Estimation of subcritical Galton Watson processes with correlated immigration

Yacouba Boubacar Mainassara, Landy Rabehasaina

Estimation of subcritical Galton Watson processes with correlated immigration

Yacouba Boubacar Maïnassara[a], Landy Rabehasaina[b]

[a] *Université Polytechnique Hauts-de-France, INSA Hauts-de-France, CERAMATHS - Laboratoire de Matériaux Céramiques et de Mathématiques, F-59313 Valenciennes, France*
[b] *Université Bourgogne Franche-Comté, Laboratoire de mathématiques de Besançon, UMR CNRS 6623, 16 Route de Gray, 25030 Besançon, France*

Abstract

We consider an observed subcritical Galton Watson process $\{Y_n,\ n \in \mathbb{Z}\}$ with correlated stationary immigration process $\{\epsilon_n,\ n \in \mathbb{Z}\}$. Two situations are presented. The first one is when $\mathrm{Cov}(\epsilon_0, \epsilon_k) = 0$ for k larger than some k_0: a consistent estimator for the reproduction and mean immigration rates is given, and a central limit theorem is proved. The second one is when $\{\epsilon_n,\ n \in \mathbb{Z}\}$ has general correlation structure: under mixing assumptions, we exhibit an estimator for the the logarithm of the reproduction rate and we prove that it converges in quadratic mean with explicit speed. In addition, when the mixing coefficients decrease fast enough, we provide and prove a two terms expansion for the estimator. Numerical illustrations are provided.

Keywords: Galton Watson processes, immigration, INAR processes

1. Introduction

Estimation of parameters in a Galton Watson process with immigration has a long history: we refer to the seminal paper Klimko and Nelson (1978) for laying the ground and expliciting conditional least square estimators for the expectation of the reproduction and immigration sequences in the subcritical case. A central limit theorem for these estimators was later proved in Venkataraman (1982), using a time series point of view. Note that the link between such processes and the so called integer-valued times series INAR(1) processes has been exploited, see e.g. Al-Osh and Alzaid (1987) which studied such a process with particular distributions for the reproduction and immigration sequences. See also McKenzie (2003) who reviews the literature on models for integer-valued time series. These above processes play an important role in many scientific disciplines and applied fields such as epidemiology, economics, finance, to name a few. Note also that the link between such processes and the so called Hawkes processes introduced by Hawkes (1971a,b) as a model for contagious processes such as measles infections or hijackings has been exploited in the literature. See Daley et al. (2003) for a reference book that covers many aspects of the Hawkes process. For instance, Daley et al. (2003, Example 6.3(c)) show that each immigrant of Hawkes process has the potential to produce descendants whose numbers in successive generations constitute a subcritical Galton–Watson branching

Email addresses: mailto:Yacouba.BoubacarMainassara@uphf.fr (Yacouba Boubacar Maïnassara), mailto:lrabehas@univ-fcomte.fr (Landy Rabehasaina)

process with Poisson offspring distribution (see also Hawkes and Oakes (1974)). Recently, it was proved in Kirchner (2016, Theorem 3) that Hawkes prcesses can be approximated by integer-valued autoregressive models of infinite order (INAR(∞)) processes, which in turn are the limit of INAR(p) processes as $p \to \infty$ see Kirchner (2017, Proposition 2.5). A certain number of extensions for the model were later devised and studied. Wei and Winnicki (1990) considered the general critical and supercritical case and proved central limit theorems for (modified) weighted least square estimators. In Barczy et al. (2021), the specific case where the immigration sequence has a regular variation distribution is considered, leading to asymptotic normality of the reproduction mean when the immigration mean is known. Generalization to two types processes have been recently investigated in Ispány et al. (2014); Körmendi and Pap (2018), including estimations for the criticality parameter. Note that these references have one of the following constraint on the immigration process: first, some of them assume that its expectation is known (and appears in the expression of the estimator for the reproduction sequence); second, this immigration process is assumed to be a sequence of *independent* random variables, in addition to be identically distributed.

We aim in this paper at considering a process which is stationary, but where the immigration process is dependent. To our knowledge, this particular feature appears not to have been studied in the literature. Thus, there are two major contributions in this work. The first one is the relaxing of the standard independence assumption on the immigration process in order to extend considerably the range of application of the Galton Watson models and to show that the moment estimation procedure can be extended to the Galton Watson models with correlated immigration to estimate the unknown mean reproduction and immigration parameters λ_0 and m_0. This goal is achieved thanks to Proposition 4.1 and Theorem 4.2 in which the consistency and the asymptotic normality are stated when the immigration sequence is no longer correlated from a certain instant. Still in this context, another contribution is to propose a weakly consistent estimator of the unknown asymptotic variance matrix (see Theorem 5.2), which may be very different from that obtained in the standard framework (see for instance Klimko and Nelson (1978, Theorem 2.2) or Al-Osh and Alzaid (1987, Section 4.2)). Thanks to this estimation of the asymptotic variance matrix, we can for instance construct a confidence region for the estimation of the parameters (the reproduction and mean immigration rates). The second major contribution is when the immigration process has general correlation structure: we exhibit an estimator for the logarithm of the reproduction rate and we prove that it converges in quadratic mean with explicit speed at most of the order $1/\ln(n)$ (see Theorem 6.1). Convergence is thus very slow, which in particular explains why we need to pick very large n in the numerical illustrations. Under mildly stronger assumption, we provide and prove a two terms expansion for the estimator (see Theorem 6.2) and thus deduce that the speed of convergence is exactly of the order $1/\ln(n)$.

The paper is organized as follows. Section 2 presents the Galton Watson process that we consider here. A series of Propositions and Lemmas which are going to be repeatedly used are presented in Section 3. It is shown in Section 4 that the moment estimators of the unknown mean reproduction and immigration parameters λ_0 and m_0 are consistent and asymptotically normally distributed when the immigration sequence is no longer correlated from a certain instant and satisfies mild mixing assumptions. The proof relies on using Herrnorf's result (see Herrndorf (1984)) that involves a truncated expansion of the stationary Galton Watson with correlated immigration process. These results use tools used previously in the time series literature in the case of uncorrelated but non-independent error terms (*i.e.* weak white noise), see Francq and Zakoïan (1998), except for some major differences: first, the model is

different from the one in a time series setting and requires a different approach for determining the estimators, and second, the asymptotic normality concerns estimators which are different from least square estimators in Francq and Zakoïan (1998). The asymptotic variance of the moment estimators may be very different in the correlated and independence immigration cases. The estimation of this unknown asymptotic covariance matrix is done in Section 5. We consider in Section 6 a more general case of correlated immigration process. Here the approach is completely novel to our knowledge. The core results of the section are the quadratic convergence of the estimator as well as a two terms stochastic expansion for the estimator of $\ln \lambda_0$. The ingredients for the proofs relies on fine estimates for the estimator coupled with a central limit theorem for mixing triangular arrays of variables whose dependence is allowed to grow with the sample size proved in Francq and Zakoïan (2005). Note that in the aforementioned results all estimators are determined from the observation of the base Galton Watson process only; in particular, the expectation of the immigration is unknown, and actually also estimated. Numerical illustrations are given in Section 7. Finally, proofs of intermediate technical results are in Appendix Appendix A.

2. Model, notation and Assumptions

We consider the following Galton Watson process with immigration as the stationary sequence $\{Y_n,\ n \in \mathbb{Z}\}$ satisfying

$$Y_{n+1} = \sum_{k=1}^{Y_n} \xi_{n+1,k} + \epsilon_{n+1}, \quad n \in \mathbb{Z}, \tag{1}$$

for some sequences $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$ and $\{\epsilon_n,\ n \in \mathbb{Z}\}$, named thereafter the *reproduction* sequence and the *immigration* sequence, which are such that

- $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$ and $\{\epsilon_n,\ n \in \mathbb{Z}\}$ are independent sequences,
- the reproduction sequence $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$ is an i.i.d. (doubly indexed) sequence, with distribution of a generic r.v. denoted by ξ,
- the immigration process $\{\epsilon_n,\ n \in \mathbb{Z}\}$ is stationary and ergodic, with distribution that of a generic r.v. denoted by ϵ,
- the moments $\lambda_0 := \mathbb{E}(\xi)$ and $m_0 := \mathbb{E}(\epsilon)$ of $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$ and $\{\epsilon_n,\ n \in \mathbb{Z}\}$ are unknown.

The aim of the paper is devoted to the problem of estimating the unknown reproduction rate λ_0 as well as the mean immigration m_0 from the observed sequence $\{Y_n,\ n \in \mathbb{Z}\}$.

Furthermore, in order for the existence of the stationary process $\{Y_n,\ n \in \mathbb{Z}\}$ to exist and be unique, it is required that the subcritical case holds, namely that

$$\lambda_0 < 1. \tag{2}$$

To control the serial dependence of the stationary process $\{\epsilon_n,\ n \in \mathbb{Z}\}$, we introduce the strong mixing coefficients $\alpha_\epsilon(h)$ defined by

$$\alpha_\epsilon(h) = \sup_{A \in \mathcal{F}_{-\infty}^{n}, B \in \mathcal{F}_{n+h}^{\infty}} \left| \mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B) \right|,$$

where $\mathcal{F}_{-\infty}^{n} = \sigma(\epsilon_u, u \leq n)$ and $\mathcal{F}_{n+h}^{\infty} = \sigma(\epsilon_u, u \geq n+h)$. Note that $\alpha_\epsilon(h)$ does not depend on $n \in \mathbb{Z}$ thanks to the stationarity of $\{\epsilon_n,\ n \in \mathbb{Z}\}$.

To establish the asymptotic properties of the proposed estimator of λ_0, the following assumptions are required.

(A1): The mixing coefficient $\alpha_\epsilon(\cdot)$ verifies the following summability condition. There exists $\beta > 2$ such that

$$\sum_{h=0}^{\infty} \lambda_0^{-h} \alpha_\epsilon(h)^{1-2/\beta} < \infty.$$

We next make an integrability assumption on the moments and covariances of the immigration process $\{\epsilon_n,\ n \in \mathbb{Z}\}$ and the reproduction sequence $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$. We use $\|\cdot\|$ to denote the Euclidean norm of a vector and for any (potentially matrix valued) random variable X, we will set $\|X\|_p^p := \mathbb{E}\|X\|^p$ its $\mathbb{L}^p$ norm, with $p \geq 1$.

(A2): The following moment conditions hold:

$$\|\epsilon\|_{2\beta} = \left[\mathbb{E}\,|\epsilon|^{2\beta}\right]^{1/(2\beta)} < \infty \quad \text{and} \quad \|\xi\|_{2\beta} = \left[\mathbb{E}\,|\xi|^{2\beta}\right]^{1/(2\beta)} < \infty.$$

(A3): The covariance of the immigration process $\nu_h := \mathrm{Cov}(\epsilon_0, \epsilon_h)$ verifies:

$$\sum_{h=0}^{\infty} h\lambda_0^{-h}\,|\nu_{h+1}| < \infty.$$

Assumption (**A3**) may be removed if we replace (**A1**) by $\sum_{h=0}^{\infty} h\lambda_0^{-h}\alpha_\epsilon(h)^{1-2/\beta} < \infty$ thanks to Davydov's inequality (see Davydov (1968, Inequality (2.2))) and (**A2**). Also note that **(A1)** implies that the mixing coefficient $\alpha_\epsilon(\cdot)$ decreases at least exponentially, with decay rate depending on the unknown parameter λ_0. This unknown decay rate may be problematic because **(A1)** may thus not be easy to verify it in practice. Note that, in the case when one imposes λ_0 to lie in a known interval $[\lambda_-, \lambda_+]$ where $0 < \lambda_- < \lambda_+ < 1$ (as is the case later on in Section 6), then **(A1)** can be replaced by the stronger assumption $\sum_{h=0}^{\infty} \lambda_-^{-h}\alpha_\epsilon(h)^{1-2/\beta} < \infty$ which may be easier to check in practice.

We need to introduce the following notation in the sequel:

$$u_k := \mathbb{E}(Y_0 Y_{k+1}) = \mathbb{E}(Y_1 Y_{-k}), \quad k \in \mathbb{N}, \tag{3}$$

$$f_\nu(x) := \sum_{j=0}^{\infty} \nu_{j+1} x^j, \quad |x| \leq \lambda_0^{-1} \tag{4}$$

$$\chi_k := -\frac{1}{\lambda_0^k} \sum_{j=k}^{\infty} \lambda_0^j \nu_{j+1}, \quad k \in \mathbb{N}, \tag{5}$$

Note that the existence of χ_k defined in (5) is an easy consequence of the convergence of the series in Assumption (**A3**), and that f_ν is a power series with convergence radius larger than λ_0^{-1} for the same reason. Observe also that $\chi_0 = -f_\nu(\lambda_0)$. Furthermore, it will be proved in Proposition 3.1 that, under Condition (2) and Assumption (**A2**), the process $\{Y_n,\ n \in \mathbb{Z}\}$ is

a second-order stationary process of which the first and second-order moments are finite and defined by:

$$C_1 := \mathbb{E}(Y_0), \quad C_2 := \mathbb{E}(Y_0^2). \tag{6}$$

Furthermore, taking the expectation in (1) with $n = 0$ and using the independence of $\{\xi_{1,j},\ j \in \mathbb{N}\}$'s and Y_0 imply the explicit expression of the first order moment of the process $\{Y_n,\ n \in \mathbb{Z}\}$:

$$C_1 = \frac{m_0}{1 - \lambda_0}. \tag{7}$$

The expression of C_2 is also available and will be derived later on in Lemma 3.1.

3. Preliminary results

3.1. Expansion for Y_n

Let us introduce for all $n \in \mathbb{Z}$ the random operator $\theta_n : \mathbb{N} \longrightarrow \mathbb{N}$ defined as

$$\theta_n \circ k := \sum_{i=1}^{k} \xi_{n,i}$$

for all $k \in \mathbb{N}$. We note that θ_n depends on the $\xi_{n,i}$, $i \in \mathbb{N}$, so that the operators $(\theta_n)_{n\in\mathbb{N}}$ are independent. Following Kevei and Wiandt (2021, Relation (4)), iteration of (1) yields the representation for the stationary distribution

$$Y_0 = \sum_{i=0}^{\infty} \theta^{(i)} \circ \epsilon_{-i} \tag{8}$$

provided that the series converges, where $\theta^{(i)}$, $i \in \mathbb{N}$, are independent random operators, such that $\theta^{(0)} = \mathrm{Id}$ and $\theta^{(i)} \overset{\mathcal{D}}{=} \theta_0 \circ \theta_{-1} \circ ... \circ \theta_{-i+1}$, $i \geq 1$, and independent from the immigration sequence $\{\epsilon_i,\ i \in \mathbb{Z}\}$. Even more generally and more precisely, let us introduce a family of i.i.d. Galton Watson processes $\{\{\theta_{j,j'}(i),\ i \in \mathbb{N}\},\ (j, j') \in \mathbb{N}^2\}$ starting from $\theta_{j,j'}(0) = 1$ and with same offspring distribution as $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$, and independent from the immigration process $\{\epsilon_n,\ n \in \mathbb{Z}\}$. We may write

$$Y_n = \sum_{i=0}^{\infty} \theta_n^{(i)} \circ \epsilon_{n-i}, \quad n \in \mathbb{Z}, \tag{9}$$

where $\{\theta_n^{(i)},\ i \in \mathbb{N},\ n \in \mathbb{N}\}$ are a set of operators defined from $\mathbb{N}$ to $\mathbb{N}$, written as

$$\theta_n^{(i)} \circ k := \sum_{j=1}^{k} \theta_{n-i,j}(i). \tag{10}$$

This way, the term $\theta_{n-i,j}(i)$ in $\theta_n^{(i)} \circ \epsilon_{n-i} = \sum_{j=1}^{\epsilon_{n-i}} \theta_{n-i,j}(i)$ in (9) may be interpreted as the number of descendants at time n of the jth immigrants arrived at time $n - i$, $j \in \{1, ...\epsilon_{n-i}\}$ and $i \in \mathbb{N}$. Note that the representation (10) in particular implies the two following important facts. First, the operators $\theta_n^{(i)}$, $i \in \mathbb{N}$, $n \in \mathbb{Z}$, are independent from the immigration sequence $\{\epsilon_i,\ i \in \mathbb{Z}\}$. Second, for all n and r in $\mathbb{Z}$ and $i \geq 0$, $j \geq 0$, $\theta_n^{(i)}$ is independent from $\theta_r^{(j)}$ as soon as $n - i \neq r - j$. This latter feature will be extensively used later on, particularly in the proofs of the central limit theorems (see upcoming Theorem 4.2 and Proposition 6.4).

Proposition 3.1 (Stationary distribution and existence of moments)**.** The stationary version of the model described by (1) exists and admits moments of order 2β under Assumption (**A2**).

The proof of Proposition 3.1 is postponed to Section Appendix A.1.

Lemma 3.1. Let us define $v_k := \mathbb{E}(\epsilon_1 Y_{-k}) = \mathbb{E}(\epsilon_{k+1} Y_0)$ for all $k \in \mathbb{N}$. Its explicit expression is given by

$$v_k = u_k - \lambda_0 u_{k-1} = \frac{m_0^2}{1-\lambda_0} - \chi_k, \quad k \in \mathbb{N} \tag{11}$$

where u_k and χ_k are given by (3) and (5). As a consequence, the second moment of Y_0 is given by

$$C_2 = \frac{1}{1-\lambda_0^2}\left[\frac{V_{0,\xi} m_0}{1-\lambda_0} + 2\lambda_0\left(\frac{m_0^2}{1-\lambda_0} + f_\nu(\lambda_0)\right) + \mathbb{E}\epsilon_1^2\right] \tag{12}$$

where $V_{0,\xi} := \mathrm{Var}(\xi)$.

The proof of Lemma 3.1 is postponed to Section Appendix A.2.

3.2. Correlations of the process $\{Y_n,\ n \in \mathbb{Z}\}$

We next study properties related to the asymptotic behaviour of $\mathrm{Cov}(Y_0, Y_n)$. The following result, proved in Appendix Appendix A.3, shows that the correlation of the process $\{Y_n,\ n \in \mathbb{Z}\}$ converges to 0 exponentially fast.

Corollary 3.2. One has the following expression

$$C_1^2 - u_k = \lambda_0^k \left(\sum_{j=0}^{k} \lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right). \tag{13}$$

In particular, one has that $\mathrm{Cov}(Y_0, Y_{k+1}) = u_k - C_1^2 = \mathrm{O}(\lambda_0^k)$ as $k \to \infty$.

One deduces from (13) the asymptotic equivalent $C_1^2 - u_k \sim \lambda_0^k \left(\sum_{j=0}^{\infty} \lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right)$ as $k \to \infty$, when the quantity $\sum_{j=0}^{\infty} \lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)$ is not zero. The latter quantity will be studied later on, hence it will be interesting to write this quantity in function of the parameters m, λ_0 and the function f_ν defined in (4). This is done as follows. Using Fubini and (5), we first obtain that

$$\begin{aligned}\sum_{j=0}^{\infty} \lambda_0^{-j}\chi_j &= -\sum_{j=0}^{\infty} \lambda_0^{-2j} \sum_{r=j}^{\infty} \lambda_0^r \nu_{r+1} = -\sum_{r=0}^{\infty} \lambda_0^r \left(\sum_{j=0}^{r} \lambda_0^{-2j}\right) \nu_{r+1} \\ &= -\sum_{r=0}^{\infty} \lambda_0^2 \frac{\lambda_0^r - \lambda_0^{-(r+2)}}{\lambda_0^2 - 1} \nu_{r+1} = \frac{\lambda_0^2}{1-\lambda_0^2}\left[f_\nu(\lambda_0) - \lambda_0^{-2} f_\nu(\lambda_0^{-1})\right]. \end{aligned} \tag{14}$$

Moreover, the expressions (7) and (12) of C_1 and C_2 the first and second moments of $\{Y_n,\ n \in \mathbb{Z}\}$ entail that

$$C_1^2 - C_2 = \frac{m_0^2}{(1-\lambda_0)^2} - \frac{1}{1-\lambda_0^2}\left[\frac{V_{0,\xi} m_0 + 2\lambda_0 m_0^2}{1-\lambda_0} + 2\lambda_0 f_\nu(\lambda_0) + V_{0,\epsilon} + m_0^2\right]$$

$$= -\frac{V_{0,\xi} m_0}{(1-\lambda_0^2)(1+\lambda_0)} - \frac{2\lambda_0}{1-\lambda_0^2} f_\nu(\lambda_0) - \frac{V_{0,\epsilon}}{1-\lambda_0^2},$$

by using $\mathbb{E}\epsilon_1^2 = V_{0,\epsilon} + m_0^2$ with $V_{0,\epsilon} := \mathrm{Var}(\epsilon)$. Combining the above expression and (14), yields

$$\sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0 (C_1^2 - C_2) = -\frac{V_{0,\xi} m_0 \lambda_0}{(1-\lambda_0^2)(1+\lambda_0)} - \frac{V_{0,\epsilon}\lambda_0}{1-\lambda_0^2} - \frac{\lambda_0^2}{1-\lambda_0^2} f_\nu(\lambda_0) - \frac{1}{1-\lambda_0^2} f_\nu(\lambda_0^{-1}). \quad (15)$$

We also mention that, still when $\sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0(C_1^2 - C_2)$ is not zero, (13) yields the following interesting limit

$$\lambda_0 = \lim_{k\to\infty} \frac{C_1^2 - u_k}{C_1^2 - u_{k-1}},$$

which is the starting point for exhibiting estimators for λ_0 in the following Section.

