \tag{5}$$

where θ_0^*, θ_1^* are the parameters belonging to a compact set $\Theta \subset \mathbb{R}^d$ ($d \in \mathbb{N}$) and k^* is a positive integer, standing for the possible instant of change. If $\theta_0^* \neq \theta_1^*$, then a structural change occurs at time k^*; otherwise, no change has occurred, and the model (5) can be simply written as

$$Y_t|\mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \text{ with } \lambda_t = f_{\theta_0^*}(Y_{t-1}, Y_{t-2}, \dots) \quad \text{for } t \in \mathbb{Z}. \tag{6}$$

We follow the paradigm of Chu et al. (1996). The general idea is to use the observations $(Y_1, \dots, Y_m)$ (called the historical data) that depend on the parameter θ_0^*. Then from the time $m+1$, one would like to test sequentially whether a structural change occurs and trigger an alarm if so, by ensuring that the probability of false alarm does not exceed a fixed level α. More precisely, $k^* > m$ and $(Y_1, \dots, Y_m)$ is assumed to be generated from the model (6), depending on θ_0^* (without change); we are going to observe new data $Y_{m+1}, Y_{m+2}, \dots, Y_{m+k}, \dots$. For each new observation Y_{m+k}, we would like to know if it is generated from a model depending on θ_0^* or from a model depending on θ_1^*, with $\theta_0^* \neq \theta_1^*$. This problem can be treated as a classical hypothesis testing:

$\mathbf{H_0}$ θ_0^* is constant over $\cdots, Y_{-1}, Y_0, Y_1, \ldots, Y_m, Y_{m+1}, \ldots$ *i.e.* $(Y_t)_{t\in\mathbb{Z}}$ satisfies (5) with $\theta_0^* = \theta_1^*$;

$\mathbf{H_1}$ the process $(Y_t)_{t\in\mathbb{Z}}$ satisfies (5) with $\theta_0^* \neq \theta_1^*$ and $k^* > m$.

Numerous works have been done in the sequential change-point detection according to such paradigm. See among others papers, Horváth et al. (2004), Gombay and Serban (2009), Na et al. (2011) and Bardet and Kengne (2014) for several test procedures for sequential change detection in a general class of time series models, including linear and GARCH-type models. Kengne (2015) proposed a fluctuation-type test procedure for sequential change detection in a large class of Poisson autoregressive model. Recently, Kirch and Kamgaing (2015) and Kirch and Weber (2018) have considered a large class of models (that includes continuous and discrete valued time series) and developed a general setup based on estimating functions for sequential change-point detection. Estimating functions is a general estimation method and some classical procedure such as, the likelihood estimator, the least square estimator,...can be treated in many cases as a particular class of estimating functions. It is well-known (Godambe 1960) that the optimal estimating function in several classical parametric models is based on the score function. In the case of infinite memory process considered here, a more complex class of estimating functions is needed; this involves some difficulties in the application of their procedure. Moreover, Kirch and Kamgaing (2015) and Kirch and Weber (2018) impose some regularity conditions on the process after the change-point. These conditions, which are not easy to verify in general, are somewhere sufficient to unify the treatment of the large class models that they have considered.

We carry out a sequential test in the spirit of Bardet and Kengne (2014), and propose an open-end and closed-end (see below) procedure for monitoring changes in the model (5). We develop a procedure where the recursive estimator is computed without the historical observations. It is shown that the detector converges to a known distribution under the null hypothesis. Under the alternative, we do not need any additional assumption on the process after the change-point. The consistency of the procedure is established even in the nonstationary setting (the previous study of inference in nonstationary models plays a key role in the proof of this result). Moreover, the test developed here is intended for early detection of change than the aforementioned procedure, since it has displayed a detection delay that can be bounded by $\mathcal{O}_P(m^{1/2+\epsilon})$ for any $\epsilon > 0$.

In the following Sect. 2, some classical assumptions on the model (2) as well as some examples are provided. The inference in the nonstationary process Y is conducted in Sect. 3. Section 4 focuses on the sequential change-point detection. Some numerical results are displayed in Sect. 5, whereas Sect. 6 is devoted to some concluding remarks. The proofs of the main results are provided in Sect. 7.

2 Assumptions and examples

2.1 Assumptions

We will use the following classical notations:

1. $\|y\| := \sum_{j=1}^{p} |y_j|$ for any $y \in \mathbb{R}^p$;
2. for any compact set $\mathcal{K} \subseteq \mathbb{R}^d$ and for any function $g : \mathcal{K} \longrightarrow \mathbb{R}^{d'}$, $\|g\|_{\mathcal{K}} = \sup_{\theta \in \mathcal{K}}(\|g(\theta)\|)$;
3. for any set $\mathcal{K} \subseteq \mathbb{R}^d$, $\mathring{\mathcal{K}}$ denotes the interior of $\mathcal{K}$;
4. $\mathbb{N} = \{1, 2, 3, \dots\}$ and $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$;
5. if Y is a random vector with finite s-order moments, we set $\|Y\|_s = (\mathbb{E}\|Y\|^s)^{1/s}$.

Throughout the sequel, we will assume that the function $\theta \mapsto f_\theta$ is twice continuously differentiable on Θ and we need the following conditions on the model (2). For $i = 0,\ 1,\ 2$, define

Assumption $\mathbf{A}_i(\Theta)$ $\|f_\theta(0)\|_\Theta < \infty$ ($\|\partial^i f_\theta(0)/\partial\theta^i\|_\Theta < \infty$ when $i = 1, 2$) and there exists a sequence of non-negative real numbers $(\alpha_k^{(i)})_{k\geq 1}$ satisfying $\sum\limits_{j=1}^{\infty} \alpha_k^{(0)} < 1$ (when $i = 0$) and $\sum\limits_{j=1}^{\infty} \alpha_k^{(i)} < \infty$ (when $i = 1, 2$) such that for all $y, y' \in (\mathbb{R}^+)^{\mathbb{N}}$,

$$\|f_\theta(y) - f_\theta(y')\|_\Theta \leq \sum_{j=i}^{\infty} \alpha_j^{(0)} |y_j - y'_j| \quad \text{or} \quad \left\| \frac{\partial^i f_\theta(y)}{\partial\theta^i} - \frac{\partial^i f_\theta(y')}{\partial\theta^i} \right\|_\Theta$$
$$\leq \sum_{k=1}^{\infty} \alpha_k^{(i)} |y_k - y'_k| \ \text{(when } i = 1, 2).$$

Under the assumption $\mathbf{A}_0(\Theta)$, Doukhan et al. (2012, 2013) proved that the model (3) has a strictly stationary solution $(\tilde{Y}_t, \tilde{\lambda}_t)_{t\in\mathbb{Z}}$, which is τ-weakly dependent with finite moment of any order (see also Doukhan and Wintenberger 2008). But, such result cannot be applied to process Y satisfying (2), since the structure of the past before k^* is unknown. The following proposition shows that if $(Y_t)_{t\leq k^*}$ has finite moments of any order, then it also holds for $(Y_t)_{t>k^*}$.

Proposition 1 *Assume* $\mathbf{A}_0(\Theta)$. *Let* $Y = (Y_t)_{t\in\mathbb{Z}}$ *satisfy* (2) *and for any* $r \geq 1$, *there exists a constant* $C_{r,0} > 0$ *such that* $\mathbb{E}Y_t^r \leq C_{r,0}$ *for all* $t \leq k^*$. *Then, there exists* $C > 0$ *such that*

$$\mathbb{E}Y_{k^*+\ell}^r \leq C \quad \text{for all } \ell \geq 1.$$

As we state above, $(Y_t)_{t\leq k^*}$ could be any integer-valued time series, and we assume in the sequel that:

$$\text{for any } r \geq 1, \ \text{ there exists } C_{r,0} > 0 \text{ such that } \mathbb{E}Y_t^r \leq C_{r,0} \text{ for all } t \leq k^*. \tag{7}$$

The conditions $A_1(\Theta)$, $A_2(\Theta)$ as well as the following assumptions $D(\Theta)$, $Id(\Theta)$ and $Var(\Theta)$ are classical for inference on such model see Doukhan and Kengne (2015).

Assumption D(Θ) $\exists \underline{c} > 0$ such that $\inf_{\theta \in \Theta}(f_\theta(y)) \geq \underline{c}$ for all $y \in (\mathbb{R}^+)^{\mathbb{N}}$.

Assumption Id(Θ For all $(\theta, \theta') \in \Theta^2$, $\Big(f_\theta(Y_{t-1}, \dots) = f_{\theta'}(Y_{t-1}, \dots)$ a.s. for some $t > k^*\Big) \Rightarrow \theta = \theta'$.

Assumption Var(Θ For all $\theta \in \Theta$ and $t > k^*$, the components of the vector $\dfrac{\partial f_\theta}{\partial \theta}(Y_{t-1,\dots})$ are a.s. linearly independent.

Also, we will assume in the sequel that the true parameter θ^* belongs to $\overset{\circ}{\Theta}$ (the interior of Θ).

2.2 Examples

2.2.1 Linear Poisson autoregression

We consider an integer-valued time series $(Y_t)_{t \in \mathbb{Z}}$ satisfying for any $t \in \mathbb{Z}$

$$Y_t | \mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \quad \text{with} \quad \lambda_t = \phi_0(\theta^*) + \sum_{k \geq 1} \phi_k(\theta^*) Y_{t-k} \tag{8}$$

with $\theta^* \in \Theta \subset \mathbb{R}^d$, where the functions $\theta \mapsto \phi_k(\theta)$ are positive, twice continuous differentiable such that $\sum_{k \geq 1} \|\phi_k(\theta)\|_\Theta < 1$, $\sum_{k \geq 1} \|\phi'_k(\theta)\|_\Theta < \infty$, $\sum_{k \geq 1} \|\phi''_k(\theta)\|_\Theta < \infty$ and $\inf_{\theta \in \Theta} \phi_0(\theta) > 0$ (see also Doukhan and Kengne 2015). Thus, Assumptions $A_i(\Theta)$ $i = 0, 1, 2$ and $D(\Theta)$ hold. Moreover, if there exists a finite subset $I \subset \mathbb{N} - \{0\}$ such that the function $\theta \mapsto (\phi_k(\theta))_{k \in I}$ is injective, then assumption $Id(\Theta)$ holds and the model (8) is identifiable. Finally, assumption $Var(\Theta)$ holds if for any $\theta \in \Theta$, there exists d functions $\phi_{k_1}, \dots, \phi_{k_d}$ such that the matrix $\left(\dfrac{\partial \phi_{k_j}}{\partial \theta}\right)_{1 \leq j \leq d}$ (computed at θ) has a full rank. This is the case in the classical useful situations, such as, for instance, the INGARCH(p, q) model below.

The classical Poisson INGARCH(p, q) (see Ferland et al. 2006 or Weiß 2009) is obtained with

$$\lambda_t = \alpha_0^* + \sum_{k=1}^{p} \alpha_k^* \lambda_{t-k} + \sum_{k=1}^{q} \beta_k^* Y_{t-k}. \tag{9}$$

The true parameter $\theta^* = (\alpha_0^*, \alpha_1^*, \dots, \alpha_p^*, \beta_1^*, \dots, \beta_q^*) \in \Theta$ where Θ is a compact subset of $(0, +\infty) \times [0, +\infty)^{p+q}$ such that $\sum_{k=1}^{p} \alpha_k + \sum_{k=1}^{q} \beta_k < 1$ for all $\theta = (\alpha_0, \alpha_1, \dots, \alpha_p, \beta_1, \dots, \beta_q) \in \Theta$. This model is a special case of the model (8) since we can find a sequence of functions $(\psi_k(\theta))_{k \geq 0}$ such that $\lambda_t = \psi_0(\theta^*) + \sum_{k \geq 1} \psi_k(\theta^*) Y_{t-k}$.

In the model (8), it often happens in practice that

$$\lambda_t = \begin{cases} \phi_0(\theta_0^*) + \sum_{k\geq 1} \phi_k(\theta_0^*) Y_{t-k} & \text{for } t \leq k^*, \\ \phi_0(\theta_1^*) + \sum_{k\geq 1} \phi_k(\theta_1^*) Y_{t-k} & \text{for } t > k^* \end{cases} \tag{10}$$

with $\theta_0^* \neq \theta_1^*$. There exists several references in the literature (see for instance Doukhan and Kengne 2015; Ahmad and Francq 2016) that address the inference on θ_0^* based on the observations of the stationary process $(Y_t)_{t\leq k^*}$. These results which are heavily based on the stationarity of the process cannot work for θ_1^*. Section 3 focuses on the estimation of θ_1^* based on the nonstationary process $(Y_t)_{t>k^*}$.

2.2.2 Threshold Poisson autoregression

We consider a threshold Poisson autoregressive model defined by:

$$\begin{aligned} Y_t|\mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \quad \text{with} \quad \lambda_t &= \phi_0(\theta^*) \\ &+ \sum_{k\geq 1} \Big(\phi_k^+(\theta^*) \max(Y_{t-k} - \ell, 0) + \phi_k^-(\theta^*) \min(Y_{t-k}, \ell) \Big) \end{aligned} \tag{11}$$

where $\phi_0(\theta) > 0$, $\phi_k^+(\theta), \phi_k^-(\theta) \geq 0$ for all $\theta \in \Theta$ and $\ell \in \mathbb{N}$. We can also write

$$\lambda_t = \phi_0(\theta^*) + \sum_{k\geq 1} \Big(\phi_k^-(\theta^*) Y_{t-k} + \big(\phi_k^+(\theta^*) - \phi_k^-(\theta^*)\big) \max(Y_{t-k} - \ell, 0) \Big).$$

This is an example of a nonlinear model called an integer-valued threshold ARCH (or INTARCH) see Doukhan and Kengne (2015); see also Franke et al. (2012) for INTARCH(1) model. Such model is often used to capture piecewise phenomenon. ℓ is the threshold parameter of the model. If the functions $\theta \mapsto \phi_k^+(\theta)$ and $\theta \mapsto \phi_k^-(\theta)$ are twice continuously differentiable such that $\sum_{k\geq 1} \max\big(\|\phi_k^+(\theta)\|_\Theta, \|\phi_k^-(\theta)\|_\Theta\big) < 1$, $\sum_{k\geq 1} \max\big(\|\frac{\partial}{\partial\theta}\phi_k^+(\theta)\|_\Theta, \|\frac{\partial}{\partial\theta}\phi_k^-(\theta)\|_\Theta, \|\frac{\partial^2}{\partial\theta^2}\phi_k^+(\theta)\|_\Theta, \|\frac{\partial^2}{\partial\theta^2}\phi_k^-(\theta)\|_\Theta\big) < \infty$, then $\mathrm{A}_i(\Theta)$ $i = 0, 1, 2$ hold. Furthermore, conditions on $\mathrm{D}(\Theta)$, $\mathrm{Id}(\Theta)$ and $\mathrm{Var}(\Theta)$ are obtained as above.