4. Ultimately uncorrelated immigration

We suppose in this section that Assumptions **(A1)**, **(A2)** and that $(\mathbf{A3})_\mathbf{1}$ defined thereafter hold.

$(\mathbf{A3})_\mathbf{1}$: $\exists k_0 \in \mathbb{N} \setminus \{0\}$ such that $\nu_k = 0$ for all $k \geq k_0$, where the instant k_0 is known.

In other words, the above assumption means that the immigration is no more correlated from an instant k_0. Also note that $(\mathbf{A3})_\mathbf{1}$ is stronger than **(A3)**.

Example 4.1. Let $\{Z_n,\ n \in \mathbb{Z}\}$ be any sequence of iid discrete (for instance: Poisson, Bernoulli, Binomial,...). For fixed $k_0 > 0$, let us assume that those random variables have finite 2β moments and let

$$\epsilon_n := \prod_{i=0}^{k_0-1} Z_{n-i}. \quad (16)$$

We may alternatively assume that they have 4β moments and let $\epsilon_n := Z_n^2 \prod_{i=1}^{k_0-1} Z_{n-i}$. Thus we have:

$$\nu_h = \begin{cases} \mathrm{Cov}\,(\epsilon_n, \epsilon_{n-h}) \neq 0 & \text{if} \quad h \leq k_0 - 1, \\ \mathrm{Cov}\,(\epsilon_n, \epsilon_{n-h}) = 0 & \text{if} \quad h \geq k_0. \end{cases} \quad (17)$$

Then the ϵ_n form an (k_0-1)-dependent sequence; that is, the mixing coefficients of the ϵ_n sequence satisfies $\alpha_\epsilon(h) = 0$ if $h > k_0 - 1$, and the assumptions **(A1)**, **(A2)** and $(\mathbf{A3})_\mathbf{1}$ are satisfied.

Example 4.2. The previous example may be generalized when $k_0 = 2$ as follows. We suppose that we are in a setting where $\{Y_n,\ n \in \mathbb{Z}\}$ described by (1) represents the number of persons infected by an illness. We consider an exogenous system that provides the contaminated immigrants $\{\epsilon_n,\ n \in \mathbb{Z}\}$ described as follows. This system consists of contaminated individuals pertaining to two classes, namely the class of contagious individuals (Class 1) and the class of non contagious individuals (Class 2). We are endowed with a family of $\mathbb{N} \setminus \{0\}$ valued, independent and identically distributed processes $\{(\zeta_{n,k})_{k\in\mathbb{N}\setminus\{0\}},\ n \in \mathbb{Z}\}$, with the assumption that $\sup_{k\in\mathbb{N}\setminus\{0\}} \|\zeta_{n,k}\|_{2\beta} < \infty$. Note that we do not assume necessarily that $(\zeta_{n,k})_{k\in\mathbb{N}\setminus\{0\}}$ is i.i.d. for fixed $n \in \mathbb{Z}$. We suppose that at time $n-1$, the k-th individual, from Class 1 contaminate $\zeta_{n,1}$ individuals of Class 1 if $k = 1$ and $\zeta_{n,k}$ individuals of Class 2 if $k \in$

$\{2, ..., \zeta_{n-1,1}\}$. Soon after contaminating those individuals, they emigrate to our system and we then set

$$\epsilon_n := \sum_{k=1}^{\zeta_{n-1,1}} \zeta_{n,k}, \quad n \in \mathbb{Z}. \tag{18}$$

An illustration of the evolution of individuals from different classes is provided in Figure 1: the Class 1 individuals are represented by crosses, the Class 1 individuals that contaminate Class 1 individuals are represented by boxed crosses, while the Class 2 individuals are represented by dots. Then, the sequence $\{\epsilon_n,\ n \in \mathbb{Z}\}$ given by (18) satisfies the k_0 dependence relation

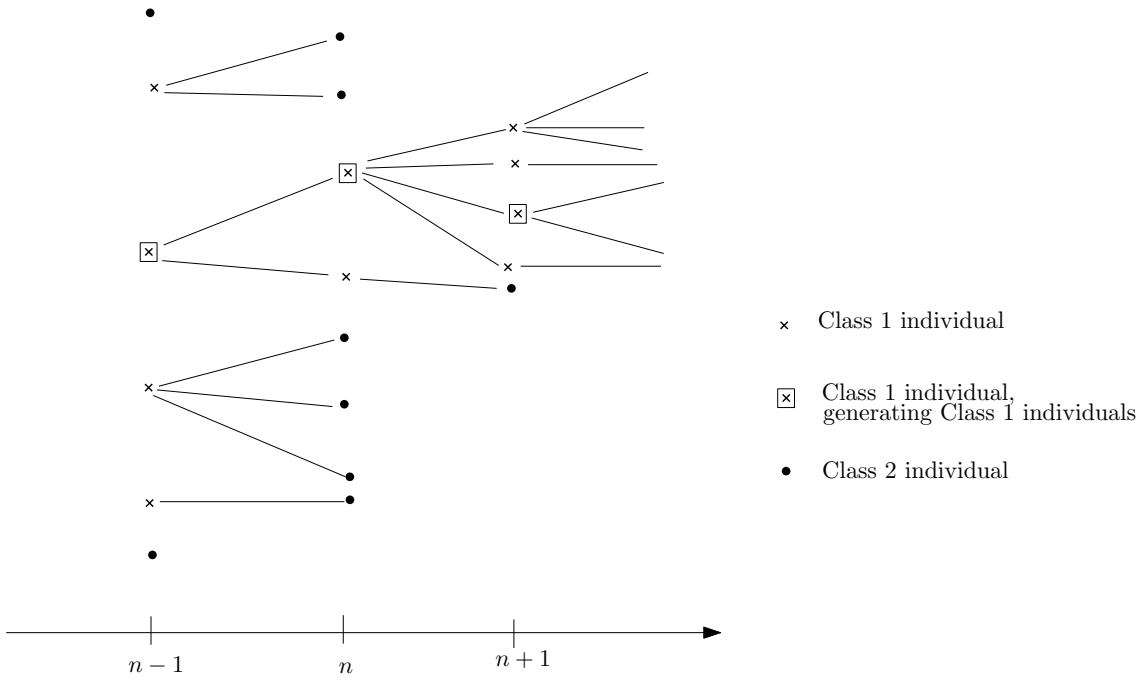

Figure 1: Infected immigrants $\{\epsilon_n,\ n \in \mathbb{Z}\}$

(17) with $k_0 = 2$. Also, we note that ϵ_n in Example 4.1 satisfies (18) with $\zeta_{n,k} = Z_n$, which is independent from k.

Under the assumption $(\mathbf{A3})_\mathbf{1}$, it is interesting to notice that χ_k defined in (5) verifies $\chi_k = 0$ for all $k \geq k_0 - 1$, so that we obtain from (13) the following moment expression for the unknown reproduction rate:

$$\lambda_0 = \frac{C_1^2 - u_{k_0-1}}{C_1^2 - u_{k_0-2}}. \tag{19}$$

For all $k \in \mathbb{N}$, we introduce the following notation:

$$\bar{Y}_n := \frac{1}{n}\sum_{j=1}^{n} Y_j,\ \bar{Y}_{k,n} := \frac{1}{n}\sum_{j=1}^{n} Y_j Y_{j+k},\ S_n(k) := \sum_{j=1}^{n} X_j(k),\ X_j(k) := (Y_j - C_1, Y_j Y_{j+k+1} - u_k)'$$

$$T_n(k) := \sum_{j=1}^{n} \mathcal{Y}_j(k), \quad \mathcal{Y}_j(k) := (Y_j - C_1, Y_j Y_{j+k-1} - u_{k-2}, Y_j Y_{j+k} - u_{k-1})'. \tag{20}$$

Note that (19) is the starting point for the estimation of the unknown parameter. The following consistance result holds.

Proposition 4.1 (Consistency)**.** Assume that (**A1**), (**A2**) and (**A3**)$_\mathbf{1}$ hold. The following estimators

$$\hat{R}_{k_0,n} := \frac{\left[\bar{Y}_n\right]^2 - \bar{Y}_{k_0,n}}{\left[\bar{Y}_n\right]^2 - \bar{Y}_{k_0-1,n}}, \quad \hat{M}_{k_0,n} := \bar{Y}_n(1-\hat{R}_{k_0,n}) \tag{21}$$

converge a.s. towards the unknown parameters λ_0 and m_0 as $n \to \infty$.

The proof of Proposition 4.1 is postponed to Appendix Appendix A.4.

Remark 4.1. Note that, when $k_0 = 1$, the estimator $\hat{R}_{k_0,n}$ in Proposition 4.1 agrees with those in the case where the immigration process $\{\epsilon_n,\ n \in \mathbb{Z}\}$ is i.i.d., see Klimko and Nelson (1978, Expression (5.1)), Al-Osh and Alzaid (1987, Expression (4.1)) and Wei and Winnicki (1990, Section 1).

The following theorem states the asymptotic distributions of the process $n^{-1/2}T_n(k_0)$, where we recall that $T_n(k_0)$ is defined in (20).

Theorem 4.2. The following central limit theorem holds:

$$\frac{T_n(k_0)}{\sqrt{n}} = \sqrt{n}\left[\bar{Y}_n - C_1, \bar{Y}_{k_0-1,n} - u_{k_0-2}, \bar{Y}_{k_0,n} - u_{k_0-1}\right]' \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \mathfrak{S}_{k_0}), \tag{22}$$

for some (non zero) semi-definite positive matrix $\mathfrak{S}_{k_0} \in \mathbb{R}^{3\times 3}$ given by

$$\mathfrak{S}_{k_0} := \sum_{j=-\infty}^{\infty} \gamma_j(k_0), \quad \text{where} \tag{23}$$

$$\begin{aligned}\gamma_j(k_0) &:= \mathbb{E}\left(\mathcal{Y}_0(k_0)\mathcal{Y}_j(k_0)'\right) \\ &= \begin{bmatrix} \mathrm{Cov}(Y_0, Y_j) & \mathrm{Cov}(Y_0, Y_jY_{j+k_0-1}) & \mathrm{Cov}(Y_0, Y_jY_{j+k_0}) \\ \mathrm{Cov}(Y_0, Y_jY_{j+k_0-1}) & \mathrm{Cov}(Y_0Y_{k_0-1}, Y_jY_{j+k_0-1}) & \mathrm{Cov}(Y_0Y_{k_0-1}, Y_jY_{j+k_0}) \\ \mathrm{Cov}(Y_0, Y_jY_{j+k_0}) & \mathrm{Cov}(Y_0Y_{k_0-1}, Y_jY_{j+k_0}) & \mathrm{Cov}(Y_0Y_{k_0}, Y_jY_{j+k_0}) \end{bmatrix}, \ j \in \mathbb{N}.\end{aligned} \tag{24}$$

The proof of Theorem 4.2 is postponed to Appendix Appendix A.5.

Let us now define the open set $\mathcal{O} := \{(a,b,c) \in \mathbb{R}^3 |\ a^2 \neq b\} \subset \mathbb{R}^3$ as well as the following function

$$\varphi : (a,b,c) \in \mathcal{O} \mapsto \left(\frac{a^2-c}{a^2-b},\ a\left(1 - \frac{a^2-c}{a^2-b}\right)\right)' \in \mathbb{R}^2, \tag{25}$$

so that $(\lambda_0, m_0)' = \varphi(C_1, u_{k_0-2}, u_{k_0-1})$ thanks to Relations (7) and (19).

We are now able to state the following proposition, which shows the asymptotic normality of the estimators proposed in Proposition 4.1. The δ-method as well as Theorem 4.2 thus yield the main following result.

Theorem 4.3 (Asymptotic normality)**.** Assume that (**A1**), (**A2**) and (**A3**)$_\mathbf{1}$ hold. The following central limit theorem holds:

$$\sqrt{n}[\hat{R}_{k_0,n} - \lambda_0, \hat{M}_{k_0,n} - m_0]' \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \Omega_{k_0}), \tag{26}$$

where $\Omega_{k_0} = \nabla\varphi(C_1, u_{k_0-2}, u_{k_0-1})'\mathfrak{S}_{k_0}\nabla\varphi(C_1, u_{k_0-2}, u_{k_0-1})$ with φ and $\mathfrak{S}_{k_0}$ are defined in (25) and in Theorem 4.2.

Remark 4.2. In the independent immigration case, Klimko and Nelson (1978, Theorem 2.2) show that the asymptotic covariance matrix is reduced as

$$\Omega_S := J^{-1}WJ^{-1}, \quad \text{where}$$

$$J := \mathbb{E}\left[\frac{\partial g(\theta_0, Y_{n-1})}{\partial \theta}\frac{\partial g(\theta_0, Y_{n-1})}{\partial \theta'}\right], \quad W := \mathbb{E}\left[\{Y_n - g(\theta_0, Y_{n-1})\}^2 \frac{\partial g(\theta_0, Y_{n-1})}{\partial \theta}\frac{\partial g(\theta_0, Y_{n-1})}{\partial \theta'}\right]$$

with $g(\theta_0, Y_{n-1}) = Y_n - \lambda_0 Y_{n-1}$ and $\theta_0 = (\lambda_0, m_0)'$. See also Al-Osh and Alzaid (1987, Section 4.2). The true asymptotic covariance matrix $\Omega_{k_0} = \nabla\varphi(C_1, u_{k_0-2}, u_{k_0-1})'\mathfrak{S}_{k_0}\nabla\varphi(C_1, u_{k_0-2}, u_{k_0-1})$ obtained in the correlated immigration framework can be very different from Ω_S and can yield inacurrate estimation results when used on a model with correlated immigration sequence, as illustrated later on in Section 7.1. This is why it is interesting to find an estimator of Ω_{k_0} which is consistent for both independent and correlated immigration cases.

5. Estimating the asymptotic variance matrix

For statistical inference problem, the asymptotic variance Ω_{k_0} has to be estimated. In particular Theorem 4.3 can be used to obtain confidence intervals.

The estimation of the long-run variance (LRV) matrix expressed by (23) and given by

$$\mathfrak{S}_{k_0} = \sum_{j=-\infty}^{\infty} \mathrm{Cov}\left(\mathcal{Y}_0(k_0), \mathcal{Y}_j(k_0)\right)$$

is complicated. In the literature, two types of estimators are generally employed: heteroskedasticity and autocorrelation consistent (HAC) estimators based on kernel methods (see Newey and West (1987) and Andrews (1991) for general references, and Francq and Zakoïan (2007) for an application to testing strong linearity in ARMA models with dependent errors) and the spectral density estimators (see e.g. Berk (1974) and den Haan and Levin (1997) for a general references; see also Boubacar Mainassara and Francq (2011) for an application to a multivariate ARMA (VARMA) model with dependent errors). In the present paper, we focus on an estimator based on a spectral density form.

Following the arguments developed in Boubacar Mainassara et al. (2012), the matrix $\mathfrak{S}_{k_0}$ can be estimated using Berk's approach (see Berk (1974)). More precisely, by interpreting $\mathfrak{S}_{k_0}/2\pi$ as the spectral density of the stationary process $\{\mathcal{Y}_n = (\mathcal{Y}_n^1, \mathcal{Y}_n^2, \mathcal{Y}_n^3)', \ n \in \mathbb{Z}\} := \{\mathcal{Y}_n(k_0), \ n \in \mathbb{Z}\}$ evaluated at frequency 0, we can use a parametric autoregressive estimate of the spectral density of $(\mathcal{Y}_n)_{n\in\mathbb{Z}}$ in order to estimate the matrix $\mathfrak{S}_{k_0}$. This is explained more precisely in the following.

First, the stationary process $\{\mathcal{Y}_n, \ n \in \mathbb{Z}\}$ is in $\mathbb{L}^2$ hence admits the following Wold decomposition

$$\mathcal{Y}_n = \varepsilon_n + \sum_{k=1}^{\infty} \Psi_k \varepsilon_{n-k}, \quad n \in \mathbb{Z}, \tag{27}$$

where $(\varepsilon_n)_{n\in\mathbb{Z}}$ is a 3−variate uncorrelated but not independent white noise with variance matrix denoted by Σ_ε. We recall e.g. from Brockwell and Davis (1991, Theorem 5.7.1 p.187) that $\varepsilon_n \in \mathbb{R}^{3\times 1}$ and $\Psi_n \in \mathbb{R}^{3\times 3}$ are expressed in function of $\{\mathcal{Y}_n, \ n \in \mathbb{Z}\}$ as follows

$$\begin{array}{rcl} \varepsilon_n & = & \mathcal{Y}_n - P_{n-1}\mathcal{Y}_n, \quad n \in \mathbb{Z}, \\ \Psi_n & = & \mathbb{E}(\mathcal{Y}_0 \varepsilon'_{-n})\Sigma_\varepsilon^{-1} = \mathbb{E}(\mathcal{Y}_n \varepsilon'_0)\Sigma_\varepsilon^{-1}, \quad n \geq 0, \end{array} \tag{28}$$

where $P_{n-1}\mathcal{Y}_n$ is the $\mathbb{L}^2$ projection of the random vector $\mathcal{Y}_n$ onto the vector space spanned by $\mathcal{Y}_k,\ k \leq n-1$. It may be verified that Σ_ε is indeed non-singular in the case when the reproduction sequence $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$ is non deterministic. We next introduce

$$\Psi(z) := \sum_{k=0}^{\infty} \Psi_k z^k, \tag{29}$$

for $z \in \mathbb{C}$ where the series converges, and where Ψ_0 is the identity matrix. We suppose in this section that Assumptions (**A1**), $(\mathbf{A2})_\mathbf{1}$, (**A3**) and (**A4**) hold where

$(\mathbf{A2})_\mathbf{1}$: The following moment condition holds

$$\|\epsilon\|_{8\vee(2\beta)} < \infty \quad \text{and} \quad \|\xi\|_{2\beta} < \infty.$$

(**A4**): The sequence $\{\Psi_k,\ k \geq 1\}$ verifies that $\sum_{k=1}^{\infty} \|\Psi_k\| < \infty$ and

$$\det\left(\Psi(z)\right) \neq 0$$

for all $z \in \mathbb{C}$ such that $|z| \leq 1$.

Under Assumption (**A4**), $\{\mathcal{Y}_n,\ n \in \mathbb{Z}\}$ admits a multivariate AR(∞) representation (see Akutowicz (1957, Section 2)) with dependent errors of the form

$$\mathcal{Y}_n - \sum_{k=1}^{\infty} \Phi_k \mathcal{Y}_{n-k} = \varepsilon_n, \tag{30}$$

where $(\Phi_k)_{k\geq 1}$ are $\mathbb{R}^{3\times 3}$ matrices such that $\sum_{k=1}^{\infty} \|\Phi_k\| < \infty$ and $\det\left(\Phi(z)\right) \neq 0$, $|z| \leq 1$, where

$$\Phi(z) := I_3 - \sum_{k=1}^{\infty} \Phi_k z^k = \Psi(z)^{-1}. \tag{31}$$

In particular, (28) implies that the $\mathbb{L}^2$ projection $P_{n-1}\mathcal{Y}_n$ is in this case expressed as the series $\sum_{k=1}^{\infty} \Phi_k \mathcal{Y}_{n-k}$. One may argue that Assumption (**A4**) is quite complicated to verify in practice, and it may be informative and useful to give some assumption on the base model (1) such that it holds. This is provided in the following proposition (proved in Appendix Appendix A.6).

Proposition 5.1. Let us define $\underline{\epsilon}_n := (\epsilon_n, \epsilon_n\epsilon_{n+k_0-1}, \epsilon_n\epsilon_{n+k_0})'$, $n \in \mathbb{Z}$, and suppose that

- the spectral density $f_{\underline{\epsilon}} : x \mapsto \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \mathrm{Cov}(\underline{\epsilon}_0, \underline{\epsilon}_h)e^{ihx}$ is such that $f_{\underline{\epsilon}}(x)$ is invertible for all $x \in [0, 2\pi]$.
- (**A1**) is replaced by the stronger assumption

$$\alpha_\epsilon(n) = \mathrm{O}\left(\lambda_0^{\frac{n}{1-2/\beta}}\right). \tag{32}$$

Then (**A4**) is verified if λ_0 is small enough, and the sequence $\{\Phi_k,\ k \geq 1\}$ in the expansion (30) verifies $\|\Phi_k\| = \mathrm{o}(1/k^2)$ as $k \to \infty$.