3 Likelihood inference

We focus on the inference for the model (2); that is, we consider the process $Y = (Y_t)_{t\in\mathbb{Z}}$ satisfying

$$Y_t|\mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \text{ with } \lambda_t = \mathbb{E}(Y_t|\mathcal{F}_{t-1}) = f_{\theta^*}(Y_{t-1}, Y_{t-2}, \ldots) \quad \text{for all } t > k^*. \tag{12}$$

Assume that a trajectory $(Y_{k^*+1}, \ldots, Y_{k^*+n})$ of the process $(Y_t)_{t>k^*}$ is observed. The conditional (log)-likelihood (up to a constant) computed on a segment $T \subset \{k^* + 1, k^* + 2, \ldots\}$ is given by

$$L_n(T, \theta) = \sum_{t\in T} (Y_t \log \lambda_t(\theta) - \lambda_t(\theta)) = \sum_{t\in T} \ell_t(\theta) \text{ with } \ell_t(\theta) = Y_t \log \lambda_t(\theta) - \lambda_t(\theta)$$

where $\lambda_t(\theta) = f_\theta(Y_{t-1}, \dots)$. In the sequel, we use the notation $f_\theta^t := f_\theta(Y_{t-1}, \dots)$. An approximation of the conditional (log)-likelihood is

$$\widehat{L}_n(T, \theta) = \sum_{t \in T} (Y_t \log \widehat{\lambda}_t(\theta) - \widehat{\lambda}_t(\theta)) = \sum_{t \in T} \widehat{\ell}_t(\theta) \quad \text{with} \quad \widehat{\ell}_t(\theta) = Y_t \log \widehat{\lambda}_t(\theta) - \widehat{\lambda}_t(\theta) \tag{13}$$

where $\widehat{\lambda}_t(\theta) := \widehat{f}_\theta^t := f_\theta(Y_{t-1}, \dots, Y_1, 0, \dots)$. The MLE of θ^* computed on T is defined by

$$\widehat{\theta}(T) = \text{ argmax}_{\theta \in \Theta}(\widehat{L}_n(T, \theta)). \tag{14}$$

For any $k, k' \in \mathbb{Z}$ such as $k \le k'$, denote

$$T_{k,k'} = \{k, k+1, \dots, k'\}.$$

The following theorem establishes that the MLE of θ^* based on the nonstationary process Y [satisfying (12)] is consistent.

Theorem 1 *Assume that* $\theta^* \in \mathring{\Theta}$, $D(\Theta)$, $\mathrm{Id}(\Theta)$, $A_0(\Theta)$ *and* (7) *hold with*

$$\alpha_j^{(0)} = O(j^{-\gamma}), \quad \text{for some } \gamma > 3/2. \tag{15}$$

Then, it holds that

$$\widehat{\theta}(T_{k^*+1,k^*+n}) \xrightarrow[n \to +\infty]{\text{a.s.}} \theta^*.$$

To address the asymptotic normality, set

$$\begin{aligned} \widetilde{\Sigma} &= E\left(\frac{1}{\widetilde{f}_{\theta^*}^0} \left(\frac{\partial}{\partial \theta} \widetilde{f}_{\theta^*}^0 \right) \left(\frac{\partial}{\partial \theta} \widetilde{f}_{\theta^*}^0 \right)' \right) \quad \text{with} \\ \widetilde{f}_{\theta^*}^0 &= f_{\theta^*}(\tilde{Y}_{-1}, \tilde{Y}_{-2} \dots), \end{aligned} \tag{16}$$

where $(\tilde{Y}_t)_{t \in \mathbb{Z}}$ is a stationary and ergodic process satisfying (3) [see also (33)] and $'$ denotes the transpose. This matrix is symmetric and positive definite (see Doukhan and Kengne 2015). According to the proof of Theorem 2, the matrix

$$\widehat{\Sigma}_n = \left(\frac{1}{n} \sum_{t=k^*+1}^{k^*+n} \frac{1}{\widehat{f}_\theta^t} \left(\frac{\partial}{\partial \theta} \widehat{f}_\theta^t \right) \left(\frac{\partial}{\partial \theta} \widehat{f}_\theta^t \right)' \right) \Bigg|_{\theta = \widehat{\theta}(T_{k^*+1,k^*+n})}$$

is a consistent estimator of $\widetilde{\Sigma}$. The asymptotic normality of the MLE is displayed in the following theorem.

Theorem 2 *Under the assumptions of Theorem* 1 *and* $Var(\Theta)$ *if* $A_i(\Theta)$ $i = 1, 2$ *hold with*

$$\alpha_j^{(i)} = O(j^{-\gamma}), \quad \text{for some } \gamma > 3/2, \tag{17}$$

then,

$$\sqrt{n}(\widehat{\theta}(T_{k^*+1,k^*+n}) - \theta^*) \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}(0, \widetilde{\Sigma}^{-1}).$$

4 Sequential change-point detection

Let $(Y_1, \dots, Y_m)$ be historical observations, assumed to be a trajectory of a stationary and ergodic time series $(Y_t)_{t\leq k^*}$ satisfying (6) with parameter θ_0^*, and with finite moments of any order (such process exists according to Doukhan et al. 2012, 2013). We consider the online change-point detection in the model (5) and focus on the hypotheses $\mathbf{H_0}$ and $\mathbf{H_1}$ (defined in the introduction).

The MLE of θ_0^*, computed on the historical observations is defined by

$$\widehat{\theta}(T_{1,m}) = \underset{\theta\in\Theta}{\text{argmax}}\,(\widehat{L}(T_{1,m}, \theta)). \tag{18}$$

According to Sect. 3 (see also Doukhan and Kengne 2015), this estimator is consistent and asymptotically normal. The asymptotic covariance matrix of $\widehat{\theta}(T_{1,m})$ is Σ^{-1} with

$$\Sigma = E\left(\frac{1}{f^0_{\theta_0^*}}\left(\frac{\partial}{\partial\theta}f^0_{\theta_0^*}\right)\left(\frac{\partial}{\partial\theta}f^0_{\theta_0^*}\right)'\right). \tag{19}$$

Recall that, the fluctuation-type test proposed by Chu et al. (1996) is based on the discrepancy between the estimators of the model's parameters. The classical idea of the fluctuation test is to evaluate at the monitoring step $m+k$, the distance between $\widehat{\theta}(T_{1,m})$ and $\widehat{\theta}(T_{1,m+k})$ by expecting this will be large enough if a change occurs at time $m+k^*$ (with $k^* < k$). Such idea has been employed by Na et al. (2011); Kengne (2015), among others. As pointed out by Bardet and Kengne (2014), the recursive estimator $\widehat{\theta}(T_{1,m+k})$ heavily depends on the historical data and the detection delay of such procedure may not be quite efficient.

We follow the ideas of Bardet and Kengne (2014) and propose a procedure which is based on the detector:

$$\widehat{C}_{k,\ell} := \sqrt{m}\,\frac{k-\ell}{k}\,\left\|\widehat{\Sigma}_m^{1/2}\big(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,m})\big)\right\|$$

defined for any $k > m$ and $\ell = m, \dots, k$, where

$$\widehat{\Sigma}_m = \left(\frac{1}{m}\sum_{t=1}^{m}\frac{1}{\widehat{f}_\theta^t}\left(\frac{\partial}{\partial\theta}\widehat{f}_\theta^t\right)\left(\frac{\partial}{\partial\theta}\widehat{f}_\theta^t\right)'\right)\Bigg|_{\theta=\widehat{\theta}(T_{1,m})}$$

is a consistent estimator of Σ (see Sect. 3 and also Doukhan and Kengne (2015)). Hence, $\widehat{\Sigma}_m$ is asymptotically positive definite, and the detector $\widehat{C}_{k,\ell}$ is well defined for m large enough.

To avoid some distortion in the computation of $\widehat{\theta}(T_{\ell,k})$ (when ℓ is close to k), we introduce a sequence of integer numbers $(v_m)_{m\in\mathbb{N}}$ with $v_m << m$ and compute $\widehat{C}_{k,\ell}$ for $\ell \in \{m - v_m, m - v_m + 1, \ldots, k - v_m\}$. Thus, for any $k > m$ denote

$$\Pi_{m,k} := \{m - v_m, m - v_m + 1, \ldots, k - v_m\}.$$

For technical consideration, we assume in this section that,

$$v_m \to \infty \quad \text{and} \quad v_m/\sqrt{m} \to 0 \quad (m \to \infty).$$

Let us stress that, if v_m is too small, the convergence of the numerical algorithm used to $\widehat{\theta}(T_{\ell,k})$ (when ℓ is close to k) is not ensured, which can introduce high distortion in the empirical level of the procedure. And conversely, if v_m is too large, under the alternative, one will wait longer in the monitoring scheme to get $k^* \in \Pi_{m,k}$ and so that the change-point is detected; that is, a procedure with high distortion in the empirical power if $\mathcal{T}$ (see below) is not too large, and large detection delay. Such sequence does not increase too fast in order to keep the accuracy of the procedure. We also refer to Remark 1 in Kengne (2012). The usual choice of such sequence is $v_m = (\log m)^\delta$, with $3/2 \leq \delta \leq 3$ (see Doukhan and Kengne 2015; Diop and Kengne 2017; Bardet and Kengne 2014; Kengne 2012). We have evaluated the procedure with $v_m = [(\log m)^2], [(\log m)^{2.25}], [(\log m)^{2.5}], [(\log m)^{2.75}], [(\log m)^3]$ for the INGARCH(1, 1) models, and have noticed that the choice $v_m = [(\log m)^{2.25}]$ displayed a good trade-off between the distortion in the empirical level and power.

Note that, for any $\ell \in \Pi_{m,k}$ both $\widehat{\theta}(T_{\ell,k})$ and $\widehat{\theta}(T_{1,m})$ are estimators of θ_0^* if change does not occur at time $k > m$, they are asymptotically close and the detector $\widehat{C}_{k,\ell}$ is not too large under H_0.

Let $\mathcal{T} > 1$ ($\mathcal{T}$ can be infinite). The monitoring scheme rejects H_0 at the first time k between m and $[\mathcal{T}m]$ where there exists $\ell \in \Pi_{m,k}$ such that $\widehat{C}_{k,\ell} > c$ for a suitable constant $c > 0$ to be computed, where $[x]$ denotes the integer part of x. The procedure is called *closed-end* (resp. *open-end*) method when $\mathcal{T} < \infty$ (resp. $\mathcal{T} = \infty$). To be more general, we will use a function $b : (0, \infty) \mapsto (0, \infty)$, called a boundary function satisfying:

Assumption B $b : (0, \infty) \mapsto (0, \infty)$ is a non-increasing and continuous function such that $\inf_{0<t<\infty} b(t) > 0$.

Then, the monitoring scheme stops at the first time k (with $m < k \leq [\mathcal{T}m] + 1$) such that $\widehat{C}_{k,\ell} > b((k - \ell)/m)$ for some $\ell \in \Pi_{m,k}$. Therefore, define the stopping time:

$$\begin{aligned}\tau(m) &:= \inf\Big\{m < k < [\mathcal{T}m] + 1 \big/ \exists \ell \in \Pi_{m,k},\ \widehat{C}_{k,\ell} > b((k - \ell)/m)\Big\} \\ &= \inf\left\{m < k < [\mathcal{T}m] + 1 \Big/ \max_{\ell \in \Pi_{m,k}} \frac{\widehat{C}_{k,\ell}}{b((k - \ell)/m)} > 1\right\},\end{aligned}$$

with the convention that $\inf\{\emptyset\} = \infty$. Therefore, we have

$$P\{\tau(m) < \infty\} = P\left\{ \max_{\ell \in \Pi_{m,k}} \frac{\widehat{C}_{k,\ell}}{b((k-\ell)/m)} > 1 \text{ for some } k \text{ between } m \text{ and } [\mathcal{T}m]+1 \right\}$$
$$= P\left\{ \sup_{m<k<[\mathcal{T}m]+1} \max_{\ell \in \Pi_{m,k}} \frac{\widehat{C}_{k,\ell}}{b((k-\ell)/m)} > 1 \right\}. \tag{20}$$

The challenge is to choose a suitable boundary function $b(\cdot)$ such that for some given $\alpha \in (0,1)$,

$$\lim_{m\to\infty} P_{H_0}\{\tau(m) < \infty\} = \alpha$$

and

$$\lim_{m\to\infty} P_{H_1}\{\tau(m) < \infty\} = 1,$$

where the hypothesis H_0 and H_1 are formulated above.

In the case of a boundary function, that is $b \equiv c$ for some $c > 0$, it holds from (20) that $P\{\tau(m) < \infty\} = P\big\{ \sup_{m<k<[\mathcal{T}m]+1} \max_{\ell \in \Pi_{m,k}} \widehat{C}_{k,\ell} > c \big\}$. Then, the aim is to compute a threshold $c = c_\alpha$ satisfying $\lim_{m\to\infty} P_{H_0}\{\tau(m) < \infty\} = \alpha$ and $\lim_{m\to\infty} P_{H_1}\{\tau(m) < \infty\} = 1$. If change is detected under H_1 *i.e.* $\tau(m) < \infty$ and $\tau(m) > k^*$, then the detection delay is defined by

$$\widehat{d}_m = \tau(m) - k^*. \tag{21}$$

$\widehat{d}_m$ is used to assess the efficiency of the procedure to early detect changes in the model. The smaller is the detection delay, the better is the efficiency under the alternative.

4.1 Asymptotic under the null hypothesis

Under H_0, all the observations are generated from the model (6) according to the parameter θ_0^*. The following theorem displays the asymptotic behavior under the null hypothesis of the detector $\widehat{C}_{m,k}$ for the open and closed-end procedure.

Theorem 3 *Assume that* $D(\Theta)$, $\mathrm{Id}(\Theta)$, $Var(\Theta)$ *and* $A_i(\Theta) i = 0, 1, 2$ *hold with*

$$\alpha_j^{(i)} = O(j^{-\gamma}), \quad \text{for some } \gamma > 3/2.$$

Under H_0 *with* $\theta_0^* \in \mathring{\Theta}$, *for the open-end* ($\mathcal{T} = \infty$) *and closed-end* ($\mathcal{T} < \infty$) *procedure it holds that*

$$\lim_{m\to\infty} P\{\tau(m) < \infty\} = P\left\{ \sup_{1<t\leq\mathcal{T}} \sup_{0<s<t-1} \frac{\|W_d(s) - sW_d(1))\|}{t\, b(s)} > 1 \right\}, \tag{22}$$

where W_d *is a* d *-dimensional standard Brownian motion.*

Assume that $b(s) = cb_0(s)$ for some function b_0 satisfying the assumption **B**, with $c > 0$. Thus, at a nominal level $\alpha \in (0, 1)$, the monitoring procedure stops and rejects H_0 at the first time k (with $1 < k \leq [\mathcal{T}m] + 1$) such that

$$\max_{\ell \in \Pi_{m,k}} \frac{\widehat{C}_{k,\ell}}{b_0((k-\ell)/m)} > c_{\alpha,d,\mathcal{T}}$$

where $c_{\alpha,d,\mathcal{T}}$ is the $(1-\alpha)$-quantile of the distribution of $\sup_{1<t\leq\mathcal{T}} \sup_{0<s<t-1} \dfrac{\|W_d(s) - sW_d(1))\|}{t\, b_0(s)}$.