Note that the proof of the above Proposition is technical and occupies most of Appendix Appendix A.6. What is essentially proved is that the spectral density is close to $f_{\underline{\epsilon}}$ when λ_0 is small enough.

Thanks to the previous remarks, the estimation of $\mathfrak{S}_{k_0}$ in (23) is therefore based on the following expression

$$\mathfrak{S}_{k_0} = \Phi^{-1}(1)\Sigma_\varepsilon \Phi^{-1'}(1) = \left[I_3 - \sum_{k=1}^{\infty} \Phi_k\right]^{-1} \Sigma_\varepsilon \left[I_3 - \sum_{k=1}^{\infty} \Phi_k\right]^{-1'} .$$

For $r \in \mathbb{N} \setminus \{0\}$, consider the least square regression of $\mathcal{Y}_n$ on $\mathcal{Y}_{n-1}, \ldots, \mathcal{Y}_{n-r}$, for all $n \in \mathbb{Z}$, defined by

$$\mathcal{Y}_n = \sum_{k=1}^{r} \Phi_{r,k} \mathcal{Y}_{n-k} + \varepsilon_{r,n}, \tag{33}$$

where $\varepsilon_{r,n}$ is uncorrelated with $\mathcal{Y}_{n-1}, \ldots, \mathcal{Y}_{n-r}$. Since $\mathcal{Y}_n$ is not observable, we introduce $\hat{\mathcal{Y}}_n \in \mathbb{R}^3$ obtained by replacing C_1 by $\bar{Y}_n$, u_{k_0-2} by $\bar{Y}_{k_0-1,n}$ and u_{k_0-1} by $\bar{Y}_{k_0,n}$ in (20):

$$\hat{\mathcal{Y}}_n := (Y_n - \bar{Y}_n, Y_n Y_{n+k_0-1} - \bar{Y}_{k_0-1,n}, Y_n Y_{n+k_0} - \bar{Y}_{k_0,n})' . \tag{34}$$

Let us define $\hat{\Phi}_r(z) := I_3 - \sum_{k=1}^{r} \hat{\Phi}_{r,k} z^k$, where $\hat{\Phi}_{r,1}, \ldots, \hat{\Phi}_{r,r}$ denote the coefficients of the least squares regression of $\hat{\mathcal{Y}}_n$ on $\hat{\mathcal{Y}}_{n-1}, \ldots, \hat{\mathcal{Y}}_{n-r}$. Let $\hat{\varepsilon}_{r,n}$ be the residuals of this regression so that $\hat{\mathcal{Y}}_n = \sum_{k=1}^{r} \hat{\Phi}_{r,k} \hat{\mathcal{Y}}_{n-k} + \hat{\varepsilon}_{r,n}$, and let $\hat{\Sigma}_{\hat{\varepsilon}_r} = \frac{1}{n} \sum_{t=1}^{n} \hat{\varepsilon}_{r,t} \hat{\varepsilon}'_{r,t}$ be the empirical variance of $\hat{\varepsilon}_{r,1}, \ldots, \hat{\varepsilon}_{r,n}$.

In the case of linear processes with independent innovations, Berk (1974) has shown that the spectral density can be consistently estimated by fitting AR models of order $r = r(n)$, whenever r tends to infinity and r^3/n tends to 0 as n tends to infinity. Here, there are differences with Berk (1974): $\{\mathcal{Y}_n,\ n \in \mathbb{Z}\}$ is multivariate, is not directly observed and is replaced by $\{\hat{\mathcal{Y}}_n,\ n \in \mathbb{Z}\}$. It is shown that this result remains valid for the multivariate linear process $\{\mathcal{Y}_n,\ n \in \mathbb{Z}\}$ with non-independent innovations (see for instance Boubacar Mainassara et al. (2012); Boubacar Mainassara and Francq (2011), for references in (multivariate) ARMA models with dependent errors).

Thus, the asymptotic study of the estimator of $\mathfrak{S}_{k_0}$ using the spectral density method is given in the following theorem.

Theorem 5.2. Assume that (**A1**) , $(\mathbf{A2})_\mathbf{1}$ hold with $\beta = 4 + 2\kappa$ for some $\kappa > 0$ and (**A3**) hold, and that the process $\{\mathcal{Y}_n,\ n \in \mathbb{Z}\}$ defined in (20) verifies (**A4**) with $\Sigma_\varepsilon = \mathrm{Var}(\varepsilon_n)$ non-singular. Assume in addition that the sequence $\{\Phi_k,\ k \geq 1\}$ defined in (30) verifies $\|\Phi_k\| = o(1/k^2)$. Then, the spectral estimator $\hat{\mathfrak{S}}_{k_0}^{\mathrm{SP}}$ of $\mathfrak{S}_{k_0}$ defined as follows verifies

$$\hat{\mathfrak{S}}_{k_0}^{\mathrm{SP}} := \hat{\Phi}_r^{-1}(1)\hat{\Sigma}_{\hat{\varepsilon}_r}\hat{\Phi}_r^{-1'}(1) \xrightarrow[n\to\infty]{\mathbb{P}} \mathfrak{S}_{k_0} = \Phi^{-1}(1)\Sigma_\varepsilon\Phi^{-1'}(1) \tag{35}$$

when $r = r(n) \to \infty$ and $r^3/n \to 0$ as $n \to \infty$.

As explained previously, Proposition 5.1 gives some conditions such that (**A4**) and $\|\Phi_k\| = o(1/k^2)$ are verified. Its proof occupies a large bulk of Appendix Appendix A.6. The proof of Theorem 5.2 is similar to that given by Boubacar Mainassara et al. (2012, Section 3.3.1) and is given in the same Appendix.

Let $\nabla\varphi(\bar{Y}_n, \bar{Y}_{k_0-1,n}, \bar{Y}_{k_0,n})$ be a consistent estimator of $\nabla\varphi(C_1, u_{k_0-2}, u_{k_0-1})$, where φ is the function defined by (25). Theorem 5.2 thus implies that

$$\hat{\Omega}_{k_0}^{\mathrm{SP}} := \left(\nabla\varphi(\bar{Y}_n, \bar{Y}_{k_0-1,n}, \bar{Y}_{k_0,n})\right)' \hat{\mathfrak{S}}_{k_0}^{\mathrm{SP}} \left(\nabla\varphi(\bar{Y}_n, \bar{Y}_{k_0-1,n}, \bar{Y}_{k_0,n})\right) \tag{36}$$

is a weakly consistent estimator of Ω_{k_0}. Some numerical illustrations are presented in Section 7.1.

6. General correlated immigration

We suppose throughout this section that (**A1**), (**A2**) and (**A3**) hold, as well as the following additional Assumptions.

(**A5**): The unknown parameters λ_0, m_0, $V_{0,\xi}$ and $V_{0,\epsilon}$ (reminding that the two latter are the variances of ξ and ϵ) belong to $\Theta := [\lambda_-, \lambda_+] \times \Theta_m \times \Theta_{V_\xi} \times \Theta_{V_\epsilon}$ for some known respective intervals $[\lambda_-, \lambda_+]$ and compact intervals Θ_m, Θ_{V_ξ} and Θ_{V_ϵ} included in $[0, \infty)$, with $0 < \lambda_- < \lambda_+ < 1$, and the generating function $f_\nu : x \in [0,1] \mapsto \sum_{s=0}^{\infty} x^s \nu_{s+1}$ belongs to some known class of function $\mathcal{F}$.

(**A6**): There exists a known quantity $K_m > 0$ such that

$$\inf_{(\lambda,\mu,V_\xi,V_\epsilon)\in\Theta,\ f\in\mathcal{F}} |\Xi(\lambda, \mu, V_\xi, V_\epsilon, f)| \geq K_m \tag{37}$$

where Ξ is the function defined on $\Theta \times \mathcal{F}$ by

$$\Xi(\lambda, \mu, V_\xi, V_\epsilon, f) := -\frac{V_\xi m_0 \lambda}{(1-\lambda^2)(1+\lambda)} - \frac{V_\epsilon \lambda}{1-\lambda^2} - \frac{\lambda^2}{1-\lambda^2} f(\lambda) - \frac{1}{1-\lambda^2} f(\lambda^{-1}) \tag{38}$$

where $(\lambda, \mu, V_\xi, V_\epsilon, f) \in \Theta \times \mathcal{F}$.

(**A7**): There exists a known constant $C_Y > 0$ such that the k-th moments $\|Y_0\|_k$ of the stationary process, $k = 1, 2$ are less than C_Y.

We give a few comments on the two last assumptions. Assumption (**A6**) may at first sight look artificial, but is actually natural for the following reason. First note that the function Ξ in (38) verifies from (15) that $\Xi(\lambda_0, m_0, V_{0,\xi}, V_{0,\epsilon}, f_\nu) = \sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0(C_1^2 - C_2)$ is non zero (a condition which will turn out to be necessary in the forthcoming Theorem 6.2), and that it is still non zero in the neighborhood $\Theta \times \mathcal{F}$ of the unknown parameters $(\lambda_0, m_0, V_{0,\xi}, V_{0,\epsilon}, f_\nu)$. So that (**A6**) is relevant because, unless in very specific circumstances, $\sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0(C_1^2 - C_2)$ is not likely to be exactly zero. This is verified for example in the following situations:

- When $\mathcal{F}$ is the set of power series with non negative coefficents, meaning that the immigration sequence $\{\epsilon_n,\ n \in \mathbb{Z}\}$ is positively correlated, i.e. $\nu_n \geq 0$ for all $n \geq 1$. This is the case when one models the evolution of a disease with constant (increasing or decreasing) trend. In that case we have that $f(\lambda)$ and $f(\lambda^{-1})$ are non negative, so that we have from (37) that

$$|\Xi(\lambda, \mu, V_\xi, V_\epsilon, f)| \geq K_m := \inf \Theta_{V_\xi} . \inf \Theta_m . \frac{\lambda_-}{(1-\lambda_-^2)(1+\lambda_+)} + \inf \Theta_{V_\epsilon} \frac{\lambda_-}{1-\lambda_-^2}.$$

- When $\mathcal{F}$ is included in the set of functions bounded by some constant $M_{\mathcal{F}}$, in which case we may set

$$K_m := \inf \Theta_{V_\xi} . \inf \Theta_m . \frac{\lambda_-}{(1-\lambda_-^2)(1+\lambda_+)} + \inf \Theta_{V_\epsilon} \frac{\lambda_-}{1-\lambda_-^2} - M_{\mathcal{F}} \left[\frac{\lambda_+^2}{1-\lambda_+^2} + \frac{1}{1-\lambda_+^2} \right],$$

  which is positive e.g. if the bound $M_{\mathcal{F}}$ is small enough.

As to (**A7**), note that the explicit upper bound C_Y in (**A7**) may be obtained from Kevei and Wiandt (2021, Section 3.1) under the assumptions that $\|\epsilon\|_2$ and $\|\xi\|_2$ are finite, so that (**A7**) is not too restrictive. The expression of the bound obtained in that latter paper may not be simple, which is why one simple example where C_Y has a nice expression may be illuminating. This is given in the following example:

Example 6.1. Assume that there exists known constants $M_\epsilon > 0$ and $M_\xi \in (0,1)$ such that $\|\epsilon\|_k = [\mathbb{E}(|\epsilon|^k)]^{1/k} \leq M_\epsilon$ and $\|\xi\|_k = [\mathbb{E}(|\xi|^k)]^{1/k} \leq M_\xi$, $k = 1, 2$. Let us then write that

$$\|Y_0\|_k = \left\| \sum_{j=1}^{Y_{-1}} \xi_{0,j} + \epsilon_0 \right\|_k \leq \left\| \sum_{j=1}^{Y_{-1}} \xi_{0,j} \right\|_k + \|\epsilon_0\|_k . \tag{39}$$

Thanks to the independence of Y_{-1} from the $\xi_{0,j}$'s, $j \geq 1$:

$$\begin{aligned} \left\| \sum_{j=1}^{Y_{-1}} \xi_{0,j} \right\|_k^k &= \mathbb{E}\left(\left[\sum_{j=1}^{Y_{-1}} \xi_{0,j} \right]^k \right) = \sum_{n=0}^{\infty} \left\| \sum_{j=1}^{n} \xi_{0,j} \right\|_k^k \mathbb{P}(Y_{-1} = n) \\ &\leq \sum_{n=0}^{\infty} n^k \|\xi\|_k^k \, \mathbb{P}(Y_{-1} = n) = \|\xi\|_k^k \, \|Y_{-1}\|_k^k = \|\xi\|_k^k \, \|Y_0\|_k^k \end{aligned}$$

which, plugged into (39) yields that the stationary process $\{Y_n,\ n \in \mathbb{Z}\}$ admits the following k-th moment upper bounds

$$\|Y_0\|_k \leq \frac{\|\epsilon\|_k}{1 - \|\xi\|_k} \leq \frac{M_\epsilon}{1 - M_\xi}, \quad k = 1, 2.$$

Here, the fact that we impose for M_ξ to be less than 1 may look demanding, but is satisfied for instance if ξ is zero with sufficiently high probability.

In order to state the main result of this section, we first introduce some auxiliary functions that will enable us to construct the estimator for the unknown parameters λ_0 and m_0. We first let $\varpi : \mathbb{R} \longrightarrow \mathbb{R}$ be a twice differentiable function such that

$$\varpi(x) \quad \begin{cases} = 1, & x \in [1, \infty), \\ = 0, & x \in (-\infty, 0], \\ \in [0,1], & x \in [0,1], \end{cases} \tag{40}$$

$$\varpi''(x) = \mathrm{o}(x), \quad x \to 0, \tag{41}$$

so that $\varpi'(x) = \mathrm{o}(x^2)$ and $\varpi(x) = \mathrm{o}(x^3)$ as $x \to 0$.

We also need a twice differentiable function $G : \mathbb{R} \longrightarrow \mathbb{R}$ with finite support such that $G(x) = 1$ for $x \in \left[0, \max\left(C_Y, C_Y^2\right)\right]$. We then let

$$\varpi_k : x \in \mathbb{R} \mapsto \varpi\left(\frac{2}{K_m \lambda_-^k} x\right), \quad k \in \mathbb{N}, \tag{42}$$

K_m being defined in (37), so that $\varpi_k(x)$ is equal to 1 on $\left[2^{-1} K_m \lambda_-^k, \infty\right)$, lies in $[0,1]$ on $\left[0, 2^{-1} K_m \lambda_-^k\right]$ and is 0 on $(-\infty, 0]$. We finally define for all $k \in \mathbb{N} \setminus \{0\}$

$$H_k : x \in \mathbb{R} \quad \mapsto \quad \frac{1}{k} \varpi_k(|x|) \ln |x|, \tag{43}$$

$$\psi_k : (a,b) \in \mathbb{R}^2 \quad \mapsto \quad G(a) G(b) H_k(a^2 - b). \tag{44}$$

Now, to motivate the expression of the upcoming estimator we first make a few comments. Since $\sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j$ is a convergent series (an easy consequence of the definition (5) of χ_k and the convergence of the series in assumption (**A2**)), and since we saw that $\sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0 (C_1^2 - C_2)$ is not zero thanks in particular to (**A6**), the first crucial observation that we may make from (13) is that

$$S_k := \frac{1}{k} \ln |C_1^2 - u_k| \longrightarrow \ln(\lambda_0), \quad k \to \infty. \tag{45}$$

This limit brings us to two deductions. First, w.l.o.g. we are going to estimate the unknown parameter $\ln \lambda_0$ rather than λ_0. Second, since for fixed $k \in \mathbb{N}$, C_1 and u_k are respectively estimated by

$$\bar{Y}_n := \frac{1}{n} \sum_{i=1}^{n} Y_i, \quad \bar{Y}_{k+1,n} := \frac{1}{n} \sum_{i=1}^{n} Y_i Y_{i+k+1}, \tag{46}$$

then it is natural to let $k = k_n$ in (45) for an integer valued sequence $(k_n)_{n \in \mathbb{N}}$ such that $\lim_{n \to \infty} k_n = \infty$, and introduce an estimator where C_1 and u_{k_n} are respectively replaced by $\bar{Y}_n$ and $\bar{Y}_{k_n+1,n}$.

We now arrive at the definition of the estimator. As previously said, and in view of (45), a first idea that springs to mind is to consider the following estimator:

$$\hat{T}_n := \frac{1}{k_n} \ln \left| [\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n} \right| \tag{47}$$

for a sequence $(k_n)_{n \in \mathbb{N}}$ with $\lim_{n \to \infty} k_n = \infty$. Equation (47) has however two drawbacks. First, by a very crude approximation we have that $[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n} \approx C_1^2 - u_{k_n}$ which tends to 0 as $n \to \infty$: this observation foretells that $[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n}$ can be arbitrarily too close to 0 with possibly large probability as n grows large, so that $\ln \left| [\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n} \right|$ is difficult to control and may e.g. tend to ∞ faster or more slowly than k_n. Second, $[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n}$ may also be arbitrarily large. Both those observations imply some potential "bad" behaviour of the estimator defined in (47), namely that it can grow large as $n \to \infty$, which will induce undesirable statistical properties. This justifies to rather consider the following regularized estimator

$$\begin{aligned} \hat{S}_n &:= \psi_{k_n}(\bar{Y}_n, \bar{Y}_{k_n+1,n}) = G(\bar{Y}_n) G(\bar{Y}_{k_n+1,n}) H_{k_n}\left(\left| [\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n} \right|\right) \\ &= G(\bar{Y}_n) G(\bar{Y}_{k_n+1,n}) \varpi_{k_n}\left(\left| [\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n} \right|\right) \frac{1}{k_n} \ln \left| [\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n} \right|, \quad n \in \mathbb{N}, \end{aligned} \tag{48}$$

for a well chosen sequence $(k_n)_{n\in\mathbb{N}}$ specified later, where ψ_k is defined in (44). One sees that $\hat{S}_n$ resembles (47), but the behaviour of $[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n}$ at 0 and infinity is "smoothed out" thanks to the functions $\varpi_{k_n}(\cdot)$ and $G(\cdot)$, see (44). The role of the function ψ_{k_n} in the definition (48) is motivated as follows. The terms $G(\bar{Y}_n)$ and $G(\bar{Y}_{k_n+1,n})$ prevent $\hat{S}_n$ from being too large if $\bar{Y}_n$ or $\bar{Y}_{k_n+1,n}$ are large. The term $H_{k_n}\left(\left|[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n}\right|\right)$ is such that it behaves like $k_n^{-1}\ln\left|[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n}\right|$ when $\left|[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n}\right|$ is far enough from 0. When the latter is close to 0, the definition (42) for ϖ_{k_n} entails that $\varpi_{k_n}\left(\left|[\bar{Y}_n]^2 - \bar{Y}_{k_n+1,n}\right|\right)$ tends to ∞ with the same order of magnitude as k_n, again preventing $\hat{S}_n$ from being too large.

The main results of this section show the convergence in $\mathbb{L}^2$ of the estimator towards the unknown parameter and gives an expansion for $\hat{S}_n$ for a suitable choice for k_n.

Theorem 6.1 (Quadratic convergence of estimator)**.** We suppose that (**A1**), (**A2**), (**A3**), (**A5**), (**A6**) and (**A7**) hold. Let us set $k_n := \lfloor c\ln n\rfloor$ where $c < -1/(2\ln\lambda_-)$. The following convergence in quadratic mean holds:

$$\left|\left|\hat{S}_n - \ln\lambda_0\right|\right|_2 = \mathrm{O}\left(\frac{1}{\ln n}\right) \longrightarrow 0, \quad n\to\infty \tag{49}$$

so that, in particular, $e^{\hat{S}_n}$ converges in probability towards λ_0 as $n\to\infty$. Furthermore, the estimator defined by $\hat{N}_n := \bar{Y}_n\left(1 - e^{\hat{S}_n}\right)$ converges in probability towards m_0 as $n\to\infty$.

The proof of Theorem 6.1 is postponed to Appendix Appendix A.9.

Remark 6.1. The sequence $(k_n)_{n\in\mathbb{N}}$ grows like $\ln n$ in the previous theorem. It may be (wrongly) guessed that such a growth is inconsequential and that any sequence satisfying $k_n = \mathrm{o}(n)$ could guarantee the convergence of the estimator $\hat{S}_n$. In fact, it turns out that this growth rate is not only significant but actually careful chosen, as it is illustrated numerically in Section 7 that the estimator performs very badly if k_n grows for example like $\sqrt{n}$ instead.

Last, a finer behaviour for the estimator can be described when the covariance of the immigration process decreases fast enough. This is explained in the following result.

Theorem 6.2 (Expansion for $\hat{S}_n$)**.** We suppose that (**A1**), (**A2**), (**A3**), (**A5**), (**A6**) and (**A7**) hold. Let us also assume here that the covariance of the immigration process satisfies $\nu_h = \mathrm{O}(\zeta^h)$ for some $\zeta < \lambda_-$ (ensuring that (**A2**) holds). Let us set $k_n := \lfloor c\ln n\rfloor$ where $c \in (-1/2\ln\zeta, -1/2\ln\lambda_-)$. Then one has the two terms expansion

$$\hat{S}_n - \ln\lambda_0 = \frac{1}{k_n}\ln\left|\sum_{j=0}^{\infty}\lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right| + \frac{1}{\sqrt{n}k_n\lambda_0^{k_n}}Z_n \tag{50}$$

where Z_n satisfies $Z_n \xrightarrow{\mathcal{D}} \mathcal{N}(0,\sigma)$, $n\to\infty$, with

$$\sigma := V\mathfrak{S}V', \tag{51}$$

$$V := \left(\frac{2C_1}{\left|\sum_{j=0}^{\infty}\lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right|}, -\frac{1}{\left|\sum_{j=0}^{\infty}\lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right|}\right)', \tag{52}$$

and where $\mathfrak{S}$ is given by

$$\mathfrak{S} = \begin{bmatrix} 2\sum_{h=1}^{\infty}(C_1^2 - u_{h-1}) & -2\frac{m_0}{(1-\lambda_0)^2}\sum_{h=0}^{\infty}\chi_h \\ & +\frac{C_1}{1-\lambda_0}(C_2 - C_1^2) \\ & +3C_1\sum_{h=1}^{\infty}(u_{h-1} - C_1^2) \\ -2\frac{m_0}{(1-\lambda_0)^2}\sum_{h=0}^{\infty}\chi_h & C_2^2 - C_1^4 + 2\sum_{h=1}^{\infty}(u_{h-1}^2 - C_1^4) \\ +\frac{C_1}{1-\lambda_0}(C_2 - C_1^2) & +2C_1^2\sum_{h=0}^{\infty}(u_{h-1} - C_1^2) \\ +3C_1\sum_{h=1}^{\infty}(u_{h-1} - C_1^2) & -2\frac{C_1 m_0}{(1-\lambda_0)^2}\sum_{t=1}^{\infty}\chi_{t-1} \end{bmatrix}. \tag{53}$$

The proof of Theorem 6.2 is postponed to Appendix Appendix A.10.

It seems that such expansions as (50) are quite rare in the literature and exist in the case of regression models, see convergence (3.6) in Taniguchi and Puri (1996, Theorem 2) for a corrected central limit theorem for an estimator in a linear model. Also note that the normalization factor in the central limit theorem involved in (50) is $\sqrt{n}k_n\lambda_0^{k_n}$ i.e. in the form $n^\delta \ln n$ with $\delta = 1/2 + c\ln\lambda_0$, which is different from the classical re-normalization in $\sqrt{n}$.

Also, a byproduct of (50) is the following refinement of (49) in Theorem 6.1:

$$\hat{S}_n - \ln\lambda_0 = \frac{1}{k_n}\ln\left|\sum_{j=0}^{\infty}\lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right| + \mathrm{O}_{\mathbb{L}^2}\left(\frac{1}{\sqrt{n}k_n\lambda_0^{k_n}}\right), \tag{54}$$

as indeed it is shown in the proof of Theorem 6.2 that $||Z_n||_2$ converges as $n \to \infty$ to some finite quantity. This makes more precise the speed of convergence of $\hat{S}_n$ towards $\ln\lambda_0$, as (54) implies that

$$\left|\left|\hat{S}_n - \ln\lambda_0\right|\right|_2 \propto \frac{1}{\ln(n)}, \tag{55}$$

i.e. that the speed of convergence is *exactly* of the order $1/\ln(n)$ here, and not just at most of order $1/\ln(n)$ as shown in (49). Convergence is thus very slow, which in particular explains why we need to pick very large n in the numerical illustrations in the forthcoming Section 7.2.

Theorems 6.1 and 6.2 will heavily rely on the two following technical results, proved respectively in Appendices Appendix A.7 and Appendix A.8.

Proposition 6.3. The following estimate holds for all k and n in $\mathbb{N}$:

$$||\psi_k(\bar{Y}_n, \bar{Y}_{k+1,n}) - \ln\lambda_0||_2 \leq K_S\left(\frac{1}{\lambda_-^k}\left|\left|\frac{S_n(k)}{n}\right|\right|_2 + \frac{1}{k}\right) \tag{56}$$

for some constant $K_S > 0$, where we recall that $S_n(k)$ is defined in (20).

Proposition 6.4. Let the sequence $(k_n)_{n\in\mathbb{N}}$ be such that $k_n = \mathrm{o}(n)$ as $n \to \infty$. Then the following convergence holds

$$\frac{1}{n}\mathbb{E}(S_n(k_n)S_n(k_n)') \longrightarrow \mathfrak{S} \tag{57}$$

as $n \to \infty$, where we recall that $S_n(\cdot)$ is defined in (20), and where $\mathfrak{S}$ is given by (53). If furthermore the sequence $(k_n)_{n\in\mathbb{N}}$ verifies

$$k_n = \mathrm{O}(n^\kappa),\ n \to \infty \text{ with } \kappa \in \left(0, 1 \wedge \frac{\beta - 1}{2(2\beta - 1)}\right), \tag{58}$$

then the following central limit theorem holds

$$\frac{S_n(k_n)}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathfrak{S}), \quad n \to \infty. \tag{59}$$

Finally, the proofs of Theorem 6.1 and Theorem 6.2 are resppectively presented in Appendices Appendix A.9 and Appendix A.10.

7. Numerical illustrations

7.1. Ultimately correlated case

We study numerically the behavior of the moment estimator for independent and correlated immigration model (1) on $N = 1000$ independent simulated trajectories to assess the performance of the moment estimator in finite sample. We generated from model (1) three trajectories of length $n = 300$, $n = 2000$ and $n = 5000$, in which the reproduction sequence $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$ is taken to be:

1) random variables $\mathcal{P}(\lambda_0)$ distributed,

2) Bernoulli random variables with expectation λ_0.

The immigration process $\{\epsilon_n,\ n \in \mathbb{Z}\}$ is taken to be:

1) $\mathcal{P}(1)$ distributed (for the independent case),

2) of the form $\epsilon_n = \prod_{i=0}^{k_0-1} Z_{n-i}$ where $\{Z_n,\ n \in \mathbb{Z}\}$ is i.i.d. with $\mathcal{P}(a_0)$ distribution (for the correlated case), for $k_0 = 2, 3$. In view of (17), we have:

$$\mathrm{Cov}\,(\epsilon_n, \epsilon_{n-h}) = \begin{cases} a_0^{2h}(a_0 + a_0^2)^{k_0-h} - a_0^{k_0} & \text{if} \quad h \leq k_0 - 1. \\ 0 & \text{if} \quad h \geq k_0 \end{cases}$$

We take $a_0 = 1$ in our numerical illustrations. Table 1 summarizes the distribution of the moment estimator $\hat{R}_{k_0,n}$ of λ_0 over these simulations experiments. As expected, Table 1 shows that the bias and the RMSE decrease when the size of the sample increases. From Table 1, the estimation of λ_0 becomes more difficult when k_0 increases. We will now numerically evaluate the potential effects of underestimation or overestimation of the instant k_0 (see Assumption $(\mathbf{A3})_1$). Table 2 summarizes the distribution of the moment estimators $\hat{R}_{1,n}$ and $\hat{R}_{3,n}$ of λ_0 for the model (1) when $\epsilon_n = Z_n Z_{n-1}$ i.e. $k_0 = 2$ with $Z_n \sim IID\mathcal{P}(1)$. As expected $\hat{R}_{1,n}$ is inconsistent and the bias of the estimation increases with increasing n. In contrast, when n increases, the distributions of $\hat{R}_{3,n}$ are close to those of Table 1 (see $k_0 = 2$) except for $\lambda_0 = 0.2$. From this example we draw the conclusion that, in the case of an overestimation of the parameter k_0 the proposed estimator $\hat{R}_{k_0,n}$ of λ_0 remain asymptotically consistent. On the other hand, if the instant k_0 is underestimated, the proposed estimator $\hat{R}_{k_0,n}$ is inconsistent even as n increases.

Figures 2, 3, 4, 5, 6 and 7 summarize via box-plots and compare the distribution of the proposed moment estimators and those proposed by Klimko and Nelson (1978) (see also Al-Osh and Alzaid (1987)) in the independent and correlated cases. For these simulations the distributions of the proposed moment estimators and those proposed by Klimko and Nelson (1978) are similar in the independent case (see the left panels of Figure 2), whereas the

distributions of our proposed estimators $\hat{R}_{k_0,n}$ and $\hat{M}_{k_0,n}$ are more accurate in the correlated case than those proposed by Klimko and Nelson (1978) (see the right panels of Figures 2 and 3). Under non-independent immigration, it appears that the estimators proposed by Klimko and Nelson (1978) (see also Al-Osh and Alzaid (1987)) are unreliable (see the right-bottom panel of Figures 2 and 3). See also the right-bottom panel of Figures 6 and 7 for which the estimation of the centered Gaussian density with the same variance plotted in dotted line is very poorly approximated because the fact that the estimators proposed by Klimko and Nelson (1978) is very biased ($\text{bias}(\lambda_0) = 0.165$ and $\text{bias}(m_0) = -0.331$) in the correlated case. For these simulations we also observe that: the precision around the estimated coefficients is better when the size of the sample increases and the distributions of $\hat{R}_{k_0,n}$ and $\hat{M}_{k_0,n}$ are more accurate in the independent case than in the correlated one. This is in accordance with the results of Romano and Thombs (1996) who showed that, with similar $(k_0 - 1)$-dependent noises, the asymptotic variance of the sample autocorrelations can be greater than 1 as well (1 is the asymptotic variance for independent white noises).

Figure 8 compares the standard estimator $\hat{\Omega}_S$ of Ω_S in Remark 4.2 by Klimko and Nelson (1978) with the proposed estimator based on spectral density estimation $\hat{\Omega}_{k_0}^{\text{SP}}$ of the asymptotic variance Ω_{k_0} in (36). We used the spectral estimator defined as (35) in Theorem 5.2. The AR order r in this theorem is taken equal to $\lfloor n^{(1/3)-\texttt{.Machine\$double.eps}} \rfloor$ where `.Machine$double.eps` is the smallest positive floating-point number x such that $n + x \neq n$. The order r can also be automatically selected by AIC or BIC, using the function `VARselect()` of the *vars* R package). In the case of independent immigration we know that the two estimators are consistent. In view of the left-top and the left-bottom panels of Figure 8, it seems that the standard estimator is most accurate than the proposed estimators in the independent case. This is not surprising because the spectral estimator is more robust, in the sense that this estimator continues to be consistent in the correlated case, contrary to the standard estimator. It is clear that in the correlated case $n\text{Var}(\hat{R}_{k_0,n} - \lambda_0)^2$ and $n\text{Var}(\hat{M}_{k_0,n} - m_0)^2$ are better estimated by $\hat{\Omega}_{k_0}^{\text{SP}}(1,1)$ and $\hat{\Omega}_{k_0}^{\text{SP}}(2,2)$ (see the box-plots 1 and 2 of the right-bottom panel of Figure 8) than by $\hat{\Omega}_S(1,1)$ and $\hat{\Omega}_S(2,2)$ (see the box-plots 1 and 2 of the right-top panel). The failure of the standard estimator of Ω_{k_0} in the correlated immigration model setting may have important consequences in terms of statistical inference.

Table 1: Sampling distribution of the moment estimator $\hat{R}_{k_0,n}$ of λ_0 for the model (1) with $\epsilon_n = \prod_{i=0}^{k_0-1} Z_{n-i}$ where $Z_n \sim IID\mathcal{P}(1)$.

			$\xi \sim \mathcal{P}(\lambda_0)$				$\xi \sim \mathcal{B}(\lambda_0)$			
			λ_0				λ_0			
			0.2	0.5	0.7	0.9	0.2	0.5	0.7	0.9
		Biais	-0.006	-0.033	-0.029	-0.029	-0.007	-0.027	-0.027	-0.034
		RMSE	0.108	0.103	0.075	0.049	0.110	0.090	0.066	0.051
		$\hat{\Omega}^{\mathrm{SP}}_{k_0}(1,1)$	4.582	1.938	1.006	0.293	4.527	2.261	0.833	0.301
		Min	0.001	0.116	0.406	0.694	0.002	0.220	0.463	0.710
$k_0 = 2$	$n = 300$	Q_1	0.108	0.402	0.627	0.846	0.109	0.414	0.637	0.844
		Q_2	0.183	0.473	0.678	0.876	0.185	0.475	0.677	0.873
		Mean	0.194	0.467	0.671	0.871	0.193	0.473	0.673	0.866
		Q_3	0.268	0.536	0.720	0.899	0.266	0.533	0.716	0.894
		Max	0.513	0.753	0.834	0.956	0.591	0.716	0.821	0.954
		Biais	-0.007	-0.007	-0.005	-0.004	-0.009	-0.007	-0.004	-0.004
		RMSE	0.057	0.037	0.026	0.014	0.057	0.035	0.023	0.012
		$\hat{\Omega}^{\mathrm{SP}}_{k_0}(1,1)$	5.935	2.417	1.202	0.337	5.690	2.081	0.938	0.240
		Min	0.002	0.374	0.599	0.857	0.031	0.372	0.598	0.856
$k_0 = 2$	$n = 2,000$	Q_1	0.158	0.470	0.678	0.887	0.156	0.470	0.683	0.889
		Q_2	0.193	0.493	0.696	0.897	0.190	0.494	0.697	0.896
		Mean	0.193	0.493	0.695	0.896	0.191	0.493	0.696	0.896
		Q_3	0.229	0.518	0.713	0.905	0.236	0.515	0.712	0.904
		Max	0.379	0.634	0.782	0.934	0.372	0.595	0.765	0.929
		Biais	0.051	-0.057	-0.040	-0.031	0.060	-0.056	-0.038	-0.030
		RMSE	0.164	0.153	0.096	0.052	0.171	0.142	0.083	0.049
		$\hat{\Omega}^{\mathrm{SP}}_{k_0}(1,1)$	14.888	4.184	1.409	1.098	14.043	3.533	1.428	0.371
		Min	0.000	0.006	0.291	0.691	0.000	0.003	0.394	0.644
$k_0 = 3$	$n = 300$	Q_1	0.133	0.356	0.607	0.843	0.129	0.363	0.617	0.849
		Q_2	0.233	0.454	0.669	0.874	0.249	0.445	0.666	0.875
		Mean	0.251	0.443	0.660	0.869	0.260	0.444	0.662	0.870
		Q_3	0.354	0.547	0.721	0.898	0.364	0.539	0.714	0.896
		Max	0.780	0.790	0.860	0.959	0.864	0.825	0.838	0.957
		Biais	-0.008	-0.009	-0.007	-0.005	0.001	-0.009	-0.006	-0.005
		RMSE	0.101	0.056	0.033	0.016	0.104	0.053	0.029	0.013
		$\hat{\Omega}^{\mathrm{SP}}_{k_0}(1,1)$	24.418	5.098	1.742	0.390	24.525	4.525	1.353	0.250
		Min	0.000	0.268	0.580	0.830	0.000	0.328	0.616	0.852
$k_0 = 3$	$n = 2,000$	Q_1	0.113	0.455	0.670	0.886	0.122	0.458	0.675	0.888
		Q_2	0.190	0.494	0.694	0.896	0.195	0.491	0.694	0.896
		Mean	0.192	0.491	0.693	0.895	0.201	0.491	0.694	0.895
		Q_3	0.259	0.526	0.715	0.905	0.273	0.524	0.714	0.904
		Max	0.525	0.688	0.782	0.947	0.581	0.670	0.797	0.933

Table 2: Sampling distribution of the moment estimator $\hat{R}_{k_0,n}$ of λ_0 for the model (1) with $\epsilon_n = Z_n Z_{n-1}$ i.e. $k_0 = 2$ where $Z_n \sim IID\mathcal{P}(1)$.

			$\xi \sim \mathcal{P}(\lambda_0)$				$\xi \sim \mathcal{B}(\lambda_0)$			
			λ_0				λ_0			
			0.2	0.5	0.7	0.9	0.2	0.5	0.7	0.9
		Biais	0.162	0.060	0.014	-0.018	0.165	0.062	0.021	-0.012
		RMSE	0.208	0.106	0.064	0.039	0.214	0.102	0.059	0.035
		Min	0.020	0.270	0.379	0.696	0.004	0.299	0.534	0.721
$k_0 = 1$	$n = 300$	Q_1	0.276	0.501	0.672	0.863	0.270	0.513	0.683	0.871
		Q_2	0.360	0.562	0.717	0.886	0.367	0.564	0.725	0.892
		Mean	0.362	0.560	0.714	0.882	0.365	0.562	0.721	0.888
		Q_3	0.451	0.617	0.757	0.906	0.460	0.618	0.762	0.910
		Max	0.732	0.858	0.865	0.957	0.847	0.797	0.863	0.961
		Biais	0.192	0.085	0.037	0.004	0.198	0.086	0.042	0.009
		RMSE	0.201	0.092	0.045	0.014	0.206	0.094	0.048	0.014
		Min	0.188	0.455	0.657	0.857	0.194	0.497	0.675	0.873
$k_0 = 1$	$n = 2,000$	Q_1	0.352	0.561	0.721	0.896	0.359	0.560	0.727	0.902
		Q_2	0.393	0.586	0.737	0.905	0.398	0.586	0.742	0.909
		Mean	0.392	0.585	0.737	0.904	0.398	0.586	0.742	0.909
		Q_3	0.431	0.608	0.754	0.914	0.436	0.609	0.758	0.916
		Max	0.645	0.726	0.824	0.947	0.587	0.715	0.829	0.939
		Biais	0.303	-0.030	-0.060	-0.035	0.305	-0.041	-0.059	-0.034
		RMSE	0.409	0.222	0.148	0.060	0.403	0.214	0.132	0.058
		Min	0.001	0.000	0.047	0.645	0.003	0.000	0.015	0.587
$k_0 = 3$	$n = 300$	Q_1	0.274	0.324	0.563	0.835	0.294	0.300	0.572	0.844
		Q_2	0.509	0.474	0.658	0.871	0.498	0.471	0.656	0.873
		Mean	0.503	0.474	0.640	0.865	0.505	0.459	0.641	0.866
		Q_3	0.737	0.633	0.739	0.900	0.705	0.609	0.722	0.896
		Max	0.996	0.989	0.938	0.964	0.998	0.975	0.961	0.961
		Biais	0.201	-0.019	-0.009	-0.006	0.202	-0.016	-0.006	-0.005
		RMSE	0.317	0.112	0.045	0.018	0.317	0.103	0.040	0.014
		Min	0.001	0.056	0.532	0.830	0.000	0.046	0.573	0.848
$k_0 = 3$	$n = 2,000$	Q_1	0.209	0.415	0.663	0.884	0.202	0.421	0.668	0.887
		Q_2	0.371	0.489	0.693	0.895	0.390	0.492	0.694	0.896
		Mean	0.401	0.481	0.691	0.894	0.402	0.484	0.694	0.895
		Q_3	0.573	0.549	0.721	0.906	0.577	0.552	0.718	0.905
		Max	0.998	0.800	0.820	0.955	0.994	0.770	0.821	0.934
		Biais	0.128	-0.010	-0.004	-0.002	0.134	-0.008	-0.003	-0.001
		RMSE	0.238	0.069	0.028	0.011	0.242	0.065	0.026	0.008
		Min	0.000	0.275	0.603	0.864	0.002	0.280	0.609	0.870
$k_0 = 3$	$n = 5,000$	Q_1	0.169	0.444	0.677	0.891	0.173	0.451	0.681	0.893
		Q_2	0.310	0.494	0.697	0.899	0.314	0.496	0.698	0.899
		Mean	0.328	0.490	0.696	0.898	0.334	0.492	0.697	0.899
		Q_3	0.465	0.535	0.714	0.905	0.468	0.537	0.714	0.904
		Max	0.952	0.688	0.774	0.938	0.999	0.689	0.785	0.929

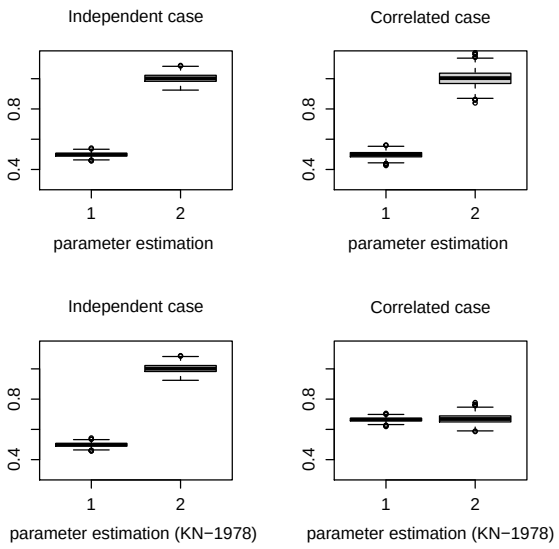

Figure 2: The moment estimators and the estimators proposed by Klimko and Nelson (1978) of $N = 1,000$ independent simulations of model (1) of size $n = 5,000$ with unknown parameter $\lambda_0 = 0.5$ and $m_0 = 1$, when the immigration is independent (left panels, with $\epsilon_n \sim IID\mathcal{P}(1)$) and when the immigration is correlated (right panels, with $\epsilon_n = Z_n Z_{n-1}$ where $Z_n \sim IID\mathcal{P}(1)$). The reproduction sequence $\xi \sim \mathcal{B}(\lambda_0)$. The panels display the distribution of the estimators $\hat{R}_{k_0,n}$ and $\hat{M}_{k_0,n}$. The two top panels (resp. two bottom panels) correspond to our proposed estimators (resp. to the estimators proposed by Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)).