In Sect. 5, for numerical convenience, we will use the simplest boundary function $b(\cdot) = c$ where c is a positive constant. In this case, it follows directly from Theorem 3 that

$$\lim_{m\to\infty} P\{\tau(m) < \infty\} = P\{U_{d,\mathcal{T}} > c\}$$

where

$$U_{d,\mathcal{T}} = \sup_{1<t\leq\mathcal{T}} \sup_{0<s<t-1} \frac{1}{t} \|W_d(s) - sW_d(1))\|. \tag{23}$$

The $(1-\alpha)$-quantile $c_{\alpha,d,\mathcal{T}}$ of the distribution of $U_{d,\mathcal{T}}$ can be computed through Monte-Carlo simulations. With $\alpha = 0.05$, we have obtained $c_{\alpha,3,1.5} = 1.679$ and $c_{\alpha,3,2} = 1.803$.

4.2 Asymptotic under the alternative

Under the alternative, a change occurs at time $k^* > m$ and contrary to some recent works (for instance: Franke et al. 2012; Doukhan and Kengne 2015; Kengne 2015; Kirch and Kamgaing 2015; Diop and Kengne 2017; Kirch and Weber 2018, ...), we do not set any additional assumption on the process after the change-point.

Many recent works impose stationarity after the change-point. This assumption is too strong for autoregressive process; note that, in model (5) with $t > k^*$,

$$\lambda_t = f_{\theta_1^*}(Y_{t-1}, Y_{t-2}, \ldots)$$

depends on θ_1^* and it is contaminated by observations which depend on θ_0^*. This shows that, stationarity assumption on the observations after the change-point is quite questionable and nonstationary approach seems to be suitable. The proof of the following theorem is heavily based on the result of Theorem 1. The results below show that the proposed monitoring procedure is consistent under the alternative for both the open-end and the closed-end methods.

Theorem 4 *Assume that* $D(\Theta)$, $\mathrm{Id}(\Theta)$, $Var(\Theta)$ *and* $A_i(\Theta)$ $i = 0, 1, 2$ *hold with*

$$\alpha_j^{(i)} = O(j^{-\gamma}), \quad \text{for some } \gamma > 3/2.$$

Under the alternative H_1, if $\theta_0^*, \theta_1^* \in \mathring{\Theta}$ and there exists $\mathcal{T}^* \in (1, \mathcal{T})$ such that $k^* = [\mathcal{T}^* m]$, for the open-end ($\mathcal{T} = \infty$) and closed-end ($\mathcal{T} < \infty$) procedure, then for $k_m = k^* + m^\delta$ with $\delta \in (1/2, 1)$, it holds that

$$\max_{\ell \in \Pi_{m,k_m}} \frac{\widehat{C}_{k_m,\ell}}{b((k_m - \ell)/m)} \xrightarrow[m\to\infty]{\text{a.s.}} \infty. \tag{24}$$

The Corollary 1 follows immediately from Theorem 4.

Corollary 1 *Under the assumptions of Theorem 4,*

$$\lim_{m\to\infty} P\{\tau(m) < \infty\} = 1.$$

Therefore, it follows from Theorem 4 that with probability one, the change is asymptotically detected both for open-end and closed-end (when $\mathcal{T}^* < \mathcal{T}$) procedures and the detection delay $\widehat{d}_n$ can be bounded by $\mathcal{O}_P(m^{1/2+\varepsilon})$ for any $\varepsilon > 0$ (or even by $\mathcal{O}_P\big(\sqrt{m}(\log m)^a\big)$ with $a > 0$ using the same kind of proof).

5 Some numerical results

In this section, we conduct a small simulation study and a real data example in order to display some empirical performances of the proposed sequential change-point procedure. We focus on the closed-end procedure with $\mathcal{T} = 1.5, 2$; that is, the historical available data are $X_1, \dots, X_m$ and the monitoring periods considered are $\{m+1, \dots, 1.5m\}$ and $\{m+1, \dots, 2m\}$. In the sequel, the detector of the sequential procedure is computed with $v_m = [(\log m)^{2.25}]$ (see Sect. 4).

5.1 Sequential change-point detection in Poisson INGARCH

We consider a Poisson INGARCH(1, 1)

$$Y_t | \mathcal{F}_{t-1} \sim \text{ Poisson } (\lambda_t) \text{ with } \lambda_t = \alpha_0^* + \alpha_1^* \lambda_{t-1} + \beta_1^* Y_{t-1} \tag{25}$$

where $\theta_0^* = (\alpha_0^*, \alpha_1^*, \beta_1^*)$ denote the parameter of the model. For any $k > m$, denote $\widehat{C}_k = \max_{\ell \in \Pi_{m,k}} \widehat{C}_{k,\ell}$. For $m = 1000$, Fig. 1 displays the statistics $(\widehat{C}_k)_{1001 \le k \le 1500}$ in a scenario without change (a) and a scenario with a change-point at $k^* = 1.25m = 1250$ (b). Figure 1a shows that the detector $\widehat{C}_k$ is under the horizontal line that defines the critical region of the test; whereas in Fig. 1b, the detector is under the horizontal line before a change occurs, and increases with a high rate until exceeds the critical value after the change-point. As pointed out by Bardet and Kengne (2014), such growth rate over a long period indicates that something is happening in the model.

Note that, a classical fluctuation test for sequential change-point detection is based on the detector (see for instance Chu et al. 1996; Leisch et al. 2000; Na et al. 2011),

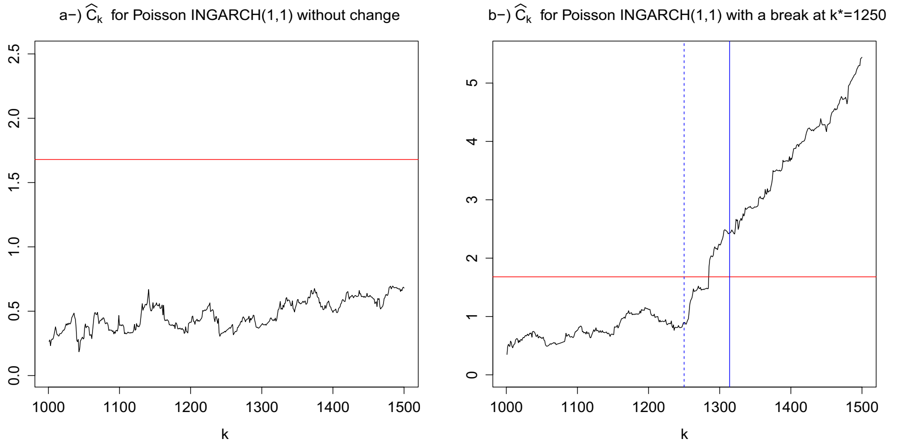

Fig. 1 A realization of the detector $(\widehat{C}_k)_{1001\le k\le 1500}$ for a Poisson INGARCH(1, 1) with $m = 1000$. **a** The parameter $\theta_0^* = (1, 0.2, 0.15)$ is constant; **b** the parameter $\theta_0^* = (1, 0.2, 0.15)$ changes to $\theta_1^* = (1, 0.2, 0.5)$ at $k^* = 1250$. The horizontal solid line represents the limit of the critical region, the vertical dotted line indicates where the change occurs and the vertical solid line indicates the time where the sequential procedure detects a break in the observations

$$\widehat{D}_k := \sqrt{m}\left\| \widehat{\Sigma}_m^{1/2}\left(\widehat{\theta}(T_{1,k}) - \widehat{\theta}(T_{1,m})\right)\right\| \quad \text{for any } k > m.$$

With this statistic and with the constant boundary function $b \equiv c > 0$, the procedure stops (and rejects H_0) at the first time k (with $m < k \le [\mathcal{T}m] + 1$) such that $\widehat{D}_k > c$. The associated stopping time is,

$$\tilde{\tau}(m) := \inf\left\{ m < k < [\mathcal{T}m] + 1,\ \widehat{D}_k > c \right\}.$$

Therefore,

$$P\{\tilde{\tau}(m) < \infty\} = P\left\{ \sup_{m<k<[\mathcal{T}m]+1} \widehat{D}_k > c \right\}.$$

This statistic has been applied by Kengne (2015) to Poisson autoregressive models (5) but without asymptotic study under H_1. Under H_0, they proved that,

$$\lim_{m\to\infty} P\{\tilde{\tau}(m) < \infty\} = P\left\{ \sup_{1<t\le\mathcal{T}} \frac{\|W_d(t) - tW_d(1))\|}{t} > c \right\}.$$

Thus, at a nominal level $\alpha \in (0, 1)$, take $c = \tilde{c}_{\alpha,d,\mathcal{T}}$, the $(1-\alpha)$-quantile of the distribution $\sup_{1<t\le\mathcal{T}}\left(\|W_d(t) - tW_d(1))\|/t\right)$. From the procedure described on page 105 in Kengne (2015), we get with $\alpha = 0.05$, $\tilde{c}_{\alpha,3,1.5} = 1.740$ and $\tilde{c}_{\alpha,3,2} = 2.130$.

Table 1 indicates the empirical levels and powers obtained after 500 replications for the procedures based on the statistics $\widehat{C}_k$ and $\widehat{D}_k$, for $m = 200, 500, 1000$.

Some elementary statistics of the empirical detection delay (defined at (21)) for the first two scenarios are summarized in Table 2. In the third scenario (with a large discrepancy between the model before and after the change-point), both the procedures based on $\widehat{C}_k$ and $\widehat{D}_k$ work very well and provide very good detection delay.

The results of Table 1 display some distortion in the empirical levels when $m = 200, 500$. But the empirical level decreases as m increases and for the three scenarios, it is close to the nominal level when $m = 1000$ for both the detectors $\widehat{C}_k$ and

Table 1 Empirical levels and powers for sequential change-point detection in Poisson INGARCH(1, 1) model, with procedures based on the statistics $\widehat{C}_k$ and $\widehat{D}_k$

		Detector	$m = 200$	$m = 500$	$m = 1000$
Empirical levels					
$\theta_0^* = (1, 0.2, 0.15)$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	0.078	0.062	0.052
		$\widehat{D}_k$	0.074	0.060	0.050
	$\mathcal{T} = 2$	$\widehat{C}_k$	0.082	0.064	0.054
		$\widehat{D}_k$	0.076	0.060	0.052
$\theta_0^* = (0.75, 0.5, 0.3)$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	0.080	0.068	0.054
		$\widehat{D}_k$	0.074	0.056	0.048
	$\mathcal{T} = 2$	$\widehat{C}_k$	0.102	0.072	0.058
		$\widehat{D}_k$	0.086	0.054	0.044
$\theta_0^* = (2.5, 0, 0.35)$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	0.076	0.060	0.042
		$\widehat{D}_k$	0.056	0.038	0.046
	$\mathcal{T} = 2$	$\widehat{C}_k$	0.088	0.066	0.058
		$\widehat{D}_k$	0.052	0.054	0.042
Empirical powers					
$\theta_0^* = (1, 0.2, 0.15)$; $\theta_1^* = (1, 0.2, 0.5)$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	0.782	0.958	0.994
		$\widehat{D}_k$	0.608	0.778	0.894
	$\mathcal{T} = 2$	$\widehat{C}_k$	0.934	0.980	1
		$\widehat{D}_k$	0.710	0.834	0.916
$\theta_0^* = (0.75, 0.5, 0.3)$; $\theta_1^* = (0.25, 0.5, 0.3)$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	0.844	0.966	1
		$\widehat{D}_k$	0.632	0.786	0.982
	$\mathcal{T} = 2$	$\widehat{C}_k$	0.900	0.988	1
		$\widehat{D}_k$	0.616	0.830	0.908
$\theta_0^* = (2.5, 0, 0.35)$; $\theta_1^* = (4.5, 0.05, 0.6)$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	0.968	1	1
		$\widehat{D}_k$	0.854	1	1
	$\mathcal{T} = 2$	$\widehat{C}_k$	0.988	1	1
		$\widehat{D}_k$	0.850	0.966	1

The empirical levels are computed when $\theta_0^* = (1, 0.2, 0.15), (0.75, 0.5, 0.3), (2.5, 0, 0.35)$ is constant (under H_0) and the empirical powers when $\theta_0^* = (1, 0.2, 0.15), (0.75, 0.5, 0.3), (2.5, 0, 0, 0.35)$ changes respectively to $\theta_1^* = (1, 0.2, 0.5), (0.25, 0.5, 0.3), (4.5, 0.05, 0.6)$ (under the alternative) at $k^* = 1.25m$

Table 2 Elementary statistics of the empirical detection delay for sequential change-point detection in a Poisson INGARCH(1, 1)

$\widehat{d}_m$		Detector	Mean	SD	Min	Q_1	Med	Q_3	Max
$\theta_0^* = (1, 0.2, 0.15)$; $\theta_1^* = (1, 0.2, 0.5)$									
$m = 200$; $k^* = 250$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	28.77	9.78	8	24	31	38	50
		$\widehat{D}_k$	29.67	11.88	7	18	28	39	50
	$\mathcal{T} = 2$	$\widehat{C}_k$	39.80	20.09	7	29	35	50	113
		$\widehat{D}_k$	49.94	29.04	7	30	45	63	142
$m = 500$; $k^* = 625$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	48.41	19.04	6	35	48	61	91
		$\widehat{D}_k$	48.19	26.25	4	26	42	55	109
	$\mathcal{T} = 2$	$\widehat{C}_k$	53.48	27.68	5	35	53	66	135
		$\widehat{D}_k$	73.55	65.23	5	38	50	81	360
$m = 1000$; $k^* = 1250$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	69.51	31.65	12	49	69	81	175
		$\widehat{D}_k$	61.29	27.78	16	36	55	69	141
	$\mathcal{T} = 2$	$\widehat{C}_k$	78.33	31.71	28	60	75	90	208
		$\widehat{D}_k$	90.65	84.67	22	49	63	96	576
$\theta_0^* = (0.75, 0.5, 0.3)$; $\theta_1^* = (0.25, 0.5, 0.3)$									
$m = 200$; $k^* = 250$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	25.06	11.73	3	19	27	34	50
		$\widehat{D}_k$	24.07	14.28	4	15	22	35	49
	$\mathcal{T} = 2$	$\widehat{C}_k$	25.65	14.98	4	20	25	29	95
		$\widehat{D}_k$	34.00	28.05	6	13	21	35	93
$m = 500$; $k^* = 625$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	41.76	20.43	5	32	39	49	102
		$\widehat{D}_k$	45.69	21.07	8	28	41	57	108
	$\mathcal{T} = 2$	$\widehat{C}_k$	46.73	29.22	3	35	40	49	244
		$\widehat{D}_k$	54.78	25.82	17	33	47	66	149
$m = 1000$; $k^* = 1250$	$\mathcal{T} = 1.5$	$\widehat{C}_k$	58.42	29.05	3	46	55	74	139
		$\widehat{D}_k$	56.81	20.63	3	38	58	67	98
	$\mathcal{T} = 2$	$\widehat{C}_k$	62.60	27.79	8	47	56	69	174
		$\widehat{D}_k$	76.41	31.17	26	50	73	104	158

$\widehat{D}_k$. Also, the empirical power increases with m and approaches one when $m = 1000$ for the procedure based on $\widehat{C}_k$, which outperforms the detector $\widehat{D}_k$; and the empirical power is overall a bit more accurate when $\mathcal{T} = 2$ than $\mathcal{T} = 1.5$.

These findings are consistent with Theorem 3 and Corollary 1.

In Table 2, for example, when $m = 200$ with a change occurred at the time $k^* = 250$ in the first scenario, this break is detected on average after a delay of 29, 30 respectively for the procedure based on $\widehat{C}_k$ and $\widehat{D}_k$ when $\mathcal{T} = 1.5$. It appears that, the detection delay when $\mathcal{T} = 2$ is slightly larger than the case when $\mathcal{T} = 1.5$; this is not surprising since the monitoring period with $\mathcal{T} = 2$ is greater. Moreover, one can notice that the detection delay displayed by the detector $\widehat{C}_k$ is overall more accurate than that based on $\widehat{D}_k$. We can also see that, for two historical sample sizes m_1 and m_2 with $m_1 < m_2$, the sequence $\widehat{d}_{m_2} - \sqrt{m_2/m_1}\widehat{d}_{m_1}$ overall decreases when m_1 and m_2 increases and it is on average, close or less than 0 when $m_1 = 500$ and $m_2 = 1000$. This is in accordance with Theorem 4 where $\widehat{d}_m$ can be bounded by $\mathcal{O}_P\left(m^{1/2+\epsilon}\right)$ for any $\epsilon > 0$.

5.2 Real data example

We consider the daily number of trades in the stock of Technofirst listed in the NYSE Euronext group. These data have been analyzed by Ahmad and Francq (2016) with the Poisson Quasi-maximum Likelihood Estimator (PQMLE); they have applied a test of nullity of the coefficient and have concluded that the INGARCH(1, 3) model is more appropriate [in comparison to INGARCH(1, 2) and INGARCH(1, 1) representation]. Diop and Kengne (2021) have applied a multiple change-point procedure with an INGARCH(1, 1) representation based on the PQMLE. We consider the data from 04 January 2010 to 05 September 2011 (see Fig. 2); there are 310 observations.

For the data from $t = 1$ to $t = 230$, Diop and Kengne (2021) have showed that the INARCH(1) representation is more appropriate and the INGARCH(1, 1) representation has been used for $t > 230$. So, we applied the Poisson INGARCH(1, 1) model and considered the observations from $t = 1$ to $t = 207$ as the historical data.

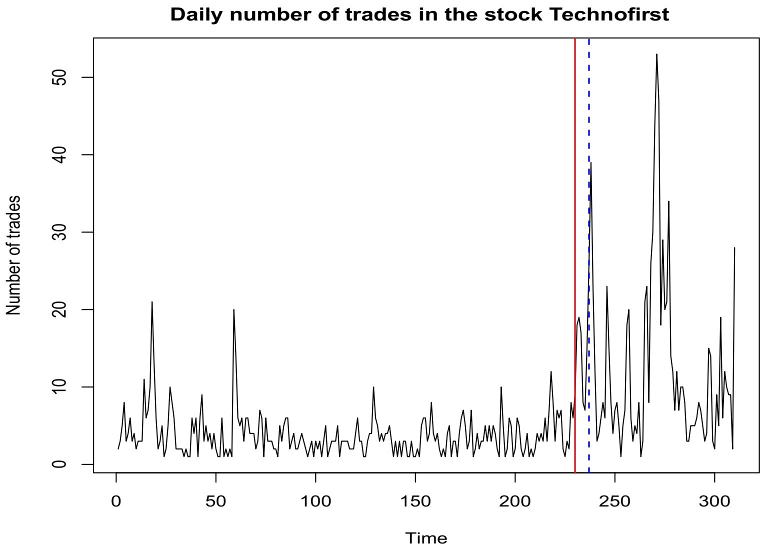

Fig. 2 Daily number of trades in the stock of Technofirst from 04 January 2010 to 05 September 2011. The solid line represents the break that has been detected by Diop and Kengne (2021) from a retrospective procedure. The dotted line indicates the stopping time of the sequential procedure proposed

We carry out the sequential procedure in the closed-end setting with $\mathcal{T} = 1.5$; so, $[\mathcal{T} \times m] = 310$. Therefore, the monitoring starts at the time $t = 208$. The estimation of the parameter computed on the historical data is $\widehat{\theta}_0 = (2.43, 2 \times 10^{-8}, 0.35)$.

Figure 3 displays the realizations of the detector $\widehat{C}_k = \max_{\ell \in \Pi_{n,k}} \widehat{C}_{k,\ell}$, with $k = 208, \ldots, 310$. One can see that the sequential procedure stops at time $t = 237$. In terms of the detection delay, it appears that the procedure works well for this real data example, in the sense that the sequential procedure stops 7 days after the break time detected by Diop and Kengne (2021).

6 Concluding remarks

This work addresses the question of inference for nonstationary time series of counts. After a time k^*, the process is a nonstationary Poisson autoregressive model with the conditional mean that depends on a parameter θ^*. We carry out an approximation study between this process and the stationary regime which allows us to establish that the MLE of θ^* computed with the nonstationary observations is consistent and asymptotically normal. We thus provide a detailed proof of an issue that has been addressed by Doukhan and Kengne (2015) (see Remark 4.1). These results are very useful for both the retrospective and the sequential change-point problem. We perform an application to sequential change-point detection and propose a consistent procedure in which the detection delay can be bounded by $\mathcal{O}_P\big(m^{1/2+\epsilon}\big)$ for any $\epsilon > 0$. Empirical studies show that the procedure works well for simulated and

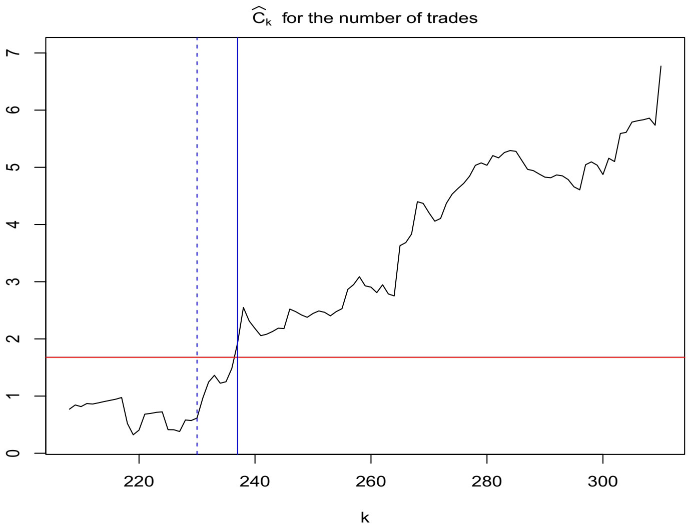

Fig. 3 Realizations of the statistics $(\widehat{C}_k)_{208 \leq k \leq 310}$ for the daily number of trades in the stock of Technofirst from 04 January 2010 to 05 September 2011; the historical data considered are the first 207 observations. The horizontal solid line represents the limit of the critical region, the vertical dotted line represents the break that has been detected by using the retrospective procedure of Diop and Kengne (2021) and the vertical solid line indicates the stopping time of the sequential procedure

real data example with satisfactory detection delay. An extension of this work is the study of the inference for nonstationary models where the conditional distribution is different from Poisson, and could be for instance negative binomial, binary, etc.

7 Proofs of main results

Let $(\psi_n)_n$ and $(r_m)_m$ be sequences of random variables or vectors. Throughout this section, we use the notation $\psi_m = o_P(r_m)$ to mean: for all $\varepsilon > 0$, $P(\|\psi_m\| \geq \varepsilon \|r_m\|) \to 0$ as $m \to \infty$. Write $\psi_m = O_P(r_m)$ to mean: for all $\varepsilon > 0$, there exists $C > 0$ such that $P(\|\psi_m\| \geq C\|r_m\|) \leq \varepsilon$ for n large enough.

Proof of Proposition 1 We will prove that, for all $r \in \mathbb{N}$, there exists $C_r > 0$ such that

$$\mathbb{E}Y_t^r \leq C_r, \quad \forall t \in \mathbb{Z}. \tag{26}$$

Recall that for all $\ell \geq 1$,

$$Y_{k^*+\ell} | \mathcal{F}_{k^*+\ell-1} \sim \text{ Poisson } (\lambda_{k^*+\ell}) \text{ with } \lambda_{k^*+\ell} = f_{\theta^*}(Y_{k^*+\ell-1}, Y_{k^*+\ell-2}, \ldots) = f_{\theta^*}^{k^*+\ell}.$$

According to the assumption $\mathbf{A}_0(\Theta)$, we have for all $\ell \geq 1$,

$$f_{\theta^*}^{k^*+\ell} \leq |f_{\theta^*}^{k^*+\ell} - f_{\theta^*}(0)| + f_{\theta^*}(0) \leq \sum_{j\geq 1} \alpha_j^{(0)} Y_{k^*+\ell-j} + f_{\theta^*}(0). \tag{27}$$

In the sequel, we set $\alpha^{(0)} = \sum_{j\geq 1} \alpha_j^{(0)}$. If (26) holds for some $r \in \mathbb{N}$, then we get from the Jensen's inequality,

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{j\geq 1} \alpha_j^{(0)} Y_{k^*+\ell-j}\right)^r\right] &= (\alpha^{(0)})^r \mathbb{E}\left[\left(\sum_{j\geq 1} \frac{\alpha_j^{(0)}}{\alpha^{(0)}} Y_{k^*+\ell-j}\right)^r\right] \\ &\leq (\alpha^{(0)})^{r-1} \sum_{j\geq 1} \alpha_j^{(0)} \mathbb{E}Y_{k^*+\ell-j}^r \leq (\alpha^{(0)})^r C_r.\end{aligned} \tag{28}$$

Moreover, under (26), for some $r \in \mathbb{N}$ since $Y_t^s \leq Y_t^r$ a.s. for any $s \leq r$, we have $\mathbb{E}Y_t^s \leq C_r$ for $s \leq r$. Thus, we can get $C_s \leq C_r$ for any $s \leq r$. Therefore, for all $\ell \geq 1$,

$$
\begin{aligned}
\mathbb{E}[(f_{\theta^*}^{k^*+\ell})^r] &\leq \mathbb{E}\left[\left(\sum_{j\geq 1}\alpha_j^{(0)}Y_{k^*+\ell-j}+f_{\theta^*}(0)\right)^r\right] \\
&\leq \sum_{s=0}^{r}\binom{r}{s}\left(f_{\theta_1^*}(0)\right)^{r-s}\mathbb{E}\left[\left(\sum_{j\geq 1}\alpha_j^{(0)}Y_{k^*+\ell-j}\right)^s\right] \\
&\leq \sum_{s=0}^{r}\binom{r}{s}\left(f_{\theta^*}(0)\right)^{r-s}(\alpha^{(0)})^s C_s \\
&\leq \sum_{s=0}^{r}C_r\binom{r}{s}\left(f_{\theta^*}(0)\right)^{r-s}(\alpha^{(0)})^s \leq C_r\left(\alpha^{(0)}+f_{\theta^*}(0)\right)^r \leq C_{r,f}
\end{aligned}
\tag{29}
$$

with $C_{r,f}=C_r\left(\alpha^{(0)}+f_{\theta^*}(0)\right)^r$.

Let us show by induction that for all $r\in\mathbb{N}$, there exists $C_r>0$ such that (26) holds. For $r=1$, if C_1 exists, we will have $\mathbb{E}Y_t\leq C_1$ for all $t\leq k^*$; and according to (27), for all $\ell\geq 1$,

$$\mathbb{E}Y_{k^*+\ell}=\mathbb{E}f_{\theta^*}^{k^*+\ell}\leq\sum_{j\geq 1}\alpha_j^{(0)}\mathbb{E}Y_{k^*+\ell-j}+f_{\theta^*}(0)\leq\alpha^{(0)}C_1+f_{\theta^*}(0).$$

Hence, (26) holds with $C_1=\max(C_{1,0},\frac{1}{1-\alpha^{(0)}}f_{\theta^*}(0))$. Indeed, from the assumption of the proposition, $\mathbb{E}Y_t\leq C_{1,0}$ for $t\leq k^*$, thus $\mathbb{E}Y_t\leq C_1$ for $t\leq k^*$. Moreover, from the above inequality, $\mathbb{E}Y_{k^*+1}\leq\sum_{j\geq 1}\alpha_j^{(0)}\mathbb{E}Y_{k^*+1-j}+f_{\theta^*}(0)\leq\alpha^{(0)}C_{1,0}+f_{\theta^*}(0)$. If $C_{1,0}\leq\frac{1}{1-\alpha^{(0)}}f_{\theta^*}(0)$, then

$$\mathbb{E}Y_{k^*+1}\leq\alpha^{(0)}\frac{1}{1-\alpha^{(0)}}f_{\theta^*}(0)+f_{\theta^*}(0)=\frac{1}{1-\alpha^{(0)}}f_{\theta^*}(0)\leq C_1;$$

else, if $\frac{1}{1-\alpha^{(0)}}f_{\theta^*}(0)\leq C_{1,0}$, then

$$\mathbb{E}Y_{k^*+1}\leq\alpha^{(0)}C_{1,0}+(1-\alpha^{(0)})C_{1,0}=C_{1,0}\leq C_1.$$

Thus, in both cases, $\mathbb{E}Y_{k^*+1}\leq C_1$. Also,

$$
\begin{aligned}
EY_{k^*+2} &\leq \sum_{j\geq 1}\alpha_j^{(0)}\mathbb{E}Y_{k^*+2-j}+f_{\theta^*}(0)=\sum_{j\geq 2}\alpha_j^{(0)}\mathbb{E}Y_{k^*+2-j}+\alpha_1^{(0)}\mathbb{E}Y_{k^*+1}+f_{\theta^*}(0) \\
&\leq \sum_{j\geq 2}\alpha_j^{(0)}C_{1,0}+\alpha_1^{(0)}C_1+f_{\theta^*}(0)\leq\sum_{j\geq 2}\alpha_j^{(0)}C_1+\alpha_1^{(0)}C_1+f_{\theta^*}(0) \\
&= \alpha^{(0)}C_1+f_{\theta^*}(0)\leq C_1;
\end{aligned}
$$

where the last inequality is obtained by considering the two cases $C_{1,0}\leq\frac{1}{1-\alpha^{(0)}}f_{\theta^*}(0)$ and $\frac{1}{1-\alpha^{(0)}}f_{\theta^*}(0)\leq C_{1,0}$ as above. Similarly, we have $EY_{k^*+\ell}\leq C_1$ for all $\ell\geq 1$. In addition to $\mathbb{E}Y_t\leq C_{1,0}\leq C_1$ for $t\leq k^*$, we get $\mathbb{E}Y_t\leq C_1$ for all $t\in\mathbb{Z}$.