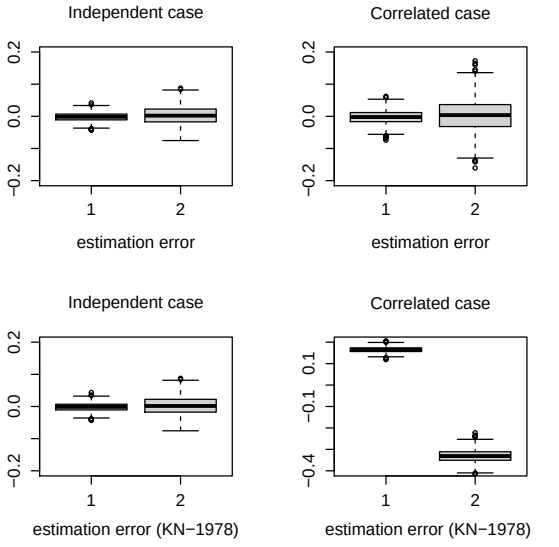

Figure 3: The panels at the top present the distribution of the estimation error of the estimates λ_0 and m_0 given in Figure 2. The left (resp. right) panel corresponds to the case of independent (correlated) immigration. Points 1 and 2, in the box-plots, display the distribution of the estimation error $\hat{R}_{k_0,n} - \lambda_0$ and $\hat{M}_{k_0,n} - m_0$. The two top panels (resp. two bottom panels) correspond to our proposed estimators (resp. to the estimators proposed by Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)).

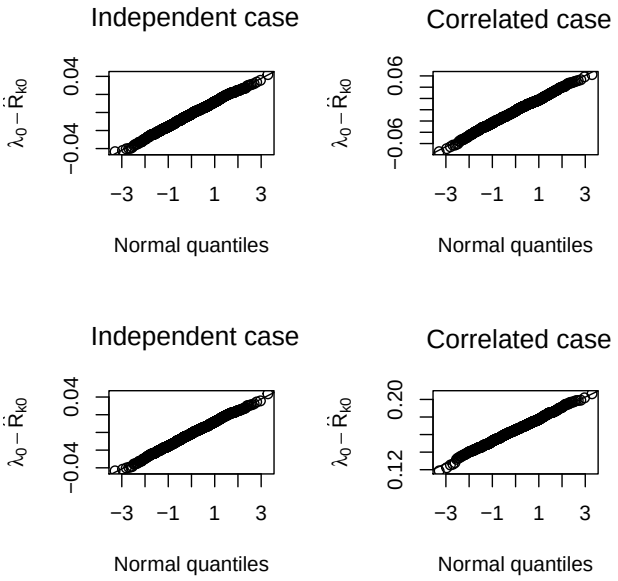

Figure 4: The panels at the top present the Q–Q plot of the estimates λ_0 given in Figure 2. The left (resp. right) panel corresponds to the case of independent (correlated) immigration. The two top panels (resp. two bottom panels) correspond to our proposed estimators (resp. to the estimators proposed by Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)).

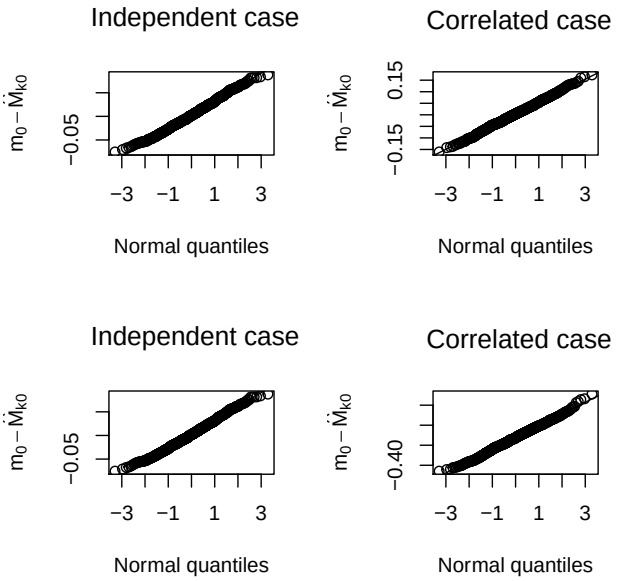

Figure 5: The panels at the top present the Q–Q plot of the estimates m_0 given in Figure 2. The left (resp. right) panel corresponds to the case of independent (correlated) immigration. The two top panels (resp. two bottom panels) correspond to our proposed estimators (resp. to the estimators proposed by Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)).

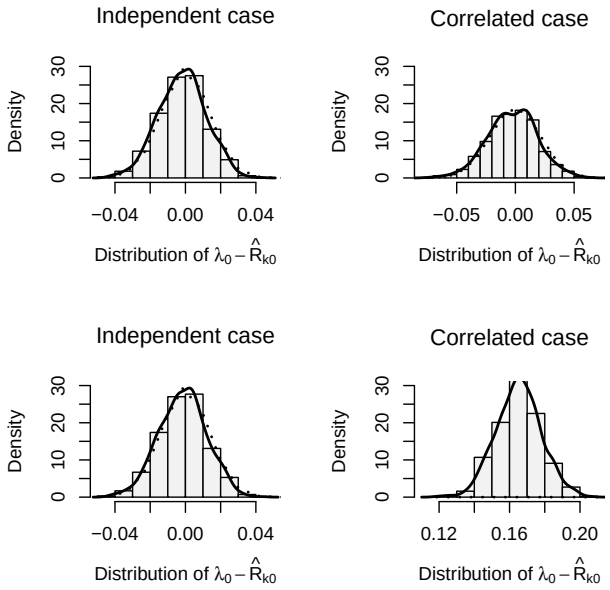

Figure 6: The panels at the top display the distribution of the estimates λ_0 given in Figure 2. The kernel density estimate is displayed in full line, and the centered Gaussian density with the same variance is plotted in dotted line. The left (resp. right) panel corresponds to the case of independent (correlated) immigration. The two top panels (resp. two bottom panels) correspond to our proposed estimators (resp. to the estimators proposed by Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)).

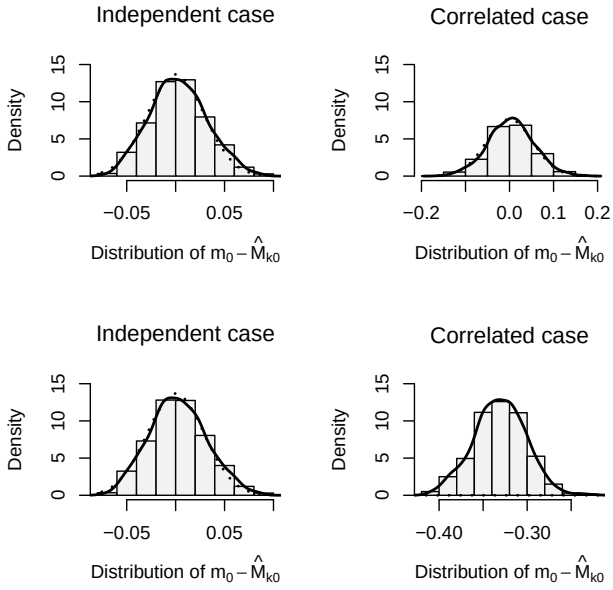

Figure 7: The panels at the top display the distribution of the estimates m_0 given in Figure 2. The kernel density estimate is displayed in full line, and the centered Gaussian density with the same variance is plotted in dotted line. The left (resp. right) panel corresponds to the case of independent (correlated) immigration. The two top panels (resp. two bottom panels) correspond to our proposed estimators (resp. to the estimators proposed by Klimko and Nelson (1978) and Al-Osh and Alzaid (1987)).

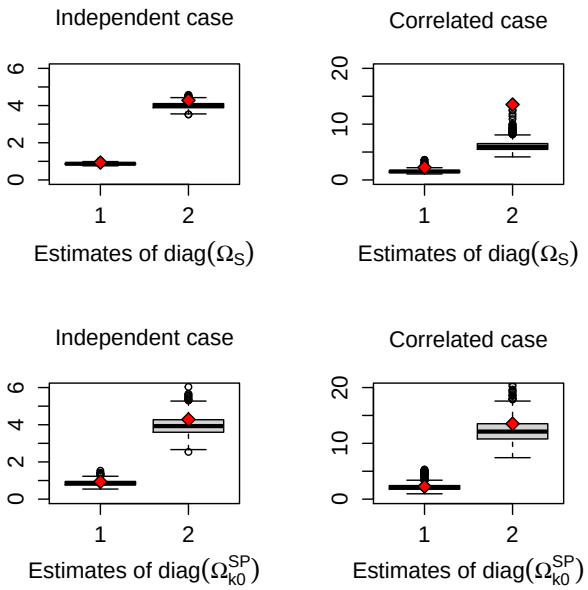

Figure 8: Comparison of standard and modified estimates of the asymptotic variance of the moment estimator of the model parameters, on the simulated models presented in Figure 2. The diamond symbols represent the mean, over $N = 1,000$ replications, of the standardized squared errors $n\left(\hat{R}_{k_0,n} - 0.5\right)^2$ for 1 (0.934 in the independent immigration case and 2.260 in the correlated immigration case) and $n\left(\hat{M}_{k_0,n} - 1\right)^2$ for 2 (4.283 in the independent immigration case and 13.519 in the correlated immigration case) .

7.2. Correlated case

We now illustrate the performance of the estimator $\hat{S}_n$ in (48) of which theorical convergence result was presented in Theorems 6.1 and 6.2.

The process considered here is the stationary version of (1), where $\xi \sim \mathcal{P}(\lambda_0)$, and where $\{\epsilon_n,\ n \in \mathbb{Z}\}$ is the stationary Markov chain with state space $\{0, 1\}$ and transition matrix $P = \begin{pmatrix} 1/2 & 1/2 \\ 1 & 0 \end{pmatrix}$. It is standard that α_ϵ is exponentially decreasing and that (**A1**) is satisfied if λ_0 is not close to 0, and all simulations below are performed such that $\Xi(\lambda_0, m_0, V_0, f_\nu) \neq 0$ (with the function Ξ defined in (38)) with $V_0 = \mathrm{Var}(\xi) = \lambda_0$ and (after some easy computation) $m_0 = \mathbb{E}(\epsilon) = 1/3$. Inspired by Theorems 6.1 and 6.2, we proceed to estimate $\mathbb{E}(\hat{S}_n)$ and

c	λ_0	$\ln(\lambda_0)$	$\mathbb{E}(\hat{S}_n)$	$\mathrm{Var}(\hat{S}_n)$	$\|\|\hat{S}_n - \ln(\lambda_0)\|\|_2$	$\mathbb{E}(\exp(\hat{S}_n))$
4.7	0.9	−0.1053605	−0.0665497	0.0001058	0.0401508	0.9356657
1.4	0.7	−0.3566749	−0.3192667	0.0066237	0.0895714	0.7292888
0.7	0.5	−0.6931472	−0.6326057	0.0369518	0.2015367	0.5423869
0.3	0.2	−1.6094379	−1.490265	0.0000447	0.1193602	0.225318

Table 3: Illustration of estimator $\hat{S}_n$ with $k_n = \lfloor c \ln n \rfloor$ (with $n = 5.10^6$)

$\mathbb{E}(\exp(\hat{S}_n))$ by Monte Carlo simulation by simulating 20 times $\hat{S}_n$ with $n = 5.10^6$ and $k_n = \lfloor c \ln n \rfloor$. The factor c is chosen purposefully close to the value $-1/(2 \ln \lambda_-)$ while assuming that λ_- is close to λ_0 in Theorem 6.1, which theoretically guarantees the convergence for $\hat{S}_n$.

The results are displayed in Table 3. We first remember that convergence is extremely slow, as was argued in (55), hence the choice of the very lengthy trajectory $n = 5.10^6$ (as opposed to the shorter lengths considered in Section 7.1). Table 3 shows that the estimator $\hat{S}_n$ performs well: in particular, $\mathbb{E}(\exp(\hat{S}_n))$ is close to λ_0 for different values of the reproduction parameter.

We next discuss the choice of the lag sequence $(k_n)_{n\in\mathbb{N}}$. To show that that the choice $k_n \propto \ln n$ (which, again, is motivated by Theorem 6.1) is efficient, we provide numerics for the case where the sequence tends to ∞ at a different speed, namely faster than $\ln n$, while still being an $o(n)$. Thus, and for comparison, Table 4 shows estimations when $k_n = \lfloor\sqrt{n}\rfloor$ (which still satisfies $k_n \ll n$): it is clear that the corresponding estimator performs very badly, with values of $\mathbb{E}(\hat{S}_n)$ being very close to 0, and $\mathbb{E}(\exp(\hat{S}_n))$ around 0.99, independently of the value of λ_0.

λ_0	$\ln(\lambda_0)$	$\mathbb{E}(\hat{S}_n)$	$\mathrm{Var}(\hat{S}_n)$	$\mathbb{E}(\exp(\hat{S}_n))$
0.9	-0.1053605	-0.0001556	$8.019.10^{-10}$	0.9998444
0.7	-0.3566749	-0.0008472	$7.539.10^{-09}$	0.9991532
0.5	-0.6931472	-0.001969	$2.651.10^{-08}$	0.998033
0.2	-1.6094379	-0.0055293	0.0000007	0.9944862

Table 4: Illustration of estimator $\hat{S}_n$ with $k_n = \lfloor\sqrt{n}\rfloor$ (with $n = 5.10^6$)

Appendix A. Appendix : Proofs of the main results

Appendix A.1. Proof of Proposition 3.1

In order to prove that Y_0 admits moments of order 2β, it suffices from (8) to prove that $\sum_{i=0}^{\infty} ||\theta^{(i)} \circ \epsilon_{-i}||_{2\beta}$ converges. This comes from the fact that

$$||\theta^{(i)} \circ \epsilon_{-i}||_{2\beta} = \mathrm{O}(\upsilon^i), \quad ||\theta_n^{(i)} \circ \epsilon_{n-i}||_{2\beta} = \mathrm{O}(\upsilon^i), \quad \forall n \in \mathbb{Z} \tag{A.1}$$

for some $\upsilon < 1$, see the proof in Kevei and Wiandt (2021, Section 3.1). Note that the authors proved this latter exponential decrease in a context where the immigration process is i.i.d., but a close look reveals that the only ingredient in the proof is the independence of $\theta^{(i)}$ from ϵ_{-i} for all $i \in \mathbb{N}$, which is the case here since $\{\xi_{n,k},\ n \in \mathbb{Z},\ k \in \mathbb{N}\}$ and $\{\epsilon_n,\ n \in \mathbb{Z}\}$ are independent sequences. □

Appendix A.2. Proof of Lemma 3.1

We introduce the quantity $\eta_k := \mathbb{E}(\epsilon_1\epsilon_{-k})$, $k \in \mathbb{N}$, and observe, using (1) as well as the independence of ϵ_1 and Y_{-k-1} from $\{\xi_{-k,j},\ j \in \mathbb{N}\}$, that

$$\begin{aligned} \eta_k &= \mathbb{E}\left(\epsilon_1\left[Y_{-k} - \sum_{j=1}^{Y_{-k-1}} \xi_{-k,j}\right]\right) = \mathbb{E}(\epsilon_1 Y_{-k}) - \lambda_0\mathbb{E}(\epsilon_1 Y_{-k-1}) \\ &= v_k - \lambda_0 v_{k+1}, \quad k \in \mathbb{N}. \end{aligned} \tag{A.2}$$

Observing now that $\eta_k = m_0^2 + \nu_{k+1}$, $k \in \mathbb{N}$, yields the following recursive first order relation from (A.2)

$$v_{k+1} = \frac{1}{\lambda_0} v_k - \frac{1}{\lambda_0}(m_0^2 + \nu_{k+1}), \quad k \in \mathbb{N},$$

leading to the expression

$$\begin{aligned} v_k &= -\sum_{j=1}^{k} \frac{1}{\lambda_0^j}(m_0^2 + \nu_{k+1-j}) + \frac{1}{\lambda_0^k} v_0 \\ &= -m_0^2 \frac{1}{\lambda_0} \frac{1 - 1/\lambda_0^k}{1 - 1/\lambda_0} - \sum_{j=1}^{k} \frac{\nu_{k+1-j}}{\lambda_0^j} + \frac{1}{\lambda_0^k} v_0 = \frac{m_0^2}{1-\lambda_0} + \frac{1}{\lambda_0^k} \left[\frac{m_0^2}{\lambda_0 - 1} - \sum_{j=0}^{k-1} \lambda_0^j \nu_{j+1} + v_0 \right] \\ &= \frac{m_0^2}{1-\lambda_0} + \frac{1}{\lambda_0^k} \left[\frac{m_0^2}{\lambda_0 - 1} - \sum_{j=0}^{\infty} \lambda_0^j \nu_{j+1} + v_0 \right] - \chi_k. \end{aligned} \tag{A.3}$$

We note now that $0 \leq v_k \leq [\mathbb{E}(\epsilon_1^2)]^{1/2} [\mathbb{E}(Y_{-k}^2)]^{1/2} = [\mathbb{E}(\epsilon_1^2)]^{1/2} C_2^{1/2}$ thanks to the Cauchy Schwarz inequality, so that the sequence $\{v_k,\ k \in \mathbb{N}\}$ is uniformly bounded. Also, one has that $|\chi_k| \leq \sum_{j=k}^{\infty} |\nu_{j+1}|$, which tends to 0 as $k \to \infty$ thanks to the (stronger) assumption (**A3**). The subcriticality condition (2) for λ_0 thus entails from (A.3) the explicit expression

$$v_0 = \frac{m_0^2}{1-\lambda_0} + \sum_{j=0}^{\infty} \lambda_0^j \nu_{j+1} = \frac{m_0^2}{1-\lambda_0} - \chi_0, \tag{A.4}$$

as well as the more general expression (11) for v_k, $k \in \mathbb{N}$.

We now come back to (12). Using (1), the independence of the sequence $\{\xi_{1,j},\ j \in \mathbb{N}\}$ as well as its independence from Y_0, yields

$$\begin{aligned} C_2 = \mathbb{E}(Y_1^2) &= \mathbb{E}\left(\left[\sum_{j=1}^{Y_0} \xi_{1,j} + \epsilon_1 \right]^2 \right) \\ &= \mathbb{E}\left(\sum_{j=1}^{Y_0} \xi_{1,j}^2 + \sum_{1 \leq r \neq r' \leq Y_0} \xi_{1,r} \xi_{1,r'} + 2\epsilon_1 \sum_{j=1}^{Y_0} \xi_{1,j} + \epsilon_1^2 \right) \\ &= \mathbb{E}(\xi^2)\mathbb{E}(Y_0) + [\mathbb{E}(\xi)]^2 \mathbb{E}(Y_0(Y_0 - 1)) + 2\mathbb{E}(\xi)\mathbb{E}(\epsilon_1 Y_0) + \mathbb{E}\epsilon_1^2 \\ &= V_{0,\xi} C_1 + \lambda_0^2 C_2 + 2\lambda_0 v_0 + \mathbb{E}\epsilon_1^2, \end{aligned}$$

which, thanks to (A.4) and the fact that $\chi_0 = -f_\nu(\lambda_0)$, yields (12). □

Appendix A.3. Proof of Corollary 3.2

(7) coupled with the relation (11) obtained in Lemma 3.1 yields $u_k - \lambda_0 u_{k-1} = C_1^2(1 - \lambda_0) - \chi_k$, from which in turn we deduce the following recursive relation

$$C_1^2 - u_k = \lambda_0 (C_1^2 - u_{k-1}) - \chi_k, \quad k \in \mathbb{N}.$$

Since $u_{-1} = C_2$, we arrive then at $C_1^2 - u_k = \sum_{j=0}^{k} \lambda_0^j \chi_{k-j} + \lambda_0^{k+1}(C_1^2 - C_2)$, hence (13). □

Appendix A.4. Proof of Proposition 4.1

Iterating Equation (1) yields that Y_n is a function of ϵ_k, $k \leq n$ and $\xi_{k,i}$, $k \leq n$, $i \in \mathbb{N}$. $\{Y_n,\ n \in \mathbb{Z}\}$ and $\{Y_n Y_{n+k_0},\ n \in \mathbb{Z}\}$ being stationary and integrable, and $\{(\epsilon_n, \xi_{n,i}),\ n \in \mathbb{Z},\ i \in \mathbb{N} \setminus \{0,\}\}$ being ergodic, yields the a.s. convergence of $\hat{Y}_n$ and $\hat{Y}_{k_0,n}$ (defined in (20)) respectively towards C_1 and u_{k_0} by Francq and Zakoïan (2019, Theorem A.2 p.344). Hence $\hat{R}_{k_0,n}$ and, in turn, $\hat{M}_{k_0,n}$, converge a.s. to λ_0 and m_0 as $n \to \infty$ based on (19) and (7). □

Appendix A.5. Proof of Theorem 4.2

We start by stating two technical lemmas.