Assume (26) holds until $r\in\mathbb{N}$. According to Lemma 1 of Ferland et al. (2006) (see also Lemma A.1. of Doukhan et al. 2012) and (29), for all $\ell\geq 1$,

$$
\begin{aligned}
\mathbb{E}Y_{k^*+\ell}^{r+1} &= \mathbb{E}\big(\mathbb{E}(Y_{k^*+\ell}^{r+1}|\mathcal{F}_{k^*+\ell-1})\big) = \sum_{s=0}^{r+1}\left\{ \begin{array}{c} r+1 \\ s \end{array} \right\}\mathbb{E}[(f_{\theta^*}^{k^*+\ell})^s] \\
&= \mathbb{E}[(f_{\theta^*}^{k^*+\ell})^{r+1}] + \sum_{s=0}^{r}\left\{ \begin{array}{c} r+1 \\ s \end{array} \right\}\mathbb{E}[(f_{\theta^*}^{k^*+\ell})^s] \qquad (30) \\
&\leq \mathbb{E}[(f_{\theta^*}^{k^*+\ell})^{r+1}] + \sum_{s=0}^{r}\left\{ \begin{array}{c} r+1 \\ s \end{array} \right\}C_{s,f}
\end{aligned}
$$

where for all $n, k \in \mathbb{N}_0$, $\left\{ \begin{array}{c} n \\ k \end{array} \right\}$ denotes the Stirling numbers of the second kind that satisfies the recurrence $\left\{ \begin{array}{c} n \\ k \end{array} \right\} = \left\{ \begin{array}{c} n-1 \\ k-1 \end{array} \right\} + k\left\{ \begin{array}{c} n-1 \\ k \end{array} \right\}$ with $\left\{ \begin{array}{c} n \\ n \end{array} \right\} = 1\ \forall n \in \mathbb{N}_0$, $\left\{ \begin{array}{c} n \\ 0 \end{array} \right\} = 0\ \forall n \in \mathbb{N}$ and $\left\{ \begin{array}{c} n \\ k \end{array} \right\} = 0$ if $k > n$. Hence, if C_{r+1} exists, it must satisfy $C_{r+1} \geq \mathbb{E}Y_t^{r+1}$ for all $t \leq k^*$; and according to (27) and (28), we have

$$
\begin{aligned}
\mathbb{E}[(f_{\theta^*}^{k^*+\ell})^{r+1}] &\leq \mathbb{E}\left[\left(\sum_{j\geq 1}\alpha_j^{(0)}Y_{k^*+\ell-j} + f_{\theta^*}(0)\right)^{r+1}\right] \\
&\leq \sum_{s=0}^{r+1}\binom{r+1}{s}\big(f_{\theta^*}(0)\big)^{r-s+1}\mathbb{E}\left[\left(\sum_{j\geq 1}\alpha_j^{(0)}Y_{k^*+\ell-j}\right)^s\right] \\
&\leq \mathbb{E}\left[\left(\sum_{j\geq 1}\alpha_j^{(0)}Y_{k^*+\ell-j}\right)^{r+1}\right] \\
&\quad + \sum_{s=0}^{r}\binom{r+1}{s}\big(f_{\theta^*}(0)\big)^{r-s+1}\mathbb{E}\left[\left(\sum_{j\geq 1}\alpha_j^{(0)}Y_{k^*+\ell-j}\right)^s\right] \\
&\leq (\alpha^{(0)})^{r+1}C_{r+1} + \sum_{s=0}^{r}\binom{r+1}{s}\big(f_{\theta^*}(0)\big)^{r-s+1}(\alpha^{(0)})^s C_s \\
&\leq (\alpha^{(0)})^{r+1}C_{r+1} + C_r\sum_{s=0}^{r}\binom{r+1}{s}\big(f_{\theta^*}(0)\big)^{r-s+1}(\alpha^{(0)})^s \\
&\leq (\alpha^{(0)})^{r+1}C_{r+1} + C_r\big((\alpha^{(0)} + f_{\theta^*}(0))^{r+1} - (\alpha^{(0)})^{r+1}\big).
\end{aligned}
$$

Hence, (30) gives

$$
\mathbb{E}Y_{k^*+\ell}^{r+1} \leq (\alpha^{(0)})^{r+1}C_{r+1} + C_r\big((\alpha^{(0)} + f_{\theta^*}(0))^{r+1} - (\alpha^{(0)})^{r+1}\big) + \sum_{s=0}^{r}\left\{ \begin{array}{c} r+1 \\ s \end{array} \right\}C_{s,f}.
$$

Thus, (26) holds with $C_{r+1} = \max\left(C_{r,0}, \frac{C_r\left((\alpha^{(0)}+f_{\theta^*}(0))^{r+1}-(\alpha^{(0)})^{r+1}\right)+\sum_{s=0}^{r}\left\{\begin{matrix} r+1 \\ s \end{matrix}\right\}C_{s,f}}{1-(\alpha^{(0)})^{r+1}} \right)$.

With this value of C_{r+1}, on can prove as above that $\mathbb{E}Y_t^{r+1} \le C_{r+1}$ for all $t \in \mathbb{Z}$. This completes the proof of the Proposition. □

As stated in the Introduction, the following approximation study to the stationary regime plays a key role in the proof Theorems 1 and 2.

7.1 Approximation with stationary solutions after the change-point

Under the Lipschitz-type $\mathbf{A}_0(\Theta)$, there exists (see Doukhan et al. 2012, 2013) a stationary and ergodic solutions of the process after k^*; that is, there exists a stationary and ergodic process $\tilde{Y} = (\tilde{Y}_t)_{t\in\mathbb{Z}}$ with finite moment of any order, satisfying:

$$\tilde{Y}_t|\tilde{\mathcal{F}}_{t-1} \sim \text{ Poisson } (\tilde{\lambda}_t) \text{ with } \tilde{\lambda}_t = f_{\theta^*}(\tilde{Y}_{t-1}, \tilde{Y}_{t-2}, \ldots) \quad \text{for } t \in \mathbb{Z} \tag{31}$$

where $\tilde{\mathcal{F}}_t = \sigma(\tilde{Y}_s, s \le t)$ is the σ-field generated by the whole past of $\tilde{Y}$.

For $T \subset \mathbb{N}$, let us consider the conditional (log)-likelihood function (up to a constant) of this stationary regime computed on T:

$$\begin{aligned}\tilde{L}(T,\theta) &= \sum_{t\in T}\left(\tilde{Y}_t \log \tilde{\lambda}_t(\theta) - \tilde{\lambda}_t(\theta)\right) = \sum_{t\in T} \tilde{\ell}_t(\theta) \text{ with } \tilde{\ell}_t(\theta) \\ &= \tilde{Y}_t \log \tilde{\lambda}_t(\theta) - \tilde{\lambda}_t(\theta)\end{aligned} \tag{32}$$

where $\tilde{\lambda}_t(\theta) = f_\theta(\tilde{Y}_{t-1}, \ldots)$; we will use the notation

$$\tilde{f}_\theta^t = f_\theta(\tilde{Y}_{t-1}, \ldots), \quad \text{for all } t \in \mathbb{Z}. \tag{33}$$

The following lemma provides an approximation of the process $(Y_t)_{t>k^*}$ to the second stationary regime.

Lemma 1 *Consider the model* (2) *and assume that the conditions of Theorem* 1 *hold. There exists* $C > 0$ *such that for all* $\ell \ge 1$,

$$\mathbb{E}|Y_{k^*+\ell} - \tilde{Y}_{k^*+\ell}| \le C\left(\inf_{1\le p\le \ell}\left\{(\alpha^{(0)})^{\ell/p} + \sum_{k\ge p}\alpha_k^{(0)}\right\}\right) \tag{34}$$

where $\alpha^{(0)} = \sum_{k\ge1}\alpha_k^{(0)}$.

Proof From the representation (4), we can write (see also Remark 4.1 of Doukhan and Kengne 2015),

$$Y_{k^*+\ell} = N_{k^*+\ell}(\lambda_{k^*+\ell}) \quad \text{with} \quad \lambda_{k^*+\ell} = f_{\theta^*}^{k^*+\ell} = f_{\theta^*}(Y_{k^*+\ell-1}, \dots)$$

and

$$\tilde{Y}_{k^*+\ell} = N_{k^*+\ell}(\tilde{\lambda}_{k^*+\ell}) \quad \text{with} \quad \tilde{\lambda}_{k^*+\ell} = \tilde{f}_{\theta^*}^{k^*+\ell} = f_{\theta^*}(\tilde{Y}_{k^*+\ell-1}, \dots).$$

Hence, we have

$$Y_{k^*+\ell} = F(Y_{k^*+\ell-1}, \dots ; N_{k^*+\ell}) \quad \text{and} \quad \tilde{Y}_{k^*+\ell} = F(\tilde{Y}_{k^*+\ell-1}, \dots ; N_{k^*+\ell})$$

where

$$F(y_1, y_2, \dots ; N_{k^*+\ell}) = N_{k^*+\ell}(f_{\theta^*}(y_1, y_2, \dots)) \quad \text{for any } y_k \in \mathbb{N}, \quad k \geq 1.$$

Therefore,

$$\begin{aligned}
\mathbb{E}|Y_{k^*+\ell} - \tilde{Y}_{k^*+\ell}| &= \mathbb{E}|F(Y_{k^*+\ell}, \dots ; N_{k^*+\ell}) - F(\tilde{Y}_{k^*+\ell}, \dots ; N_{k^*+\ell})| \\
&= \mathbb{E}|N_{k^*+\ell}(f_{\theta^*}^{k^*+\ell}) - N_{k^*+\ell}(\tilde{f}_{\theta^*}^{k^*+\ell})| \\
&= \mathbb{E}\big[\mathbb{E}[|N_{k^*+\ell}(f_{\theta^*}^{k^*+\ell}) - N_{k^*+\ell}(\tilde{f}_{\theta^*}^{k^*+\ell})| \mid \mathcal{F}_{k^*+\ell-1}, \tilde{\mathcal{F}}_{k^*+\ell-1}]\big] \\
&= \mathbb{E}|f_{\theta^*}^{k^*+\ell} - \tilde{f}_{\theta^*}^{k^*+\ell}| \\
&= \mathbb{E}|f_{\theta^*}(Y_{k^*+\ell-1}, \dots) - f_{\theta^*}(\tilde{Y}_{k^*+\ell-1}, \dots)| \\
&\leq \sum_{k \geq 1} \alpha_k^{(0)} \mathbb{E}|Y_{k^*+\ell-k} - \tilde{Y}_{k^*+\ell-k}|
\end{aligned} \tag{35}$$

where the third equality holds since $|N_{k^*+\ell}(f_{\theta^*}^{k^*+\ell}) - N_{k^*+\ell}(\tilde{f}_{\theta^*}^{k^*+\ell})| \mid \mathcal{F}_{k^*+\ell-1}, \tilde{\mathcal{F}}_{k^*+\ell-1}$ can also be considered as a number of events N_t that occur in the time interval $[0, |f_{\theta^*}^{k^*+\ell} - \tilde{f}_{\theta^*}^{k^*+\ell}|]$.

For all $\ell \in \mathbb{N}$, set $u_\ell := \mathbb{E}|Y_{k^*+\ell} - \tilde{Y}_{k^*+\ell}|$. According to Proposition 1, we can find a constant $C_1 > 0$ satisfying $\mathbb{E}Y_t \leq C_1$ for all $t \in \mathbb{Z}$. Hence, since the process $\tilde{Y}$ is stationary, we have for any $\ell \in \mathbb{N}$, $u_\ell \leq \mathbb{E}Y_{k^*+\ell} + \mathbb{E}\tilde{Y}_{k^*+\ell} \leq C_1 + \mathbb{E}\tilde{Y}_0$. Set $C = C_1 + \mathbb{E}\tilde{Y}_0$. Let us show by induction on ℓ that for any $\ell \in \mathbb{N}$,

$$u_\ell \leq C\left(\inf_{1 \leq p \leq \ell} \left\{ (\alpha^{(0)})^{\ell/p} + \frac{1}{1 - \alpha^{(0)}} \sum_{k \geq p} \alpha_k^{(0)} \right\} \right). \tag{36}$$

For $\ell = 1$, (36) holds according to (35). Assume that (36) holds until ℓ. Let $1 \leq p \leq \ell + 1$. From (35), we have

$$\begin{aligned}
u_{\ell+1} &\leq \sum_{k=1}^{p-1} \alpha_k^{(0)} u_{\ell-k+1} + \sum_{k\geq p} \alpha_k^{(0)} u_{\ell-k+1} \\
&\leq C \sum_{k=1}^{p-1} \alpha_k^{(0)} \left((\alpha^{(0)})^{(\ell-k+1)/p} + \frac{1}{1-\alpha^{(0)}} \sum_{i\geq p} \alpha_i^{(0)} \right) + C \sum_{k\geq p} \alpha_k^{(0)} \\
&\leq C \sum_{k=1}^{p-1} \alpha_k^{(0)} (\alpha^{(0)})^{(\ell-k+1)/p} + C \frac{\alpha^{(0)}}{1-\alpha^{(0)}} \sum_{i\geq p} \alpha_i^{(0)} + C \sum_{k\geq p} \alpha_k^{(0)} \qquad (37) \\
&\leq C (\alpha^{(0)})^{(\ell-(p-1)+1)/p} \alpha^{(0)} + C \frac{1}{1-\alpha^{(0)}} \sum_{k\geq p} \alpha_k^{(0)} \\
&\leq C \left((\alpha^{(0)})^{(\ell+1)/p} + \frac{1}{1-\alpha^{(0)}} \sum_{k\geq p} \alpha_k^{(0)} \right).
\end{aligned}$$

Therefore (36) holds for $\ell + 1$. Thus, (34) holds. Note that, in the inequality (37), we have applied

$$u_{\ell-k+1} \leq C \left((\alpha^{(0)})^{(\ell-k+1)/p} + \frac{1}{1-\alpha^{(0)}} \sum_{i\geq p} \alpha_i^{(0)} \right)$$

even when $p \geq \ell - k + 1$. Indeed, in this case, $(\ell - k + 1)/p \leq 1$ and since $\alpha^{(0)} \in (0, 1)$, we have $\alpha^{(0)} \leq (\alpha^{(0)})^{(\ell-k+1)/p}$, then it holds that

$$\begin{aligned}
u_{\ell-k+1} &\leq C \sum_{i\geq 1} \alpha_i^{(0)} \leq C\alpha^{(0)} \leq C(\alpha^{(0)})^{(\ell-k+1)/p} \\
&\leq C \left((\alpha^{(0)})^{(\ell-k+1)/p} + \frac{1}{1-\alpha^{(0)}} \sum_{i\geq p} \alpha_i^{(0)} \right).
\end{aligned}$$

□

Proof of Theorem 1 Let us prove that

$$\frac{1}{n} \left\| \widehat{L}\big(T_{k^*+1,k^*+n}, \theta\big) - \tilde{L}\big(T_{k^*+1,k^*+n}, \theta\big) \right\|_{\Theta} \xrightarrow[n\to\infty]{\text{a.s}} 0. \qquad (38)$$

Indeed, consider the function $\tilde{\mathcal{L}} : \theta \mapsto \mathbb{E}\tilde{\ell}_0(\theta)$; where $\tilde{\ell}_0$ is defined in (32). From the proof of Theorem 3.1 of Doukhan and Kengne (2015), we have $\mathbb{E}\left(\sup_{\theta\in\Theta} |\tilde{\ell}_0(\theta)| \right) < \infty$

,

$$\left\| \frac{1}{n} \tilde{L}\big(T_{k^*+1,k^*+n}, \theta\big) - \tilde{\mathcal{L}}(\theta) \right\|_{\Theta} \xrightarrow[m\to\infty]{\text{a.s}} 0,$$

and that the function $\tilde{\mathcal{L}}$ has a unique maximum at θ^*. If (38) holds, we will get

$$\Big\| \frac{1}{n}\widehat{L}\big(T_{k^*+1,k^*+n},\theta\big) - \tilde{\mathcal{L}}(\theta)\Big\|_\Theta \xrightarrow[m\to\infty]{\text{a.s}} 0;$$

and standard arguments can be used to conclude that $\widehat{\theta}(T_{k^*+1,k^*+n}) \xrightarrow[n\to\infty]{\text{a.s}} \theta^*$. Thus, to complete the proof of the Theorem, it suffices to prove (38).

In the sequel, C denotes a positive constant whom value may differ from an inequality to another. We have

$$\begin{aligned}
&\frac{1}{n}\|\widehat{L}\big(T_{k^*+1,k^*+n},\theta\big) - \tilde{L}\big(T_{k^*+1,k^*+n},\theta\big)\|_\Theta \\
&\quad\le \frac{1}{n}\sum_{t\in T_{k^*+1,k^*+n}} \|\widehat{l}_t(\theta) - \tilde{l}_t(\theta)\|_\Theta \\
&\quad\le \frac{1}{n}\sum_{t=1}^{n} \|\widehat{l}_{k^*+t}(\theta) - \tilde{l}_{k^*+t}(\theta)\|_\Theta.
\end{aligned}$$

Let $0 < r < 1$. According to Kounias and Weng (1969), it suffices to show that

$$\sum_{\ell\ge 1}\Big(\frac{1}{\ell}\Big)^r \mathbb{E}\Big[\|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\|_\Theta^r\Big] < \infty. \tag{39}$$

By using the inequality $|a_1b_1 - a_2b_2| \le |a_1||b_1 - b_2| + |b_2||a_1 - a_2|$ $\forall a_1, a_2, b_1, b_2 \in \mathbb{R}$, we get for all $\ell \in \mathbb{N}$ and $\theta \in \Theta$,

$$\begin{aligned}
\Big|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\Big| &= \Big|Y_{k^*+\ell}\log\widehat{f}_\theta^{k^*+\ell} - \widehat{f}_\theta^{k^*+\ell} - \tilde{Y}_{k^*+\ell}\log\tilde{f}_\theta^{k^*+\ell} + \tilde{f}_\theta^{k^*+\ell}\Big| \\
&\le \Big|Y_{k^*+\ell}\log\widehat{f}_\theta^{k^*+\ell} - \tilde{Y}_{k^*+\ell}\log\tilde{f}_\theta^{k^*+\ell}\Big| + \Big|\widehat{f}_\theta^{k^*+\ell} - \tilde{f}_\theta^{k^*+\ell}\Big| \\
&\le Y_{k^*+\ell}\Big|\log\widehat{f}_\theta^{k^*+\ell} - \log\tilde{f}_\theta^{k^*+\ell}\Big| + \Big|\log\tilde{f}_\theta^{k^*+\ell}\Big||Y_{k^*+\ell} \\
&\quad - \tilde{Y}_{k^*+\ell}| + \Big|\widehat{f}_\theta^{k^*+\ell} - \tilde{f}_\theta^{k^*+\ell}\Big|.
\end{aligned}$$

By applying the mean value theorem at the function $x \mapsto \log x$ on $[\underline{c}, +\infty[$ (since $\widehat{f}_\theta^{k^*+\ell}, \tilde{f}_\theta^{k^*+\ell} \ge \underline{c}$ from the assumption D(Θ)), we get $|\log\widehat{f}_\theta^{k^*+\ell} - \log\tilde{f}_\theta^{k^*+\ell}| \le \frac{1}{\underline{c}}|\widehat{f}_\theta^{k^*+\ell} - \tilde{f}_\theta^{k^*+\ell}|$. Moreover, from the inequality $\log x \le x - 1, \forall x \ge 1$, we have $|\log\tilde{f}_\theta^{k^*+\ell}| = |\log\frac{\tilde{f}_\theta^{k^*+\ell}}{\underline{c}} + \log\underline{c}| \le |\frac{\tilde{f}_\theta^{k^*+\ell}}{\underline{c}} - 1| + |\log\underline{c}|$. Hence,

$$\begin{aligned}
\|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\|_\Theta \le C\big(1 + Y_{k^*+\ell} + \|\tilde{f}_\theta^{k^*+\ell}\|_\Theta\big)\big(|Y_{k^*+\ell} - \tilde{Y}_{k^*+\ell}| \\
+ \|\widehat{f}_\theta^{k^*+\ell} - \tilde{f}_\theta^{k^*+\ell}\|_\Theta\big),
\end{aligned}$$

and from the Hölder's inequality, we get

$$\begin{aligned}
\mathbb{E}\|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\|_\Theta^r \le C\big(\mathbb{E}\big[1 + Y_{k^*+\ell} + \|\tilde{f}_\theta^{k^*+\ell}\|_\Theta\big]^{\frac{r}{1-r}}\big)^{1-r} \\
\big(\mathbb{E}[|Y_{k^*+\ell} - \tilde{Y}_{k^*+\ell}| + \|\widehat{f}_\theta^{k^*+\ell} - \tilde{f}_\theta^{k^*+\ell}\|_\Theta]\big)^r.
\end{aligned}$$

From Proposition 1, and arguments of its proof, for all $s > 0$, we can find a constant $C > 0$ such that $\mathbb{E}Y^s_{k^*+\ell} \le C$, $\mathbb{E}\|\widehat{f}^{k^*+\ell}_\theta\|^s_\Theta \le C$ and $\mathbb{E}\|\tilde{f}^{k^*+\ell}_\theta\|^s_\Theta \le C$. Therefore,

$$\mathbb{E}\|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\|^r_\Theta \le C\big(\mathbb{E}|Y_{k^*+\ell} - \tilde{Y}_{k^*+\ell}| + \mathbb{E}\|\widehat{f}^{k^*+\ell}_\theta - \tilde{f}^{k^*+\ell}_\theta\|_\Theta\big)^r, \quad (40)$$

and also $\mathbb{E}\|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\|^r_\Theta < \infty$ for all $\ell \ge 1$. Hence, it suffices to establish (39) with the sum on $\ell \ge e^2$. Let $\ell \ge e^2$, according to the assumption $\mathbf{A}_0(\Theta)$ and (34), we get

$$\begin{aligned}
\mathbb{E}\|\widehat{f}^{k^*+\ell}_\theta - \tilde{f}^{k^*+\ell}_\theta\|_\Theta &\le \sum_{j\ge 1}\alpha^{(0)}_j \mathbb{E}|Y_{k^*+\ell-j} - \tilde{Y}_{k^*+\ell-j}| \\
&\le \sum_{j=1}^{\ell/2-1} \alpha^{(0)}_j \mathbb{E}|Y_{k^*+\ell-j} - \tilde{Y}_{k^*+\ell-j}| + C\sum_{j\ge \ell/2}\alpha^{(0)}_j \\
&\le C\sum_{j=1}^{\ell/2-1} \alpha^{(0)}_j \left(\inf_{1\le p\le \ell-j}\left\{(\alpha^{(0)})^{(\ell-j)/p} + \sum_{i\ge p}\alpha^{(0)}_i\right\}\right) \\
&\quad + C\sum_{j\ge \ell/2}\alpha^{(0)}_j \\
&\le C\left(\inf_{1\le p\le \ell/2}\left\{(\alpha^{(0)})^{\ell/(2p)} + \sum_{i\ge p}\alpha^{(0)}_i\right\} + \sum_{j\ge \ell/2}\alpha^{(0)}_j\right).
\end{aligned}$$

Thus, (40) and (34) imply

$$\begin{aligned}
\mathbb{E}\|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\|^r_\Theta &\le C\Bigg(\inf_{1\le p\le \ell}\left\{(\alpha^{(0)})^{\ell/p} + \sum_{j\ge p}\alpha^{(0)}_j\right\} \\
&\qquad + \inf_{1\le p\le \ell/2}\left\{(\alpha^{(0)})^{\ell/(2p)} + \sum_{j\ge p}\alpha^{(0)}_j\right\} + \sum_{j\ge \ell/2}\alpha^{(0)}_j\Bigg)^r \\
&\le C\left(\inf_{1\le p\le \ell/2}\left\{(\alpha^{(0)})^{\ell/(2p)} + \sum_{j\ge p}\alpha^{(0)}_j\right\} + \sum_{j\ge \ell/2}\alpha^{(0)}_j\right)^r \\
&\le C'\left((\alpha^{(0)})^{\ell/(2p_\ell)} + \sum_{j\ge p_\ell}\alpha^{(0)}_j\right)^r \\
&\le C'\left((\alpha^{(0)})^{r\ell/(2p_\ell)} + \left(\sum_{j\ge p_\ell}\alpha^{(0)}_j\right)^r\right)
\end{aligned}$$

with $p_\ell = \ell/\log\ell$ and $C' = 2^r C$. Hence,

$$
\begin{aligned}
&\sum_{\ell \geq e^2} \Big(\frac{1}{\ell}\Big)^r \mathbb{E}\big[\|\widehat{l}_{k^*+\ell}(\theta) - \tilde{l}_{k^*+\ell}(\theta)\|_\Theta^r\big] \\
&\quad \leq C' \sum_{\ell \geq e^2} \Big(\frac{1}{\ell}\Big)^r \Bigg((\alpha^{(0)})^{r\ell/(2p_\ell)} \Bigg(\sum_{j \geq p_\ell} \alpha_j^{(0)} \Bigg)^r \Bigg) \\
&\quad \leq C' \sum_{\ell \geq e^2} \frac{1}{\ell^r} (\alpha^{(0)})^{\frac{r}{2}\log \ell} + C' \sum_{\ell \geq e^2} \frac{1}{\ell^r} \Bigg(\sum_{j \geq p_\ell} \alpha_j^{(0)} \Bigg)^r \\
&\quad \leq C' \sum_{\ell \geq e^2} \frac{1}{\ell^{r - \frac{r}{2}\log \alpha^{(0)}}} + C' \sum_{\ell \geq e^2} \frac{1}{\ell^r} \Bigg(\frac{1}{(p_\ell)^{\gamma - 1}} \Bigg)^r \\
&\quad \leq C' \sum_{\ell \geq e^2} \frac{1}{\ell^{\frac{r}{2}(2 - \log \alpha^{(0)})}} + C' \sum_{\ell \geq e^2} \frac{(\log \ell)^{r(\gamma - 1)}}{\ell^{r\gamma}}.
\end{aligned}
\tag{41}
$$

For $r \in (\max(\frac{2}{3}, \frac{2}{2-\log \alpha^{(0)}})\,, 1)$, we have $\frac{r}{2}(2 - \log \alpha^{(0)}) > 1$ and $r\gamma > 1$; which ensures that the sums on the right-hand side of (41) are finite. Thus, one can find $r \in (0, 1)$ such that (39) holds; which achieves the proof of (38) and completes the proof of the Theorem. $\square$

Proof of Theorem 2 For any $1 \leq i \leq d$, from the Taylor expansion to the function $\frac{\partial}{\partial \theta_i}\widehat{L}_n(T_{k^*+1,k^*+n}, \theta)$, there exists $\theta_{n,i}$ between $\widehat{\theta}(T_{k^*+1,k^*+n})$ and θ^* such that

$$
\begin{aligned}
\frac{\partial}{\partial \theta_i}\widehat{L}_n(T_{k^*+1,k^*+n}, \widehat{\theta}(T_{k^*+1,k^*+n})) = {}& \frac{\partial}{\partial \theta_i}\widehat{L}_n(T_{k^*+1,k^*+n}, \theta^*) \\
&+ \frac{\partial^2}{\partial \theta \partial \theta_i}\widehat{L}_n(T_{k^*+1,k^*+n}, \theta_{n,i}) \\
&\cdot (\widehat{\theta}(T_{k^*+1,k^*+n}) - \theta^*).
\end{aligned}
$$