Lemma Appendix A.1. Let k, j and m in $\mathbb{N}$ with $m > j > k$ and φ_k^- (resp. φ_m^+) be a $\sigma(\epsilon_n,\ n \leq k) \otimes \sigma(\xi_{n,i}, n \leq k, i \in \mathbb{N})$ measurable (resp. $\sigma(\epsilon_n,\ n \geq m)$ measurable) random variables belonging to $\mathbb{L}^{(1-2/\beta)\vee 4}$. The following decompositions hold:

$$\begin{aligned}
\mathrm{Cov}(\varphi_k^-, Y_j Y_m) &= \sum_{s=0}^{m-j-1} \lambda_0^s \left[\sum_{t=0}^{j-k-1} \lambda_0^t \mathrm{Cov}(\varphi_k^-, \epsilon_{j-t}\epsilon_{m-s}) + \lambda_0^{j-k} \mathrm{Cov}(\varphi_k^-, Y_k \epsilon_{m-s}) \right] \\
&\quad + \lambda_0^{m-j} \mathrm{Cov}(\varphi_k^-, Y_j^2), \qquad \text{(A.5)} \\
\mathrm{Cov}(\varphi_k^-, Y_j^2 \varphi_m^+) &= \sum_{s=0}^{j-k-1} \lambda_0^{2s} f_{j-s}(k,m) + \lambda_0^{2(j-k)} \mathrm{Cov}(\varphi_k^-, Y_k^2 \varphi_m^+), \quad \text{where} \qquad \text{(A.6)} \\
f_j(k,m) &= V_{0,\xi} \mathrm{Cov}(\varphi_k^-, Y_{j-1}\varphi_m^+) + \mathrm{Cov}(\varphi_k^-, \epsilon_j^2 \varphi_m^+) + 2\lambda_0 \mathrm{Cov}(\varphi_k^-, Y_{j-1}\epsilon_j \varphi_m^+) \qquad \text{(A.7)}
\end{aligned}$$

with $V_{0,\xi} = \mathrm{Var}(\xi)$.

Proof. In view of (1), the independence of $\xi_{m,p}$ from φ_k^-, Y_j yields:

$$\mathrm{Cov}(\varphi_k^-, Y_j Y_m) = \mathrm{Cov}\left(\varphi_k^-, Y_j \left[\sum_{p=1}^{Y_{m-1}} \xi_{m,p} + \epsilon_m\right]\right) = \lambda_0 \mathrm{Cov}(\varphi_k^-, Y_j Y_{m-1}) + \mathrm{Cov}(\varphi_k^-, Y_j \epsilon_m)$$

which, by direct induction, gives

$$\mathrm{Cov}(\varphi_k^-, Y_j Y_m) = \sum_{s=0}^{m-j-1} \lambda_0^s \mathrm{Cov}(\varphi_k^-, Y_j \epsilon_{m-s}) + \lambda_0^{m-j} \mathrm{Cov}(\varphi_k^-, Y_j^2).$$

A similar argument (involving this time the independence of φ_k^- and ϵ_{m-s} from $\{\xi_{r,p},\ p \geq 1\}$, $r = k+1, ..., j$), yields that $\mathrm{Cov}(\varphi_k^-, Y_j \epsilon_{m-s}) = \sum_{t=0}^{j-k-1} \lambda_0^t \mathrm{Cov}(\varphi_k^-, \epsilon_{j-t}\epsilon_{m-s}) + \lambda_0^{j-k} \mathrm{Cov}(\varphi_k^-, Y_k \epsilon_{m-s})$ which, plugged in the above equality, yields (A.5).

In order to prove (A.6), we first expand $Y_j^2 = \left[\sum_{p=1}^{Y_{j-1}} \xi_{j,p} + \epsilon_j\right]^2 = \sum_{p=1}^{Y_{j-1}} \xi_{j,p}^2 + \sum_{1 \leq p \neq p' \leq Y_{j-1}} \xi_{j,p}\xi_{j,p'} + \epsilon_j^2 + 2\epsilon_j \sum_{p=1}^{Y_{j-1}} \xi_{j,p}$, which, still by independence arguments and since $\mathbb{E}(\xi^2) = V_{0,\xi} + \lambda_0^2$, yields

$$\begin{aligned}
\mathrm{Cov}(\varphi_k^-, Y_j^2 \varphi_m^+) &= (V_{0,\xi} + \lambda_0^2) \mathrm{Cov}(\varphi_k^-, Y_{j-1}\varphi_m^+) + \lambda_0^2 \mathrm{Cov}(\varphi_k^-, Y_{j-1}(Y_{j-1}-1)\varphi_m^+) \\
&\quad + \mathrm{Cov}(\varphi_k^-, \epsilon_j^2 \varphi_m^+) + 2\lambda_0 \mathrm{Cov}(\varphi_k^-, \epsilon_j Y_{j-1} \varphi_m^+) \\
&= \lambda_0^2 \mathrm{Cov}(\varphi_k^-, Y_{j-1}^2 \varphi_m^+) + f_j(k,m)
\end{aligned}$$

where $f_j(k,m)$ is defined by (A.7). Direct induction in the above equality yields (A.6). □

Lemma Appendix A.2. Let k, j in $\mathbb{N}$ with $j > k$ and φ_t^- (resp. φ_t^+) be a $\sigma(\epsilon_n,\ n \leq t,\ \xi_{n,i}, n \leq t, i \in \mathbb{N})$ measurable (resp. $\sigma(\epsilon_n,\ n \geq t)$ measurable) random variables belonging to $\mathbb{L}^{\beta \vee 4}$, uniformly in $t = k, ..., j$. The following upper bound holds

$$\left| \sum_{s=0}^{j-k-1} \lambda_0^s \mathrm{Cov}(\varphi_k^-, \varphi_{j-s}^+) \right| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{j-k-1}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{j-k-1}{2} \rfloor} \right], \tag{A.8}$$

where $\beta > 2$ is defined in (**A1**). As a consequence, the following similar bound holds for all $k < j < m$ in $\mathbb{N}$:

$$\left| \mathrm{Cov}(\varphi_k^-, Y_j^\ell \varphi_m^+) \right| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{j-k-1}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{j-k-1}{2} \rfloor} \right] \tag{A.9}$$

for $\ell = 1, 2$.

Proof. Since $\{\xi_{m,i}, m \in \mathbb{N}, i \in \mathbb{N}\}$ is independent from $\{\epsilon_n,\ n \in \mathbb{Z}\}$, Davydov's inequality (see Davydov (1968)) as well as the $\mathbb{L}^\beta$ assumption on φ_t^- and φ_t^+, $t = k, ..., j$ (see Assumption **(A2)**), yields that $|\lambda_0^s \mathrm{Cov}(\varphi_k^-, \varphi_{j-s}^+)|$ is less than $K\ \alpha_\epsilon(j-s-k)^\beta$ or $K\ \lambda_0^s$ for $s = 0, ..., j-k-1$ (thanks to the Cauchy Scwharz inequality thanks to the $\mathbb{L}^4$, hence $\mathbb{L}^2$, assumption), for some constant $K > 0$. We split the sum on the lefthanside of (A.8) in $s = 0, ..., \lfloor (j-k-1)/2 \rfloor$ and $s = \lfloor (j-k-1)/2 \rfloor, ..., j-k-1$. When $s = 0, ..., \lfloor (j-k-1)/2 \rfloor$ we get that $j - s \geq \lfloor (j-k-1)/2 \rfloor$, so that Davydov's inequality reads

$$\lambda_0^s \left| \mathrm{Cov}(\varphi_k^-, \varphi_{j-s}^+) \right| \leq C \lambda_0^s \alpha_\epsilon \left(\left\lfloor \frac{j-k-1}{2} \right\rfloor \right)^{1-2/\beta} \left\| \varphi_k^- \right\|_2 \cdot \left\| \varphi_{j-s}^+ \right\|_2 \leq K \lambda_0^s \alpha_\epsilon \left(\left\lfloor \frac{j-k-1}{2} \right\rfloor \right)^{1-2/\beta} \tag{A.10}$$

for some universal positive constants C and K. So that summing (A.10) yields

$$\left| \sum_{s=0}^{\lfloor (j-k-1)/2 \rfloor} \lambda_0^s \mathrm{Cov}(\varphi_k^-, \varphi_{j-s}^+) \right| \leq K \alpha_\epsilon \left(\left\lfloor \frac{j-k-1}{2} \right\rfloor \right)^{1-2/\beta} \tag{A.11}$$

for some (different) constant $K > 0$. When $s = \lfloor (j-k-1)/2 \rfloor + 1, ...j-k-1$, one easily gets

$$\left| \sum_{s=\lfloor (j-k-1)/2 \rfloor + 1}^{j-k-1} \lambda_0^s \mathrm{Cov}(\varphi_k^-, \varphi_{j-s}^+) \right| \leq K \lambda_0^{\lfloor \frac{j-k-1}{2} \rfloor}, \tag{A.12}$$

so that combining (A.11) and (A.12) yields (A.8).
We next prove (A.9) for $s = 2$ ($s = 1$ being proved similarly).

Observing each term on the righthandside of (A.7), we have (still by standard arguments) that

$$|f_j(k,m)| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{j-1}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{j-1}{2} \rfloor} \right].$$

Hence, splitting the sum on the righthandside of (A.6) in $s = 0, ..., \lfloor (j-k-1)/2 \rfloor$ and $s = \lfloor (j-k-1)/2 \rfloor + 1, ..., j-k-1$ and using the fact that $\alpha_\epsilon(\cdot)$ is decreasing, yields

$$\left| \sum_{s=0}^{j-k-1} \lambda_0^{2s} f_{j-s}(k,m) \right| \leq \sum_{s=0}^{\lfloor \frac{j-k-1}{2} \rfloor} |\lambda_0^{2s} f_{j-s}(k,m)| + \sum_{s=\lfloor \frac{j-k-1}{2} \rfloor + 1}^{j-k-1} |\lambda_0^{2s} f_{j-s}(k,m)|$$

$$\leq K\left[\alpha_\epsilon\left(\left\lfloor\frac{j-k-1}{2}\right\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{j-k-1}{2}\rfloor}\right]+K\lambda_0^{2\lfloor\frac{j-k-1}{2}\rfloor}$$
$$\leq K\left[\alpha_\epsilon\left(\left\lfloor\frac{j-k-1}{2}\right\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{j-k-1}{2}\rfloor}\right],$$

where the constant $K>0$ changes at each inequality. Combined with the fact that:

$$\left|\lambda_0^{2(j-k)}\mathrm{Cov}(\varphi_k^-,Y_k^2\varphi_m^+)\right|\leq K\lambda_0^{2(j-k)},$$

(a consequence of the Cauchy-Schwarz inequality, thanks to the $\mathbb{L}^{\beta\vee 4}$ integrability assumption of φ_k^-, φ_m^+ and the $\mathbb{L}^4$ property of Y_j), we thus obtain (A.9) from (A.6). □

Proof of Theorem 4.2. The proof follows the line of proof in Francq and Zakoïan (1998), with some noticeable technical differences. More precisely, we intend to apply Herrndorf (1984) to a truncation of Y_n expressed as the series (9).

Appendix A.5.0.1. ⋄ *Step 1:.* we prove that $\mathfrak{S}_{k_0}$ defined in (23) is a convergent series. To prove this, and in view of (24), we prove here only that $\sum_{j=1}^\infty \mathrm{Cov}(Y_0Y_{k_0},Y_jY_{j+k_0})$ converges, the series involving the other terms in (24) being proved similarly. Using (A.5) with $k:=k_0$, $m:=j+k_0$, $\varphi_k^-:=Y_0Y_{k_0}$, yields for $j>k_0$

$$\mathrm{Cov}(Y_0Y_{k_0},Y_jY_{j+k_0})=\sum_{s=0}^{k_0-1}\lambda_0^s\left[\sum_{t=0}^{j-k_0-1}\lambda_0^t\mathrm{Cov}(Y_0Y_{k_0},\epsilon_{j-t}\epsilon_{j+k_0-s})+\lambda_0^{j-k_0}\mathrm{Cov}(Y_0Y_{k_0},Y_{k_0}\epsilon_{j+k_0-s})\right]$$
$$+\lambda_0^{k_0}\mathrm{Cov}(Y_0Y_{k_0},Y_j^2),\quad\text{(A.13)}$$

of which different terms are appropriately upper bounded. Applying inequalities (A.8), the Cauchy Schwarz inequality (under the assumptions (**A2**) and thanks to Proposition 3.1) and (A.9) give respectively

$$\begin{aligned}\sum_{t=0}^{j-k_0-1}\lambda_0^t\left|\mathrm{Cov}(Y_0Y_{k_0},\epsilon_{j-t}\epsilon_{j+k_0-s})\right| &\leq K\left[\alpha_\epsilon\left(\left\lfloor\frac{j-k_0-1}{2}\right\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{j-k_0-1}{2}\rfloor}\right],\\ \left|\mathrm{Cov}(Y_0Y_{k_0},Y_{k_0}\epsilon_{j+k_0-s})\right| &\leq K,\\ \left|\mathrm{Cov}(Y_0Y_{k_0},Y_j^2)\right| &\leq K\left[\alpha_\epsilon\left(\left\lfloor\frac{j-k_0-1}{2}\right\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{j-k_0-1}{2}\rfloor}\right],\end{aligned}$$

for some constant $K>0$, which, plugged into (A.13), gives

$$\left|\mathrm{Cov}(Y_0Y_{k_0},Y_jY_{j+k_0})\right|\leq K\left[\alpha_\epsilon\left(\left\lfloor\frac{j-k_0-1}{2}\right\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{j-k_0-1}{2}\rfloor}+\lambda_0^{j-k_0}\right]\quad\text{(A.14)}$$

for $j>k_0$, yielding that $\sum_{j=1}^\infty\left|\mathrm{Cov}(Y_0Y_{k_0},Y_jY_{j+k_0})\right|$ is indeed finite thanks to Condition (2) and Assumption (**A1**).

Appendix A.5.0.2. $\diamond$ *Step 2:.* Recall that Y_n may be expanded as (9). We then further decompose (9) for all $m \in \mathbb{N}$ into

$$Y_n = Y_{n,m} + Z_{n,m}, \quad Y_{n,m} := \sum_{i=0}^{m} \theta_n^{(i)} \circ \epsilon_{n-i}, \quad Z_{n,m} := \sum_{i=m+1}^{\infty} \theta_n^{(i)} \circ \epsilon_{n-i}, \tag{A.15}$$

and $\mathcal{Y}_n(k_0) = [Y_n - C_1, Y_n Y_{n+k_0-1} - u_{k_0-1}, Y_n Y_{n+k_0} - u_{k_0}]'$ into

$$\begin{aligned}
\mathcal{Y}_n(k_0) &= \mathcal{Y}_{n,m}(k_0) + \mathcal{Z}_{n,m}(k_0), \\
\mathcal{Y}_{n,m}(k_0) &:= [Y_{n,m} - \mathbb{E}(Y_{n,m}), Y_{n,m}Y_{n+k_0-1,m} - \mathbb{E}(Y_{n,m}Y_{n+k_0-1,m}), \\
&\quad Y_{n,m}Y_{m+k_0,m} - \mathbb{E}(Y_{n,m}Y_{n+k_0,m})]', \\
\mathcal{Z}_{n,m}(k_0) &:= [Z_{n,m} - \mathbb{E}(Z_{n,m}), Y_{n,m}Z_{n+k_0-1,m} + Z_{n,m}Y_{n+k_0-1,m} + Z_{n,m}Z_{n+k_0-1,m} \\
&\quad -\mathbb{E}(Y_{n,m}Z_{n+k_0-1,m} + Z_{n,m}Y_{n+k_0-1,m} + Z_{n,m}Z_{n+k_0-1,m}), \\
&\quad Y_{n,m}Z_{n+k_0,m} + Z_{n,m}Y_{n+k_0,m} + Z_{n,m}Z_{n+k_0,m} \\
&\quad -\mathbb{E}(Y_{n,m}Z_{n+k_0,m} + Z_{n,m}Y_{n+k_0,m} + Z_{n,m}Z_{n+k_0,m}),]'
\end{aligned} \tag{A.16}$$

for all $m \in \mathbb{N}$. We also recall from the proof of Proposition 3.1 that $\|\theta_n^{(i)} \circ \epsilon_{n-i}\|_{2\beta} = \|\theta^{(i)} \circ \epsilon_{-i}\|_{2\beta}$ is an $\mathrm{O}(\upsilon^i)$ for some $\upsilon < 1$, so that $\|Y_n\|_{2\beta}$, $\|Y_{n,m}\|_{2\beta}$, $\|Z_{n,m}\|_{2\beta}$ are finite for all $n \in \mathbb{Z}$ and $m \in \mathbb{N}$. We prove in this Step that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \mathcal{Y}_{i,m}(k_0) \xrightarrow[n\to\infty]{\mathcal{D}} \mathcal{N}(0, \mathfrak{S}_{k_0}^m), \tag{A.17}$$

for some semi-definite positive matrix $\mathfrak{S}_{k_0}^m \in \mathbb{R}^{3\times 3}$, for all $m \in \mathbb{N}$ given by

$$\mathfrak{S}_{k_0}^m := \mathrm{Var}(\mathcal{Y}_{0,m}(k_0)) + \sum_{h=1}^{\infty} [\mathrm{Cov}(\mathcal{Y}_{0,m}(k_0), \mathcal{Y}_{h,m}(k_0)) + \mathrm{Cov}(\mathcal{Y}_{0,m}(k_0), \mathcal{Y}_{h,m}(k_0))'] \tag{A.18}$$

For this we verify Assumption (1.2) in Herrndorf (1984) which says that the following limit exists

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\left(\left(\sum_{i=1}^{n} \mathcal{Y}_{i,m}(k_0)\right)\left(\sum_{i=1}^{n} \mathcal{Y}_{i,m}(k_0)\right)'\right) &= \frac{1}{n}\mathbb{E}\left(\sum_{1\le i,j\le n} \mathcal{Y}_{i,m}(k_0)'\mathcal{Y}_{j,m}(k_0)\right) \\
&= \frac{1}{n}\sum_{h=-n}^{n} (n - |h|)\mathrm{Cov}(\mathcal{Y}_{0,m}(k_0), \mathcal{Y}_{h,m}(k_0)) \xrightarrow[n\to\infty]{} \mathfrak{S}_{k_0}^m,
\end{aligned} \tag{A.19}$$

as indeed, if we prove that $\mathfrak{S}_{k_0}^m$ in (A.18) converges absolutely then the limit (A.19) exists by the dominated convergence theorem. Then, under Assumption **(A1)** and $m \in \mathbb{N}$ fixed, we prove that the process $(\mathcal{Y}_{n,m}(k_0))_{n\in\mathbb{Z}}$ is strongly mixing (see Davidson (1994, Theorem 14.1)), i.e. that the following condition holds:

$$\sum_{h=1}^{\infty} \alpha_{m,\mathcal{Y}}(h)^{1-2/\beta} < \infty \tag{A.20}$$

where $\alpha_{m,\mathcal{Y}}(h) := \sup\{|\mathbb{P}(A \bigcap B) - \mathbb{P}(A)\mathbb{P}(B)|,\ A \in \sigma(\mathcal{Y}_{n,m}(k_0),\ 1 \le n \le r),\ B \in \sigma(\mathcal{Y}_{n,m}(k_0),\ n \ge r+h)\}$ is the mixing coefficient associated to the sequence $\{\mathcal{Y}_{i,m}(k_0),\ i \in \mathbb{N}\}$.