Hence,

$$
\begin{aligned}
n\widehat{G}_n \cdot (\widehat{\theta}(T_{k^*+1,k^*+n}) - \theta^*) = {}& \frac{\partial}{\partial \theta}\widehat{L}_n(T_{k^*+1,k^*+n}, \theta^*) \\
&- \frac{\partial}{\partial \theta}\widehat{L}_n(T_{k^*+1,k^*+n}, \widehat{\theta}(T_{k^*+1,k^*+n})),
\end{aligned}
\tag{42}
$$

with

$$
\widehat{G}_n = -\frac{1}{n}\Bigg(\frac{\partial^2}{\partial \theta \partial \theta_i}\widehat{L}_n(T_{k^*+1,k^*+n}, \theta_{n,i}) \Bigg)_{1 \leq i \leq d}.
$$

Since $\widehat{\theta}(T_{k^*+1,k^*+n}) \underset{n \to \infty}{\longrightarrow} \theta^*$ and $\theta^* \in \mathring{\Theta}$, for n large enough, $\widehat{\theta}(T_{k^*+1,k^*+n}) \in \mathring{\Theta}$ and $\frac{\partial}{\partial \theta}\widehat{L}_n(T_{k^*+1,k^*+n}, \widehat{\theta}(T_{k^*+1,k^*+n})) = 0$. Therefore, (42) gives

$$
n\widehat{G}_n \cdot (\widehat{\theta}(T_{k^*+1,k^*+n}) - \theta^*) = \frac{\partial}{\partial \theta}\widehat{L}_n(T_{k^*+1,k^*+n}, \theta^*). \tag{43}
$$

The following convergences hold (see the proofs in the supplementary material).

$$\frac{1}{n}\left\| \frac{\partial^2}{\partial\theta\partial\theta'}\widehat{L}\big(T_{k^*+1,k^*+n},\theta\big) - \frac{\partial^2}{\partial\theta\partial\theta'}\tilde{L}\big(T_{k^*+1,k^*+n},\theta\big)\right\|_{\Theta} \xrightarrow[n\to\infty]{\text{a.s}} 0; \tag{44}$$

$$\frac{1}{n}\left\| \sum_{t=k^*+1}^{k^*+n}\frac{1}{\widehat{f}^t_\theta}\Big(\frac{\partial}{\partial\theta}\widehat{f}^t_\theta\Big)\Big(\frac{\partial}{\partial\theta}\widehat{f}^t_\theta\Big)' - \sum_{t=k^*+1}^{k^*+n}\frac{1}{\tilde{f}^t_\theta}\Big(\frac{\partial}{\partial\theta}\tilde{f}^t_\theta\Big)\Big(\frac{\partial}{\partial\theta}\tilde{f}^t_\theta\Big)'\right\|_{\Theta} \xrightarrow[n\to\infty]{\text{a.s}} 0; \tag{45}$$

$$\mathbb{E}\left(\frac{1}{\sqrt{n}}\left\|\frac{\partial}{\partial\theta}\widehat{L}\big(T_{k^*+1,k^*+n},\theta\big) - \frac{\partial}{\partial\theta}\tilde{L}\big(T_{k^*+1,k^*+n},\theta\big)\right\|_{\Theta}\right) \xrightarrow[n\to\infty]{} 0. \tag{46}$$

From Lemma 7.2 and the proof of Theorem 3.2 of Doukhan and Kengne (2015), it follows that:

$$\frac{1}{n}\left\| \frac{\partial^2}{\partial\theta\partial\theta'}\tilde{L}\big(T_{k^*+1,k^*+n},\theta\big) - \mathbb{E}\left(\frac{\partial^2}{\partial\theta\partial\theta'}\tilde{\ell}_0(\theta)\right)\right\|_{\Theta} \xrightarrow[n\to\infty]{\text{a.s}} 0; \tag{47}$$

$$\left\| \sum_{t=k^*+1}^{k^*+n}\frac{1}{\tilde{f}^t_\theta}\Big(\frac{\partial}{\partial\theta}\tilde{f}^t_\theta\Big)\Big(\frac{\partial}{\partial\theta}\tilde{f}^t_\theta\Big)' - \mathbb{E}\left(\frac{1}{\tilde{f}^0_\theta}\Big(\frac{\partial}{\partial\theta}\tilde{f}^0_\theta\Big)\Big(\frac{\partial}{\partial\theta}\tilde{f}^0_\theta\Big)'\right)\right\|_{\Theta} \xrightarrow[n\to\infty]{\text{a.s}} 0; \tag{48}$$

$$\mathbb{E}\left(\frac{1}{\tilde{f}^0_{\theta^*}}\Big(\frac{\partial}{\partial\theta}\tilde{f}^0_{\theta^*}\Big)\Big(\frac{\partial}{\partial\theta}\tilde{f}^0_{\theta^*}\Big)'\right) = -\mathbb{E}\left(\frac{\partial^2}{\partial\theta\partial\theta'}\tilde{\ell}_0(\theta^*)\right) = \widetilde{\Sigma}; \tag{49}$$

$$\frac{1}{\sqrt{n}}\frac{\partial}{\partial\theta}\tilde{L}\big(T_{k^*+1,k^*+n},\theta^*\big) = \frac{1}{\sqrt{n}}\sum_{t=k^*+1}^{k^*+n}\frac{\partial}{\partial\theta}\ell_t(\theta^*) \xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}(0,\widetilde{\Sigma}). \tag{50}$$

According to Theorem 1, (44), (45), (47), (48) and (49), we get $\widehat{G}_n \xrightarrow[n\to\infty]{\text{a.s}} \widetilde{\Sigma}$ and also $\widehat{\Sigma}_n \xrightarrow[n\to\infty]{\text{a.s}} \widetilde{\Sigma}$. Hence, for n large enough, $\widehat{G}_n$ is invertible, therefore in addition to (43), (46) and (50) it holds that

$$\begin{aligned}\sqrt{n}(\widehat{\theta}(T_{k^*+1,k^*+n}) - \theta^*) &= \frac{1}{\sqrt{n}}\widehat{G}_n^{-1}\frac{\partial}{\partial\theta}\widehat{L}_n(T_{k^*+1,k^*+n},\theta^*)\\ &= \frac{1}{\sqrt{n}}\widetilde{\Sigma}^{-1}\frac{\partial}{\partial\theta}\widetilde{L}_n(T_{k^*+1,k^*+n},\theta^*) + o_P(1)\\ &\xrightarrow[n\to+\infty]{\mathcal{D}} \mathcal{N}(0,\widetilde{\Sigma}^{-1}).\end{aligned}$$

□

Let $k > m$ and $T_{1,m} = \{1, \ldots, m\}$, $T_{\ell,k} = \{\ell, \ell+1, \ldots, k\}$ with $\ell \in \Pi_{m,k} = \{m - v_m, v_m + 1, \ldots, k - v_m\}$, define

$$C_{k,\ell} := \sqrt{m}\,\frac{k-\ell}{k}\left\|\Sigma^{1/2}\cdot\left(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,m})\right)\right\|,$$

with $\widehat{\theta}$ defined in (14).

Lemma 2 *Under the assumptions of Theorem 3,*

$$\sup_{k>m}\max_{\ell\in\Pi_{m,k}}\frac{1}{b((k-\ell)/m)}\left|\widehat{C}_{k,\ell} - C_{k,\ell}\right| = o_P(1) \quad \text{as} \quad n\to\infty.$$

Proof For any $m \geq 1$, we have

$$\sup_{k>m}\max_{\ell\in\Pi_{m,k}}\frac{1}{b((k-\ell)/m)}\left|\widehat{C}_{k,\ell} - C_{k,\ell}\right| \leq \frac{1}{\inf_{s>0} b(s)}\sup_{k>m}\max_{\ell\in\Pi_{m,k}}\left|\widehat{C}_{k,\ell} - C_{k,\ell}\right|.$$

Therefore, similar arguments as in the proof of Lemma 7.3 of Doukhan and Kengne (2015) lead to conclusion. □

Proof of Theorem 3 Recall that

$$P\{\tau(m) < \infty\} = P\left\{\sup_{m<k\leq[Tm]+1}\max_{\ell\in\Pi_{m,k}}\frac{\widehat{C}_{k,\ell}}{b((k-\ell)/m)} > 1\right\}.$$

Hence, it suffices to show that

$$\sup_{m<k\leq[Tm]+1}\max_{\ell\in\Pi_{m,k}}\frac{1}{b((k-\ell)/m)}\widehat{C}_{k,\ell}\xrightarrow[m\to+\infty]{\mathcal{D}}\sup_{1<t\leq T}\sup_{0<s<t-1}\frac{\|W_d(s) - sW_d(1))\|}{t\,b(s)}. \tag{51}$$

According to Lemma 2, it is enough to show that

$$\sup_{m<k\leq[Tm]+1}\max_{\ell\in\Pi_{m,k}}\frac{1}{b((k-\ell)/m)}C_{k,\ell}\xrightarrow[m\to+\infty]{\mathcal{D}}\sup_{1<t\leq T}\sup_{0<s<t-1}\frac{\|W_d(s) - sW_d(1))\|}{t\,b(s)}. \tag{52}$$

Let $k > m$ and $\ell \in \Pi_{m,k}$. From the proof of Theorem 4.1 of Doukhan and Kengne (2015), it holds that, as $m \to \infty$

$$\Sigma(\widehat{\theta}(T_{1,m}) - \theta_0^*) = \frac{1}{m}\frac{\partial}{\partial\theta}L_m(T_{1m}, \theta_0^*) + o_P\left(\frac{1}{\sqrt{m}}\right) \quad \text{and} \quad \Sigma(\widehat{\theta}(T_{\ell,k}) - \theta_0^*)$$
$$= \frac{1}{k-\ell}\frac{\partial}{\partial\theta}L_m(T_{\ell,k}, \theta_0^*) + o_P\left(\frac{1}{\sqrt{k-\ell}}\right).$$

Therefore,

$$\Sigma(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,m})) = \frac{1}{k-\ell}\left(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{m}\frac{\partial}{\partial\theta}L_m(T_{1,m},\theta_0^*)\right) + o_P\left(\frac{1}{\sqrt{k-\ell}} + \frac{1}{\sqrt{m}}\right).$$

This implies

$$\sqrt{m}\frac{k-\ell}{k}\Sigma^{1/2}(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,m})) = \frac{\sqrt{m}}{k}\Sigma^{-1/2}\Big(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{m}\frac{\partial}{\partial\theta}L_m(T_{1,m},\theta_0^*)\Big) + o_P(1).$$

Hence,

$$\begin{aligned}
&\sup_{m<k\leq[Tm]+1}\max_{\ell\in\Pi_{m,k}}\frac{1}{b((k-\ell)/m)}\Big\|\sqrt{m}\frac{k-\ell}{k}\Sigma^{1/2}(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,m})) \\
&\quad - \frac{\sqrt{m}}{k}\Sigma^{-1/2}\left(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{m}\frac{\partial}{\partial\theta}L_m(T_{1,m},\theta_0^*)\right)\Big\| \\
&\leq \frac{1}{\inf_{s>0} b(s)}\sup_{m<k\leq[Tm]+1}\max_{\ell\in\Pi_{m,k}}\Big\|\sqrt{m}\frac{k-\ell}{k}\Sigma^{1/2}(\widehat{\theta}(T_{\ell,k}) - \widehat{\theta}(T_{1,m})) \\
&\quad - \frac{\sqrt{m}}{k}\Sigma^{-1/2}\left(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{m}\frac{\partial}{\partial\theta}L_m(T_{1,m},\theta_0^*)\right)\Big\| \\
&= o_P(1).
\end{aligned}$$

Thus, to complete the proof of the theorem, we will prove that

$$\begin{aligned}
&\sup_{m<k\leq[Tm]+1}\max_{\ell\in\Pi_{m,k}}\frac{1}{b((k-\ell)/m)}\frac{\sqrt{m}}{k}\Big\|\Sigma^{-1/2}\Big(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) \\
&\quad - \frac{k-\ell}{m}\frac{\partial}{\partial\theta}L_m(T_{1,m},\theta_0^*)\Big)\Big\| \\
&\xrightarrow[m\to+\infty]{\mathcal{D}} \sup_{1<t\leq T}\sup_{0<s<t-1}\frac{\|W_d(s) - sW_d(1))\|}{t\,b(s)}.
\end{aligned} \tag{53}$$

Let $k > m$ and $\ell \in \Pi_{m,k}$. We have

$$\begin{aligned}
&\frac{\sqrt{m}}{k}\left(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{m}\frac{\partial}{\partial\theta}L_m(T_{1,m},\theta_0^*)\right) \\
&\quad = -\frac{m}{k}\frac{1}{\sqrt{m}}\left(\sum_{i=\ell}^{k}\frac{\partial l_i(\theta_0^*)}{\partial\theta} - \frac{k-\ell}{m}\sum_{i=1}^{m}\frac{\partial l_i(\theta_0^*)}{\partial\theta}\right).
\end{aligned}$$

Let us consider the following cases.

(i) Closed-end procedure.

Let $1 < \mathcal{T} < \infty$. Define the set $S := \{(t,s) \in [1,\mathcal{T}] \times [0,\mathcal{T}-1]/s < t\}$. According to Doukhan and Kengne (2015), $\left(\frac{\partial l_i(\theta_0^*)}{\partial \theta}, \mathcal{F}_i\right)_{i\in\mathbb{Z}}$ is a stationary ergodic square integrable martingale difference sequence with covariance matrix Σ. By the Cramér-Wold device (see Billingsley 1968), it holds that

$$\frac{1}{\sqrt{m}} \sum_{i=[ms]}^{[mt]} \frac{\partial l_i(\theta_0^*)}{\partial \theta} \xrightarrow[m\to\infty]{\mathcal{D}(S)} W_\Sigma(t-s)$$

where $\xrightarrow[m\to\infty]{\mathcal{D}(S)}$ denotes the weak convergence on the Skorohod space $\mathcal{D}(S)$ and W_Σ is a centered Gaussian process such that $\mathbb{E}\left(W_\Sigma(s), W_\Sigma(\tau)'\right) = \min(s,\tau)\Sigma$. Therefore

$$\frac{1}{\sqrt{m}}\left(\sum_{i=[ms]}^{[mt]} \frac{\partial l_i(\theta_0^*)}{\partial \theta} - \frac{[mt]-[ms]}{m}\sum_{i=1}^{m} \frac{\partial l_i(\theta_0^*)}{\partial \theta}\right)$$
$$\xrightarrow[m\to\infty]{\mathcal{D}(S)} W_\Sigma(t-s) - (t-s)B_\Sigma(1);$$

and

$$\frac{1}{\sqrt{m}}\Sigma^{-1/2}\left(\sum_{i=[ms]}^{[mt]} \frac{\partial l_i(\theta_0^*)}{\partial \theta} - \frac{[mt]-[ms]}{m}\sum_{i=1}^{m} \frac{\partial l_i(\theta_0^*)}{\partial \theta}\right)$$
$$\xrightarrow[m\to\infty]{\mathcal{D}(S)} W_d(t-s) - (t-s)B_d(1);$$

Hence

$$
\begin{aligned}
&\sup_{m<k<[\mathcal{T}m]+1} \max_{\ell\in\Pi_{m,k}} \frac{1}{b((k-\ell)/m)} \frac{\sqrt{m}}{k} \\
&\quad \left\| \Sigma^{-1/2}\left(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{n}\frac{\partial}{\partial\theta}L_m(T_{1,n},\theta_0^*)\right)\right\| \\
&= \sup_{m<k<[\mathcal{T}m]+1} \max_{\ell\in\Pi_{m,k}} \frac{1}{b((k-\ell)/m)} \frac{\sqrt{m}}{k} \\
&\quad \left\| \Sigma^{-1/2}\left(\sum_{i=\ell}^{k}\frac{\partial l_i(\theta_0^*)}{\partial\theta} - \frac{k-\ell}{n}\sum_{i=1}^{m}\frac{\partial l_i(\theta_0^*)}{\partial\theta}\right)\right\| \\
&= \sup_{t\in\left\{1+\frac{1}{m},1+\frac{2}{m},\ldots,\frac{[\mathcal{T}m]}{m}\right\}} \max_{s\in\left\{1-\frac{v_m}{m},1+\frac{1}{m}-\frac{v_m}{m},1+\frac{2}{m}-\frac{v_m}{m},\ldots,\frac{[mt]}{m}-\frac{v_m}{m}\right\}} \\
&\quad \left[\frac{1}{b(([mt]-[ms])/m)}\frac{m}{[mt]}\right. \\
&\quad \times\left.\left\|\frac{1}{\sqrt{m}}\Sigma^{-1/2}\left(\sum_{i=[ms]}^{[mt]}\frac{\partial l_i(\theta_0^*)}{\partial\theta} - \frac{[mt]-[ms]}{m}\sum_{i=1}^{m}\frac{\partial l_i(\theta_0^*)}{\partial\theta}\right)\right\|\right] \\
&\xrightarrow[m\to+\infty]{\mathcal{D}} \sup_{1<t<\mathcal{T}}\sup_{1<s<t}\frac{\|W_d(t-s)-(t-s)W_d(1)\|}{t\,b(t-s)} \overset{\mathcal{D}}{=} \sup_{1<t<\mathcal{T}}\sup_{0<s<t-1} \\
&\quad \frac{\|W_d(s)-s\,W_d(1)\|}{t\,b(s)}.
\end{aligned}
\tag{54}
$$

(ii) Open-end procedure. We proceed as in proof of Lemma 6.3 of Bardet and Kengne (2014). Thus, according to (53) and (i), it suffices to show that the limit in distribution (as $m,\mathcal{T}\to\infty$) of

$$
\sup_{k>[\mathcal{T}m]} \max_{\ell\in\Pi_{m,k}} \frac{1}{b((k-\ell)/m)} \frac{\sqrt{m}}{k}\left\| \Sigma^{-1/2}\left(\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*) - \frac{k-\ell}{m}\frac{\partial}{\partial\theta}L_m(T_{1,m},\theta_0^*)\right)\right\|
$$

exists and is equal to the limit in distribution (as $\mathcal{T}\to\infty$) of

$$
\sup_{t>\mathcal{T}}\sup_{0<s<t-1}\frac{\|W_d(s)-sW_d(1))\|}{t\,b(s)}.
$$

Let $k>[\mathcal{T}m]$. For some $\ell_k\in\Pi_{m,k}$, we have

$$
\max_{\ell\in\Pi_{m,k}} \frac{1}{b((k-\ell)/m)}\frac{\sqrt{m}}{k}\left\|\frac{\partial}{\partial\theta}L_m(T_{\ell,k},\theta_0^*)\right\| = \frac{1}{b((k-\ell_k)/m)}\frac{\sqrt{m}}{k}\left\|\sum_{i=\ell_k}^{k}\frac{\partial l_i(\theta_0^*)}{\partial\theta}\right\|.
$$

From the Hájek-Rényi-Chow inequality (see Chow 1960), we get

$$\forall x > 0, \lim_{\mathcal{T}\to\infty} \limsup_{m\to\infty} P\left(\sup_{k>[\mathcal{T}m]} \frac{1}{b((k-\ell_k)/m)} \frac{\sqrt{m}}{k} \left\| \sum_{i=\ell_k}^{k} \frac{\partial l_i(\theta_0^*)}{\partial \theta} \right\| > x \right) = 0. \tag{55}$$

Moreover, since the function $b(\cdot)$ is non-increasing, we have for any $m, \mathcal{T} > 1$

$$\begin{aligned}
&\sup_{k>m\mathcal{T}} \max_{\ell\in\Pi_{m,k}} \frac{1}{b((k-\ell)/m)} \frac{\sqrt{m}}{k} \left\| \frac{k-\ell}{m} \frac{\partial}{\partial\theta} L_m(T_{1,m}, \theta_0^*) \right\| \\
&= \left\| \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \frac{\partial l_i(\theta_0^*)}{\partial\theta} \right\| \times \sup_{k>\mathcal{T}m} \max_{\ell\in\Pi_{m,k}} \frac{1}{b((k-\ell)/m)} \frac{k-\ell}{k} \\
&= \left\| \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \frac{\partial l_i(\theta_0^*)}{\partial\theta} \right\| \times \sup_{k>\mathcal{T}m} \frac{1}{b((k-v_m)/m)} \frac{k-v_m}{k} \\
&= \frac{1}{\inf_{s>0} b(s)} \left\| \frac{1}{\sqrt{m}} \sum_{i=1}^{m} \frac{\partial l_i(\theta_0^*)}{\partial\theta} \right\| \\
&\xrightarrow[m\to+\infty]{\mathcal{D}} \frac{1}{\inf_{s>0} b(s)} \|W_\Sigma(1)\|,
\end{aligned} \tag{56}$$

where the latter convergence holds from the Cramér-Wold device and the central limit theorem applied to the martingale difference sequence $\left(\frac{\partial l_i(\theta_0^*)}{\partial\theta}, \mathcal{F}_i \right)_{i\in\mathbb{Z}}$. According to (55) and (56), it follows that

$$\begin{aligned}
&\sup_{k>\mathcal{T}m} \max_{\ell\in\Pi_{m,k}} \frac{1}{b((k-\ell)/m)} \frac{\sqrt{n}}{k} \left\| \Sigma^{-1/2} \left(\frac{\partial}{\partial\theta} L_m(T_{\ell,k}, \theta_0^*) - \frac{k-\ell}{m} \right.\right. \\
&\left.\left. \frac{\partial}{\partial\theta} L_m(T_{1,m}, \theta_0^*) \right) \right\| \xrightarrow[m\to+\infty]{\mathcal{D}} \frac{1}{\inf_{s>0} b(s)} \|W_d(1)\|.
\end{aligned} \tag{57}$$

On the other hand, from the proof of Lemma 6.3 of Bardet and Kengne (2014), we get

$$\sup_{t>\mathcal{T}} \sup_{0<s<t-1} \frac{\|W_\Sigma(s) - sW_\Sigma(1)\|}{t\, b(s)} \xrightarrow[\mathcal{T}\to+\infty]{\mathcal{D}} \frac{1}{\inf_{s>0} b(s)} \|W_\Sigma(1)\|.$$

This implies

$$\sup_{t>\mathcal{T}} \sup_{0<s<t-1} \frac{\|W_d(s) - sW_d(1)\|}{t\, b(s)} \xrightarrow[\mathcal{T}\to+\infty]{\mathcal{D}} \frac{1}{\inf_{s>0} b(s)} \|W_d(1)\|. \tag{58}$$

Equations (57) and (58) complete the proof in the case of the open-end procedure. □

Proof of Theorem 4 In the sequel, C denotes a positive constant whom value may differ from an inequality to another. Denote $k_m = k^* + m^\delta$ for $\delta \in (1/2, 1)$. For m large enough, we have $m \le k_m \le [\mathcal{T}m] + 1$ for both open-end and closed-end procedure; moreover, $v_n << n^{\delta}$ and $k^* \in \Pi_{m,k_m}$. Hence, according to assumption **B**, we can find a constant $C > 0$ such that

$$\begin{aligned}
\max_{\ell \in \Pi_{m,k_m}} \frac{\widehat{C}_{k_m,\ell}}{b((k_m-\ell)/m)} &= \max_{\ell \in \Pi_{m,k_m}} \frac{1}{b((k_m-\ell)/m)} \sqrt{m}\frac{k_m-\ell}{k_m} \big\| \widehat{\Sigma}_m^{1/2} \big(\widehat{\theta}(T_{\ell,k_m}) - \widehat{\theta}(T_{1,m})\big) \big\| \\
&\ge \frac{1}{b((k_m-k^*)/m)} \sqrt{m}\frac{k_m-k^*}{k_m} \big\| \widehat{\Sigma}_m^{1/2} \big(\widehat{\theta}(T_{k^*,k_m}) - \widehat{\theta}(T_{1,m})\big) \big\| \\
&\ge C\sqrt{m}\, \frac{m^{\delta}}{[T^*m] + m^{\delta}} \big\| \widehat{\Sigma}_m^{1/2} \big(\widehat{\theta}(T_{k^*,k_m}) - \widehat{\theta}(T_{1,m})\big) \big\| \\
&\ge C\, m^{\delta-1/2} \big\| \widehat{\Sigma}_m^{1/2} \big(\widehat{\theta}(T_{k^*,k_m}) - \widehat{\theta}(T_{1,m})\big) \big\|.
\end{aligned} \tag{59}$$

From Doukhan and Kengne (2015), we get $\widehat{\Sigma}_m^{1/2} \xrightarrow[m\to\infty]{a.s} \Sigma^{1/2}$ and $\widehat{\theta}(T_{1,m}) \xrightarrow[m\to\infty]{a.s} \theta_0^*$. Moreover, from Theorem 1, $\widehat{\theta}(T_{k^*,k_m}) \xrightarrow[m\to\infty]{a.s} \theta_1^*$. Thus, since Σ is symmetric positive definite, $\theta_0^* \neq \theta_1^*$ and $\delta > 1/2$, (59) implies

$$\max_{\ell \in \Pi_{m,k_m}} \frac{\widehat{C}_{k_m,\ell}}{b((k_m-\ell)/m)} \xrightarrow[m\to\infty]{a.s} \infty.$$

□

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10463-021-00815-1.

Acknowledgements The authors are grateful to the three anonymous Referees for many relevant suggestions and comments which helped to improve the contents of this paper.

References

Ahmad, A., Francq, C. (2016). Poisson QMLE of count time series models. *Journal of Time Series Analysis, 37*(3), 291–314.

Bardet, J.-M., Kengne, W. (2014). Monitoring procedure for parameter change in causal time series. *Journal of Multivariate Analysis, 125*, 204–221.

Billingsley, P. (1968). *Convergence of probability measures*. New York: Wiley.

Chow, Y. (1960). A martingale inequality and the law of large numbers. *Proceedings of the American Mathematical Society, 11*(1), 107–111.

Chu, C.-S.J., Stinchcombe, M., White, H. (1996). Monitoring structural change. *Econometrica*: *Journal of the Econometric Society*, *64*(5), 1045–1065.

Diop, M. L., Kengne, W. (2017). Testing parameter change in general integer-valued time series. *Journal of Time Series Analysis, 38*(6), 880–894.

Diop, M. L., Kengne, W. (2021). Piecewise autoregression for general integer-valued time series. *Journal of Statistical Planning and Inference, 211*, 271–286.

Doukhan, P., Kengne, W. (2015). Inference and testing for structural change in general Poisson autoregressive models. *Electronic Journal of Statistics, 9*, 1267–1314.

Doukhan, P., Fokianos, K., Tjøstheim, D. (2012). On weak dependence conditions for Poisson autoregressions. *Statistics& Probability Letters, 82*(5), 942–948.

Doukhan, P., Fokianos, K., Tjøstheim, D. (2013). Correction to “on weak dependence conditions for poisson autoregressions’’ [statist. probab. lett. 82 (2012) 942–948]. *Statistics& Probability Letters, 83*(8), 1926–1927.

Doukhan, P., Wintenberger, O. (2008). Weakly dependent chains with infinite memory. *Stochastic Processes and their Applications, 118*(11), 1997–2013.

Ferland, R., Latour, A., Oraichi, D. (2006). Integer-valued GARCH process. *Journal of Time Series Analysis, 27*(6), 923–942.

Fokianos, K., Tjøstheim, D. (2012). Nonlinear Poisson autoregression. *Annals of the Institute of Statistical Mathematics, 64*(6), 1205–1225.

Fokianos, K., Rahbek, A., Tjøstheim, D. (2009). Poisson autoregression. *Journal of the American Statistical Association, 104*(488), 1430–1439.

Franke, J., Kirch, C., Kamgaing, J. T. (2012). Changepoints in times series of counts. *Journal of Time Series Analysis, 33*(5), 757–770.

Godambe, V. P. (1960). An optimum property of regular maximum likelihood estimation. *The Annals of Mathematical Statistics, 31*(4), 1208–1211.

Gombay, E., Serban, D. (2009). Monitoring parameter change in AR(p) time series models. *Journal of Multivariate Analysis, 100*(4), 715–725.

Horváth, L., Hušková, M., Kokoszka, P., Steinebach, J. (2004). Monitoring changes in linear models. *Journal of Statistical Planning and Inference, 126*(1), 225–251.

Kang, J., Lee, S. (2014). Parameter change test for Poisson autoregressive models. *Scandinavian Journal of Statistics, 41*(4), 1136–1152.

Kengne, W. (2015). Sequential change-point detection in Poisson autoregressive models. *Journal de la Société Française de Statistique, 156*(4), 98–112.

Kengne, W. C. (2012). Testing for parameter constancy in general causal time-series models. *Journal of Time Series Analysis, 33*(3), 503–518.

Kirch, C., Kamgaing, J. T. (2015). On the use of estimating functions in monitoring time series for change points. *Journal of Statistical Planning and Inference, 161*, 25–49.

Kirch, C., Weber, S. (2018). Modified sequential change point procedures based on estimating functions. *Electronic Journal of Statistics, 12*(1), 1579–1613.

Kounias, E. G., Weng, T.-S. (1969). An inequality and almost sure convergence. *The Annals of Mathematical Statistics, 40*(3), 1091–1093.

Leisch, F., Hornik, K., Kuan, C.-M. (2000). Monitoring structural changes with the generalized fluctuation test. *Econometric Theory, 16*(06), 835–854.

Na, O., Lee, Y., Lee, S. (2011). Monitoring parameter change in time series models. *Statistical Methods& Applications, 20*(2), 171–199.

Weiß, C. H. (2009). Modelling time series of counts with overdispersion. *Statistical Methods and Applications, 18*(4), 507–519.