Now, in view of the definition (A.16) of $\mathcal{Y}_{n,m}(k_0)$, we first prove the following limit

$$\sum_{h=0}^{\infty}\left|\mathrm{Cov}(Y_{0,m}Y_{k_0-t_1,m}^{t_2},Y_{h,m}Y_{h+k_0-t_1,m}^{t_2})-\mathrm{Cov}(Y_0Y_{k_0-t_1}^{t_2},Y_hY_{h+k_0-t_1}^{t_2})\right|\xrightarrow[m\to\infty]{}0, \tag{A.21}$$

for all t_1, t_2 in $\{0,1\}$. Indeed, one deduces easily from this both that $\mathfrak{S}_{k_0}^m$ in (A.18) is finite for all m, and that it converges to $\mathfrak{S}_{k_0}$ defined in (23). We prove (A.21) only for $t_1=0$ and $t_2=1$, the other cases being similar. For these t_1 and t_2, we decompose the summands in the above series as

$$\mathrm{Cov}(Y_{0,m}Y_{k_0,m},Y_{h,m}Y_{h+k_0,m})-\mathrm{Cov}(Y_0Y_{k_0},Y_hY_{h+k_0})=\mathrm{Cov}([Y_{0,m}-Y_0]Y_{k_0,m},Y_{h,m}Y_{h+k_0,m})$$
$$+\mathrm{Cov}(Y_0[Y_{k_0,m}-Y_{k_0}],Y_{h,m}Y_{h+k_0,m})+\mathrm{Cov}(Y_0Y_{k_0},[Y_{h,m}-Y_h]Y_{h+k_0,m})+\mathrm{Cov}(Y_0Y_{k_0},Y_h[Y_{h+k_0,m}-Y_{h+k_0}])$$

and we focus on the first term in the righhandside, the other terms being tackled with similarly, so that we prove that

$$\sum_{h=0}^{\infty}|\mathrm{Cov}([Y_{0,m}-Y_0]Y_{k_0,m},Y_{h,m}Y_{h+k_0,m})|\xrightarrow[m\to\infty]{}0. \tag{A.22}$$

For this, we observe that, letting

$$\varphi_{k_0,m}^{-}:=Z_{0,m}Y_{k_0,m}, \tag{A.23}$$

and in view of (A.15) and for $h/2>k_0$:

$$\begin{aligned}-\mathrm{Cov}([Y_{0,m}-Y_0]Y_{k_0,m},Y_{h,m}Y_{h+k_0,m}) &= \mathrm{Cov}(Z_{0,m}Y_{k_0,m},Y_{h,m}Y_{h+k_0,m})=\mathrm{Cov}(\varphi_{k_0,m}^{-},Y_{h,m}Y_{h+k_0,m})\\ &= \sum_{i_1=0}^{m}\sum_{i_2=0}^{m}\mathrm{Cov}\left(\varphi_{k_0,m}^{-},[\theta_h^{(i_1)}\circ\epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)}\circ\epsilon_{h+k_0-i_2}]\right), \qquad \text{(A.24)}\\ &= \sum_{j=1}^{4}I_j(h,m),\end{aligned}$$

where

$$\begin{aligned}I_1(h,m) &= \sum_{\substack{0\le i_1\le\min(\lfloor h/2\rfloor-k_0,m)\\0\le i_2\le\min(\lfloor h/2\rfloor,m)}}\mathrm{Cov}\left(\varphi_{k_0,m}^{-},[\theta_h^{(i_1)}\circ\epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)}\circ\epsilon_{h+k_0-i_2}]\right),\\ I_2(h,m) &= \sum_{\substack{\lfloor h/2\rfloor-k_0<i_1\le m\\0\le i_2\le\lfloor h/2\rfloor}}\mathrm{Cov}\left(\varphi_{k_0,m}^{-},[\theta_h^{(i_1)}\circ\epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)}\circ\epsilon_{h+k_0-i_2}]\right),\\ I_3(h,m) &= \sum_{\substack{0\le i_1\le\lfloor h/2\rfloor-k_0\\\lfloor h/2\rfloor<i_2\le m}}\mathrm{Cov}\left(\varphi_{k_0,m}^{-},[\theta_h^{(i_1)}\circ\epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)}\circ\epsilon_{h+k_0-i_2}]\right),\\ I_4(h,m) &= \sum_{\substack{\lfloor h/2\rfloor-k_0<i_1\le m\\\lfloor h/2\rfloor<i_2\le m}}\mathrm{Cov}\left(\varphi_{k_0,m}^{-},[\theta_h^{(i_1)}\circ\epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)}\circ\epsilon_{h+k_0-i_2}]\right),\end{aligned}$$

where a sum over an empty set is equal to zero by convention. We notice that $\varphi_{k_0,m}^{-}$ depends on $(\theta_0^{(i)})_{i\in\mathbb{N}}$, $(\theta_{k_0}^{(i)})_{i\in\mathbb{N}}$ and ϵ_j, $j\le k_0$. Also, we recall from (10) that $\theta_0^{(i)}$ and $\theta_{k_0}^{(i)}$ are independent

from $\theta_h^{(i_1)}$ and $\theta_{h+k_0}^{(i_2)}$ for all $i \in \mathbb{N}$ as soon as $[\{-i, k_0-i\} \cap \{h-i_1, h+k_0-i_2\} = \emptyset,\ \forall i \in \mathbb{N}] \iff [k_0 < h - i_1,\ k_0 < h + k_0 - i_2] \iff [i_1 < h - k_0,\ i_2 < h]$. The mixing coefficient $\alpha_\theta(h) := \sup\{|\mathbb{P}(A \bigcap B) - \mathbb{P}(A)\mathbb{P}(B)|,\ A \in \sigma((\theta_n^{(i)})_{i\in\mathbb{N}}, 1 \leq n \leq r),\ B \in \sigma(\theta_n^{(i)}, n - i + 1 \geq r + h)\}$ verifies $\alpha_\theta(h) = 0$ as soon as $h \geq 1$ for similar independence reasons. And, we remember that $\{(\theta_n^{(i)})_{i\in\mathbb{N}},\ n \in \mathbb{Z}\}$ and $\{\epsilon_n,\ n \in \mathbb{Z}\}$ are independent. Thus, the following upper bound holds, using Davydov's inequality (see Davydov (1968)):

$$|I_1(h,m)| \leq \sum_{\substack{0\leq i_1\leq \lfloor h/2\rfloor - k_0 \\ 0\leq i_2 \leq \lfloor h/2\rfloor}} K \left\|\varphi_{k_0,m}^- - \mathbb{E}(\varphi_{k_0,m}^-)\right\|_\beta \left\|[\theta_h^{(i_1)} \circ \epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)} \circ \epsilon_{h+k_0-i_2}] - \mathbb{E}\left([\theta_h^{(i_1)} \circ \epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)} \circ \epsilon_{h+k_0-i_2}]\right)\right\|_\beta \alpha_\epsilon(\lfloor h/2\rfloor)^{1-2/\beta}, \tag{A.25}$$

for some constant K. The terms involved on the right-hand side of the above inequality are bounded as follows. First, thanks to Minkoswki's inequality followed by Cauchy Schwarz as well as (A.1), we get

$$\begin{aligned}
\left\|\varphi_{k_0,m}^- - \mathbb{E}(\varphi_{k_0,m}^-)\right\|_\beta &\leq \|Z_{0,m}Y_{k_0,m}\|_\beta + |\mathbb{E}(Z_{0,m}Y_{k_0,m})| && \text{(A.26)}\\
&\leq \|Z_{0,m}\|_{2\beta}\|Y_{k_0,m}\|_{2\beta} + \|Z_{0,m}\|_2\|Y_{k_0,m}\|_2 \\
&\leq \left(\sum_{i=m+1}^{\infty}\left\|\theta^{(i)}\circ\epsilon_{-i}\right\|_{2\beta}\right)\left(\sum_{i=0}^{m}\left\|\theta^{(i)}\circ\epsilon_{-i}\right\|_{2\beta}\right) \\
&\quad + \left(\sum_{i=m+1}^{\infty}\left\|\theta^{(i)}\circ\epsilon_{-i}\right\|_{2}\right)\left(\sum_{i=0}^{m}\left\|\theta^{(i)}\circ\epsilon_{-i}\right\|_{2}\right) = \mathrm{O}(\upsilon^m). && \text{(A.27)}
\end{aligned}$$

Similarly,

$$\left\|\left[\theta_h^{(i_1)} \circ \epsilon_{h-i_1}\right]\left[\theta_{h+k_0}^{(i_2)} \circ \epsilon_{h+k_0-i_2}\right] - \mathbb{E}\left(\left[\theta_h^{(i_1)} \circ \epsilon_{h-i_1}\right]\left[\theta_{h+k_0}^{(i_2)} \circ \epsilon_{h+k_0-i_2}\right]\right)\right\|_\beta \leq \mathrm{O}(\upsilon^{i_1+i_2}) \tag{A.28}$$

Plugging (A.27) and (A.28) into (A.25) yields their upper bound

$$|I_1(h,m)| \leq K\upsilon^m\alpha_\epsilon\left(\lfloor h/2\rfloor\right)^{1-2/\beta}, \tag{A.29}$$

for some constant K. Similarly, using again (A.27) and (A.28) yields

$$\begin{aligned}
|I_2(h,m)| &\leq K\upsilon^{\lfloor h/2\rfloor}\upsilon^m, && \text{(A.30)}\\
|I_3(h,m)| &\leq K\upsilon^{\lfloor h/2\rfloor}\upsilon^m, && \text{(A.31)}\\
|I_4(h,m)| &\leq K\upsilon^{2\lfloor h/2\rfloor}\upsilon^m, && \text{(A.32)}
\end{aligned}$$

for some constant K. Using (A.24) and plugging the bounds, (A.29), (A.30), (A.31) and (A.32) we thus obtain

$$\begin{aligned}
&\sum_{h/2>k_0}|\mathrm{Cov}([Y_{0,m} - Y_0]Y_{k_0,m}, Y_{h,m}Y_{h+k_0,m})| \\
&\quad \leq K\upsilon^m\sum_{h/2>k_0}\left[\alpha_\epsilon\left(\lfloor h/2\rfloor\right)^{1-2/\beta} + \upsilon^{\lfloor h/2\rfloor} + \upsilon^{2\lfloor h/2\rfloor}\right] \xrightarrow[m\to\infty]{} 0, && \text{(A.33)}
\end{aligned}$$

for some constant K. Since it is easy to check that

$$|\mathrm{Cov}\left([Y_{0,m}-Y_0]Y_{k_0,m}, Y_{h,m}Y_{h+k_0,m}\right)| \xrightarrow[m\to\infty]{} 0,$$

for the finite number of terms h such that $h/2 \leq k_0$, one then deduces that (A.22) holds. And all in all, we thus proved (A.21), ensuring that $\mathfrak{S}_{k_0}^m \longrightarrow \mathfrak{S}_{k_0}$ as $m \to \infty$.
In order to prove (A.20), we notice from (A.16) that $\mathcal{Y}_{n,m}(k_0)$ depends on ϵ_i, $i = n-m, ..., n+k_0$, and $\theta_n^{(i)}$, $\theta_{n+k_0-1}^{(i)}$, $\theta_{n+k_0}^{(i)}$, $i = 0, ..., m$. We already observed that the operators $\theta_q^{(i)}$, $i \in \mathbb{N}$, $q \in \mathbb{Z}$ and $\theta_p^{(j)}$, $j \in \mathbb{N}$, $p \in \mathbb{Z}$ are independent as soon as $q - i \neq p - j$. Remembering that $\{(\theta_n^{(i)})_{i\in\mathbb{N}},\ n \in \mathbb{Z}\}$ and $\{\epsilon_n,\ n \in \mathbb{Z}\}$ are independent, we hence obtain that

$$\alpha_{m,\mathcal{Y}}(h) \leq \alpha_\epsilon(h - k_0 - m) \tag{A.34}$$

as soon as $k_0+m < h$, so that (A.20) holds thanks to Assumption (**A1**). All in all, the finiteness of $\mathfrak{S}_{k_0}^m$ and (A.20) imply the convergence (A.17) thanks to Herrndorf (1984, Corollary 1).

Appendix A.5.0.3. $\diamond$ *Step 3:.* We then prove that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Z}_{i,m}(k_0)\right\| > \epsilon\right) = 0, \quad \forall \epsilon > 0, \tag{A.35}$$

where the norm $\|\cdot\|$ is defined here on $\mathbb{R}^3$ by $\|u\|^2 = u'u$, $u \in \mathbb{R}^3$ a column vector. The result (A.35) follows from a straightforward adaptation of Anderson (1971, Theorem 7.7.1 and Corollary 7.7.1, pages 425-426). Indeed, Markov's inequality as well as the stationarity of $\{\mathcal{Z}_{i,m}(k_0),\ i \in \mathbb{Z}\}$ yields

$$\mathbb{P}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Z}_{i,m}(k_0)\right\| > \epsilon\right) \leq \frac{1}{n\epsilon^2}\mathbb{E}\left(\left\|\sum_{i=1}^n \mathcal{Z}_{i,m}(k_0)\right\|^2\right) = \frac{1}{n\epsilon^2}\mathbb{E}\left(\sum_{1\leq i,j\leq n} \mathcal{Z}_{i,m}(k_0)'\mathcal{Z}_{j,m}(k_0)\right)$$
$$\xrightarrow[n\to\infty]{} \frac{1}{\epsilon^2}\sum_{h=-\infty}^{\infty} \mathbb{E}\left(\mathcal{Z}_{0,m}(k_0)'\mathcal{Z}_{h,m}(k_0)\right), \tag{A.36}$$

where the latter limit is obtained similarly as in (A.19), provided that $\sum_{h=1}^\infty \mathbb{E}\left(\mathcal{Z}_{0,m}(k_0)'\mathcal{Z}_{h,m}(k_0)\right)$ is a convergent series that tends to 0 as $m \to \infty$. In view of the definition (A.15), this is the case if we prove that the following series:

$$\sum_{h=1}^\infty \mathrm{Cov}(Z_{0,m}Y_{k_0-t_2,m}^{t_1}, \Upsilon_{h,m}^{t_3}), \quad \sum_{h=1}^\infty \mathrm{Cov}(Y_{0,m}Z_{k_0-t_2,m}, \Upsilon_{h,m}^{t_3}), \quad \sum_{h=1}^\infty \mathrm{Cov}(Z_{0,m}Z_{k_0-t_2,m}, \Upsilon_{h,m}^{t_3}),$$
$$\text{where } \Upsilon_{h,m}^{t_3} \in \{Z_{h,m}Y_{h+k_0-t_3,m}, Y_{h,m}Z_{h+k_0-t_3,m}, Z_{h,m}Y_{h+k_0-t_3,m}\},$$

converge to 0 as $m \to \infty$ for t_1,t_2 and t_3 in $\{0,1\}$. We proceed to do prove this convergence for the series with general term $\mathrm{Cov}(Z_{0,m}Z_{k_0,m}, Z_{h,m}Z_{h+k_0,m})$, the other cases being similar. As previously while studying (A.24), we write that

$$\mathrm{Cov}(Z_{0,m}Z_{k_0,m}, Z_{h,m}Z_{h+k_0,m}) = \sum_{i_1=m+1}^\infty \sum_{i_2=m+1}^\infty \mathrm{Cov}\left(Z_{0,m}Z_{k_0,m}, [\theta_h^{(i_1)} \circ \epsilon_{h-i_1}][\theta_{h+k_0}^{(i_2)} \circ \epsilon_{h+k_0-i_2}]\right).$$

Splitting the above double sum into i_1 and i_2 less or larger than $h/2$, the following equivalent of (A.33) may be obtained:

$$\sum_{h/2>k_0} |\mathrm{Cov}(Z_{0,m}Z_{k_0,m}, Z_{h,m}Z_{h+k_0,m})| \leq K\upsilon^m \sum_{h/2>k_0} \left[\alpha_\epsilon(\lfloor h/2 \rfloor)^{1-2/\beta} + \upsilon^{\lfloor h/2 \rfloor} + \upsilon^{2\lfloor h/2 \rfloor}\right] = \mathrm{O}(\upsilon^m) \tag{A.37}$$

for some constant $K > 0$. It can be also proved that for $h/2 \leq k_0$ we have $|\mathrm{Cov}(Z_{0,m}Z_{k_0,m}, Z_{h,m}Z_{h+k_0,m})| = \mathrm{O}(\upsilon^m)$, which, combined with (A.37), yields $\sum_{h=1}^{\infty} |\mathrm{Cov}(Z_{0,m}Z_{k_0,m}, Z_{h,m}Z_{h+k_0,m})| = \mathrm{O}(\upsilon^m)$, from which we in turn get

$$\sum_{h=-\infty}^{\infty} \mathbb{E}\left(\mathcal{Z}_{0,m}(k_0)'\mathcal{Z}_{h,m}(k_0)\right) = \mathrm{O}(\upsilon^m). \tag{A.38}$$

Combining the above with (A.36) thus yields (A.35).

Appendix A.5.0.4. $\diamond$ *Step 4:.* Combining (A.17) and (A.35) yields by Brockwell and Davis (1991, Proposition 6.3.9 p.208) the final result (22). $\square$

As a side result of the previous proof, let us prove that the following convergence holds:

$$\frac{1}{n}\mathbb{E}(T_n(k_0)T_n(k_0)') \xrightarrow[n\to\infty]{} \mathfrak{S}_{k_0} \tag{A.39}$$

which will be useful later on. For this, we write, remembering the notation (20) and (A.16),

$$\begin{aligned}
\left\|\mathbb{E}\left(\frac{T_n(k_0)T_n(k_0)'}{n}\right) - \mathfrak{S}_{k_0}\right\| &\leq \left\|\mathbb{E}\left(\frac{T_n(k_0)T_n(k_0)'}{n}\right) - \mathfrak{S}_{k_0}^m\right\| + \left\|\mathfrak{S}_{k_0}^m - \mathfrak{S}_{k_0}\right\| \\
&\leq \left\|\mathbb{E}\left(\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0) + \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Z}_{i,m}(k_0)\right]\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0) + \frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Z}_{i,m}(k_0)\right]'\right) - \mathfrak{S}_{k_0}^m\right\| \\
&\qquad + \left\|\mathfrak{S}_{k_0}^m - \mathfrak{S}_{k_0}\right\| \\
&\leq \left\|\mathbb{E}\left(\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0)\right]\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0)\right]'\right) - \mathfrak{S}_{k_0}^m\right\| \\
&\qquad + 2\left\|\mathbb{E}\left(\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0)\right]\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Z}_{i,m}(k_0)\right]'\right)\right\| + \left\|\mathfrak{S}_{k_0}^m - \mathfrak{S}_{k_0}\right\|.
\end{aligned} \tag{A.40}$$

We now observe that, thanks to (A.19),

$$\left\|\mathbb{E}\left(\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0)\right]\left[\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0)\right]'\right) - \mathfrak{S}_{k_0}^m\right\| \xrightarrow[n\to\infty]{} 0, \quad \forall m \in \mathbb{N}. \tag{A.41}$$

Next, we have that $\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Y}_{i,m}(k_0)\right\|_2$ is upper bounded by a constant C independent from n and m thanks to (A.19) and the fact that $\mathfrak{S}_{k_0}^m$ converges to the finite limit $\mathfrak{S}_{k_0}$ as $m \to \infty$. Also a byproduct of (A.36) and (A.38) is that $\lim_{n\to\infty}\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^n \mathcal{Z}_{i,m}(k_0)\right\|_2 = \sum_{h=-\infty}^{\infty} \mathbb{E}\left(\mathcal{Z}_{0,m}(k_0)'\mathcal{Z}_{h,m}(k_0)\right) = \mathrm{O}(\upsilon^m)$. Thus, Cauchy Schwarz's inequality yields that

$$\left\|\mathbb{E}\left(\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{Y}_{i,m}(k_0)\right]\left[\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{Z}_{i,m}(k_0)\right]'\right)\right\| \leq \mathbb{E}\left(\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{Y}_{i,m}(k_0)\right\|\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{Z}_{i,m}(k_0)\right\|\right)$$
$$\leq \left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{Y}_{i,m}(k_0)\right\|_2\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{Z}_{i,m}(k_0)\right\|_2 \leq C\left\|\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\mathcal{Z}_{i,m}(k_0)\right\|_2$$
$$\xrightarrow[n\to\infty]{} C\sum_{h=-\infty}^{\infty}\mathbb{E}\left(\mathcal{Z}_{0,m}(k_0)'\mathcal{Z}_{h,m}(k_0)\right) = \mathrm{O}(\upsilon^m). \quad \text{(A.42)}$$

Thus, (A.41), (A.42) along with $\lim_{m\to\infty}\mathfrak{S}_{k_0}^m = \mathfrak{S}_{k_0}$ yields from (A.40) that

$$\limsup_{n\to\infty}\left\|\mathbb{E}\left(\frac{T_n(k_0)T_n(k_0)'}{n}\right) - \mathfrak{S}_{k_0}\right\| \leq \mathrm{O}(\upsilon^m) + \left\|\mathfrak{S}_{k_0}^m - \mathfrak{S}_{k_0}\right\| \xrightarrow[m\to\infty]{} 0,$$

which implies (A.39).

Appendix A.6. Proof of Proposition 5.1 and Theorem 5.2

Proof of Proposition 5.1. Remembering the expression (20), we recall that $\mathcal{Y}_n = \mathcal{Y}_n(k_0) = (Y_n - C_1, Y_nY_{n+k_0-1} - u_{k_0-2}, Y_nY_{n+k_0} - u_{k_0-1})'$. Writing the $\mathbb{R}^3$ vector $\Sigma_\varepsilon^{-1}\varepsilon_0$ in the form $\Sigma_\varepsilon^{-1}\varepsilon_0 = \left(\varepsilon_0^{(1)}, \varepsilon_0^{(2)}, \varepsilon_0^{(3)}\right)'$, we thus have from (28) that

$$\Psi_n = \mathrm{Cov}(\mathcal{Y}_n, \Sigma_\varepsilon^{-1}\varepsilon_0),$$

which is a $\mathbb{R}^{3\times 3}$ matrix. We first prove that

$$\sum_{n=1}^{\infty}\|\Psi_n\| < \infty. \quad \text{(A.43)}$$

The first row, third line of $\|\Psi_n\|$ is equal to $\mathrm{Cov}(\varepsilon_0^{(1)}, Y_nY_{n+k_0})$. Since $\varepsilon_0^{(1)}$ depends on Y_j, $j \leq k_0$, an argument similar to (A.14) yields that

$$\left|\mathrm{Cov}(\varepsilon_0^{(1)}, Y_nY_{n+k_0})\right| \leq K\left[\alpha_\epsilon\left(\left\lfloor\frac{n-k_0-1}{2}\right\rfloor\right)^{1-2/\beta} + \lambda_0^{\left\lfloor\frac{n-k_0-1}{2}\right\rfloor} + \lambda_0^n\right]$$

And, more generally one obtains a similar bound for all coefficients of Ψ_n, which thus implies (A.43).

We next introduce the spectral density associated to the process $\{\mathcal{Y}_n,\ n\in\mathbb{Z}\} := \{\mathcal{Y}_n(k_0),\ n\in\mathbb{Z}\}$ defined as

$$f_{\mathcal{Y}}(x) := \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}C(h)e^{ihx} \quad \text{where } C(h) := \mathbb{E}(\mathcal{Y}_0\mathcal{Y}_h') = \mathrm{Cov}(\mathcal{Y}_0, \mathcal{Y}_h). \quad \text{(A.44)}$$

We note that the above series is indeed convergent thanks to arguments similar to the ones leading to the convergence of (A.43). Also note that, still thanks to that latter convergence, we have the relation

$$f_{\mathcal{Y}}(x) = \overline{\Psi(e^{ix})}\Sigma_\varepsilon\Psi(e^{ix})', \quad \forall x\in\mathbb{R}, \quad \text{(A.45)}$$

where we recall that $\Psi(\cdot)$ is defined by (29). We first prove that

$$\det\left(\Psi(z)\right) \neq 0, \quad \forall z \in \mathbb{C} \text{ such that } |z| = 1. \tag{A.46}$$

In view of (A.45), the above is satisfied if and only if we show that $f_{\mathcal{Y}}(x)$ is invertible for all $x \in [0, 2\pi]$, which we proceed to do if λ_0 is small enough. For this we will momentarily write the dependence of the spectral density of $\{\mathcal{Y}_n,\ n \in \mathbb{Z}\}$ and write $f_{\mathcal{Y}}(\lambda_0, x)$ instead of $f_{\mathcal{Y}}(x)$, and we are going to prove that

$$\sup_{x\in[0,2\pi]} \|f_{\mathcal{Y}}(\lambda_0, x) - f_{\underline{\epsilon}}(x)\| \longrightarrow 0, \quad \lambda_0 \to 0. \tag{A.47}$$

In view of the definitions of those spectral densities, it is not difficult to check that this amounts to show that the quantities

$$\begin{array}{c} \sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0, Y_h) - \mathrm{Cov}(\epsilon_0, \epsilon_h)|, \\ \sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0, Y_h Y_{h+k_0-1}) - \mathrm{Cov}(\epsilon_0, \epsilon_h \epsilon_{h+k_0-1})|, \\ \sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0, \epsilon_h \epsilon_{h+k_0})|, \\ \sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0-1}, Y_h Y_{h+k_0-1}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0-1}, \epsilon_h \epsilon_{h+k_0-1})|, \\ \sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0-1}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0-1}, \epsilon_h \epsilon_{h+k_0})|, \\ \sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})| \end{array} \tag{A.48}$$

tend to 0 as λ_0 tends to 0: we prove this only for $\sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})|$, the other cases being dealt with similarly. We first note that

$$\begin{aligned} \mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) &= \lambda_0 [\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0-1}) + \mathrm{Cov}(Y_0 Y_{k_0}, Y_{h-1} \epsilon_{h+k_0}) \\ &+ \mathrm{Cov}(Y_0 Y_{k_0-1}, \epsilon_h \epsilon_{h+k_0}) + \mathrm{Cov}(Y_{-1} \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})] + \mathrm{Cov}(\epsilon_{k_0} \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0}), \quad h > k_0. \end{aligned} \tag{A.49}$$

Now, (A.14) implies that $|\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0-1})| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{h-k_0-1}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\left\lfloor \frac{h-k_0-1}{2} \right\rfloor} \right]$ where K can be verified to be uniformly bounded in $\lambda_0 \in [\lambda_-, \lambda_+]$. This inequality along with the assumption (32) entails that $|\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0-1})| \leq K \lambda_+^{\left\lfloor \frac{h-k_0-1}{2} \right\rfloor}$ for $h > k_0$, so that $\sum_{h=k_0+1}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0-1})|$ is finite and bounded in $\lambda_0 \in [\lambda_-, \lambda_+]$. A similar argument and bound holds for $\sum_{h=k_0+1}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_{h-1} \epsilon_{h+k_0})|$, $\sum_{h=k_0}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0-1}, \epsilon_h \epsilon_{h+k_0})|$ and $\sum_{h=k_0+1}^{\infty} |\mathrm{Cov}(Y_{-1} \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})|$. All in all, one thus deduce from (A.49) that

$$\sum_{h=k_0+1}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})| = \mathrm{O}(\lambda_0).$$

Besides, one can verify that $|\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})| = \mathrm{O}(\lambda_0)$ for $h = 0, ..., k_0$ so that one finally obtains that

$$\sum_{h=0}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})| = \mathrm{O}(\lambda_0).$$

A similar argument holds for $\sum_{h=-\infty}^{0} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})|$, so that we eventually have that $\sum_{h=-\infty}^{\infty} |\mathrm{Cov}(Y_0 Y_{k_0}, Y_h Y_{h+k_0}) - \mathrm{Cov}(\epsilon_0 \epsilon_{k_0}, \epsilon_h \epsilon_{h+k_0})| = \mathrm{O}(\lambda_0)$ as λ_0 tends to 0. As explained earlier, all series in (A.48) can be shown to be also $\mathrm{O}(\lambda_0)$, so that (A.47)

is proved. This in turn implies that $\lim_{\lambda_0 \to 0} \sup_{x \in [0,2\pi]} |\det f_{\mathcal{Y}}(\lambda_0, x) - \det f_{\underline{\epsilon}}(x)| = 0$. Hence we have that, for λ_0 small enough:

$$\inf_{x \in [0,2\pi]} |\det f_{\mathcal{Y}}(\lambda_0, x)| \geq \inf_{x \in [0,2\pi]} |\det f_{\underline{\epsilon}}(x)| - \sup_{x \in [0,2\pi]} |\det f_{\mathcal{Y}}(\lambda_0, x) - \det f_{\underline{\epsilon}}(x)| > 0,$$

so that (A.46) is proved. We now prove that (**A4**) holds, i.e. that $\det(\Psi(z)) \neq 0$ not just on z such that $|z| = 1$ but on the broader set $\{z \in \mathbb{C} |\ |z| \leq 1\}$. This is the most delicate part of the proof, and the main argument is an analysis related to the measurability with respect to the immigration sequences. We first observe that (A.46) implies that the expansion (27) can be inverted and there exists a sequence of $\mathbb{R}^{3\times 3}$ matrices $(\phi_n)_{n \in \mathbb{Z}}$ such that

$$\varepsilon_n = \sum_{j \in \mathbb{Z}} \phi_j \mathcal{Y}_{n-j}, \quad n \in \mathbb{Z}, \tag{A.50}$$

and we are going to prove that $\phi_j = 0$ for all $j \leq -1$, so that the expansion (30) holds with $\phi_0 = \mathbf{I_3}$ and $\Phi_j = -\phi_j$, $j \geq 1$, which in turn is equivalent to Condition (**A4**). We first recall that, for all $n \in \mathbb{Z}$, $\mathcal{Y}_n$ depends on the random variables ϵ_j and $\{\xi_{j,i}\ i \in \mathbb{N}\}$, $j \leq n + k_0$ (see (1) as well as the expression (20) for $\mathcal{Y}_n$). Next, we write, using the definition (28) for ε_n and the expansion (A.50):

$$\mathcal{Y}_n - P_{n-1}\mathcal{Y}_n - \sum_{j=0}^{\infty} \phi_j \mathcal{Y}_{n-j} = \sum_{j=-\infty}^{-1} \phi_j \mathcal{Y}_{n-j}. \tag{A.51}$$

We now prove that the righthanside of (A.51) is 0 by using arguments related to the indepence of the sequences $\{\xi_{n,k},\ k \in \mathbb{N}\}$ when $n \in \mathbb{N}$. To avoid tedious notation we will denote by $\xi_{n,.}$ the latter sequence in the rest of the proof. We define the following $\sigma-$ algebras

$$\mathcal{G}_n := \sigma(Y_k, k \leq n,\ \xi_{j,.}, j \geq n+1), \quad n \in \mathbb{Z}, \tag{A.52}$$

and we proceed to express the conditional expectation of $\mathcal{Y}_{n-j}$ given $\mathcal{G}_n$ in function of $j \in \mathbb{Z}$ in order to take the conditional expectation on both sides of (A.51). We observe that

$$\mathbb{E}(Y_{n-j}|\mathcal{G}_n) = \begin{cases} Y_{n-j} & \text{if } j \geq 0, \\ G^{(1)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j) & \text{if } j \leq -1. \end{cases} \tag{A.53}$$

Here $G^{(1)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j)$ is a random variable that depends on the sequences $\xi_{k,.}, k = n+1, ..., n-j$ as well as on the random variables Y_k, $k \leq n$; however, only the dependence on $\xi_{k,.}$ is mentioned in order to avoid tedious notation. Similarly, we have

$$\mathbb{E}(Y_{n-j}Y_{n-j+k_0-1}|\mathcal{G}_n) = \begin{cases} Y_{n-j}Y_{n-j+k_0-1} & \text{if } j \geq k_0 - 1, \\ G^{(2)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j+k_0-1) & \text{if } j \leq -1, \\ Y_{n-j}\tilde{G}^{(2)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j+k_0-1) & \text{if } j = 0, ..., k_0-2, \end{cases} \tag{A.54}$$

$$\mathbb{E}(Y_{n-j}Y_{n-j+k_0}|\mathcal{G}_n) = \begin{cases} Y_{n-j}Y_{n-j+k_0} & \text{if } j \geq k_0, \\ G^{(3)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j+k_0) & \text{if } j \leq -1, \\ Y_{n-j}\tilde{G}^{(3)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j+k_0) & \text{if } j = 0, ..., k_0-1, \end{cases} \tag{A.55}$$

where $\tilde{G}^{(2)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j+k_0-1) = \mathbb{E}(Y_{n-j+k_0-1}|\mathcal{G}_n)$ and $\tilde{G}^{(3)}_{-j}(\xi_{k,.}, k = n+1, ..., n-j+k_0) = \mathbb{E}(Y_{n-j+k_0}|\mathcal{G}_n)$. We now note that, by definition of the projection operator P_{n-1},

there exists a sequence of combinations $\left(\sum_{j=1}^{N_m}\alpha_j^{(m)}\mathcal{Y}_{n-j}\right)_{m\in\mathbb{N}}$ such that, $\alpha_j^{(m)}$ is a $\mathbb{R}^{3\times 3}$ matrix and $\sum_{j=1}^{N_m}\alpha_j^{(m)}\mathcal{Y}_{n-j}\xrightarrow{\mathbb{L}^2}P_{n-1}\mathcal{Y}_n$ as $m\to\infty$. Thus, the lefthanside of (A.51) is the limit in $\mathbb{L}^2$ of an (potentially infinite) combination $\sum_{j=1}^{\infty}\phi_j^{(m)}\mathcal{Y}_{n-j}$ as $m\to\infty$, where $(\phi_j^{(m)})_{j\in\mathbb{N}}$ are $\mathbb{R}^{3\times 3}$ matrices given by $\phi_0^{(m)}=I_3-\alpha_0^{(m)}$, $\phi_j^{(m)}=-\alpha_j^{(m)}-\phi_j$, $j=1,...,N_m$, and $\phi_j^{(m)}=-\phi_j$, $j>N_m$. Then, taking the conditional expectation in (A.51) reads

$$\lim_{m\to\infty}\sum_{j=0}^{\infty}\phi_j^{(m)}\mathbb{E}(\mathcal{Y}_{n-j}|\mathcal{G}_n)=\sum_{j=-\infty}^{-1}\phi_j\mathbb{E}(\mathcal{Y}_{n-j}|\mathcal{G}_n). \tag{A.56}$$

Let us now write the matrices ϕ_j, $j\le -1$, and $\phi_j^{(m)}$, $j\ge 0$, in the form of $\phi_j=[C_j^{(1)},C_j^{(2)},C_j^{(3)}]$ and $\phi_j^{(m)}=[C_j^{(1)(m)},C_j^{(2)(m)},C_j^{(3)(m)}]$ where $C_j^{(i)}$, $C_j^{(i)(m)}$, $i=1,2,3$, are 3×1 column vectors. We next identify in both sides of (A.56) the terms that feature the sequence $\xi_{n+1,.}$, and only this sequence, among $\xi_{k,.}$, $k\ge n+1$. From the expressions (A.53), (A.54) and (A.55), we may see that the only term on the righthandside of (A.56) with that property is $C_{-1}^{(1)}G_{-1}^{(1)}(\xi_{n+1,.})=C_{-1}^{(1)}\mathbb{E}(Y_{n+1}|\mathcal{G}_n)=C_{-1}^{(1)}\left[\sum_{r=1}^{Y_n}\xi_{n+1,r}+\mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n)\right]$, with $\mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n)$ depending only on Y_j, $j\le n$ because of the independence of ϵ_{n+1} with $\xi_{j,.}$, $j\ge n+1$. On the lefthandside, and in view of the expressions (A.54) and (A.55), the only case for $\mathbb{E}(Y_{n-j}Y_{n-j+k_0-1}|\mathcal{G}_n)$ and $\mathbb{E}(Y_{n-j}Y_{n-j+k_0}|\mathcal{G}_n)$ to depend on $\xi_{n+1,.}$ is when j is respectively equal to k_0-2 and k_0-1: in that case $\mathbb{E}(Y_{n-(k_0-2)}Y_{n+1}|\mathcal{G}_n)=Y_{n-(k_0-2)}\tilde{G}^{(2)}_{-(k_0-2)}(\xi_{n+1,.})=Y_{n-(k_0-2)}\mathbb{E}(Y_{n+1}|\mathcal{G}_n)=Y_{n-(k_0-2)}\left[\sum_{r=1}^{Y_n}\xi_{n+1,r}+\mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n)\right]$ and $\mathbb{E}(Y_{n-(k_0-1)}Y_{n+1}|\mathcal{G}_n)=Y_{n-(k_0-1)}\tilde{G}^{(3)}_{-(k_0-2)}(\xi_{n+1,.})=Y_{n-(k_0-1)}\left[\sum_{r=1}^{Y_n}\xi_{n+1,r}+\mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n)\right]$. Isolating these particular terms, one thus checks that (A.56) can be equivalently written as

$$\begin{aligned}\mathcal{E}_m:=&-C_{-1}^{(1)}G_{-1}^{(1)}(\xi_{n+1,.})+C_{-(k_0-2)}^{(2)(m)}Y_{n-(k_0-2)}\tilde{G}^{(2)}_{-(k_0-2)}(\xi_{n+1,.})\\&+C_{-(k_0-1)}^{(3)(m)}Y_{n-(k_0-1)}\tilde{G}^{(3)}_{-(k_0-1)}(\xi_{n+1,.})+X^{(m)}\\=&\left[-C_{-1}^{(1)}+C_{-(k_0-2)}^{(2)(m)}Y_{n-(k_0-2)}+C_{-(k_0-1)}^{(3)(m)}Y_{n-(k_0-1)}\right]\left[\sum_{r=1}^{Y_n}\xi_{n+1,r}+\mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n)\right]\\&+X^{(m)}\xrightarrow{\mathbb{L}^2}0,\quad m\to\infty,\end{aligned} \tag{A.57}$$

where $X^{(m)}$ is a random vector which is the sum of

- a random vector that involves the terms $\mathbb{E}(Y_{n-j}|\mathcal{G}_n)$, $j\ge 0$, $\mathbb{E}(Y_{n-j}Y_{n-j+k_0-1}|\mathcal{G}_n)$, $j\ge k_0-1$, $\mathbb{E}(Y_{n-j}Y_{n-j+k_0}|\mathcal{G}_n)$, $j\ge k_0$, that are respectively equal to Y_{n-j}, $Y_{n-j}Y_{n-j+k_0-1}$ and $Y_{n-j}Y_{n-j+k_0}$, and that thus depend on Y_k, $k\le n$ only,
- and of a random vector that figures terms of the form $\mathbb{E}(Y_{n+j}|\mathcal{G}_n)$, $\mathbb{E}(Y_{n+j-(k_0-1)}Y_{n+j}|\mathcal{G}_n)$, $\mathbb{E}(Y_{n+j-(k_0-2)}Y_{n+j}|\mathcal{G}_n)$, $j\ge 2$, that thus depend on Y_k, $k\le n$ and $\xi_{k,.}$, $k\ge n+1$.

Let us now suppose that

$$\limsup_{m\to\infty}\|-C_{-1}^{(1)}+C_{-(k_0-2)}^{(2)(m)}Y_{n-(k_0-2)}+C_{-(k_0-1)}^{(3)(m)}Y_{n-(k_0-1)}\|>0 \tag{A.58}$$

on a non negligible set. Then, from (A.57), the 3×1 vectors $D_m := -C_{-1}^{(1)} + C_{-(k_0-2)}^{(2)(m)} Y_{n-(k_0-2)} + C_{-(k_0-1)}^{(3)(m)} Y_{n-(k_0-1)}$, $\mathcal{E}_m$ and $X^{(m)}$ are such that there exists $i_0 \in \{1,2,3\}$ and $\limsup_{m\to\infty} |D_{m,i_0}| > 0$, where D_{m,i_0} is the i_0th entries . With the same notation for $\mathcal{E}_{m,i_0}$ and $X_{i_0}^{(m)}$, we obtain by taking the limsup in (A.57) (observing that $\sum_{r=1}^{Y_n} \xi_{n+1,r} + \mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n)$ is non negative) that

$$\left[\limsup_{m\to\infty} |D_{m,i_0}|\right] \cdot \left[\sum_{r=1}^{Y_n} \xi_{n+1,r} + \mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n)\right] = \limsup_{m\to\infty} \left|\mathcal{E}_{m,i_0} - X_{i_0}^{(m)}\right|. \tag{A.59}$$

Up to a subsequence, we have from (A.57) that $\mathcal{E}_m \longrightarrow 0$ a.s. as $m \to \infty$; Thus we have that $\mathcal{E}_{m,i_0} \longrightarrow 0$. So that, on a non negligible set, (A.59) yields

$$\sum_{r=1}^{Y_n} \xi_{n+1,r} + \mathbb{E}(\epsilon_{n+1}|\mathcal{G}_n) = \frac{\limsup_{m\to\infty} \left|X_{i_0}^{(m)}\right|}{\limsup_{m\to\infty} |D_{m,i_0}|}. \tag{A.60}$$

Let us now define the operator $\mathcal{F}(\xi_{j,.}, \epsilon_j) : y \in \mathbb{N} \mapsto \sum_{r=1}^{y} \xi_{j,r} + \epsilon_j = \theta_j \circ y + \epsilon_j$ for all $j \in \mathbb{Z}$. We may thus write that, for all $k \geq 2$,

$$Y_{n+k} = \mathcal{F}(\xi_{n+k,.}, \epsilon_{n+k}) \circ \ldots \circ \mathcal{F}(\xi_{n+1,.}, \epsilon_{n+1})(Y_n), \tag{A.61}$$

so that, obviously, Y_{n+k}, $Y_{n+k-(k_0-2)}$, $Y_{n+k-(k_0-1)}$, and thus also their conditional expectation given $\mathcal{G}_n$, depend on the sequences $\xi_{n+1,.}$, $\xi_{n+2,.}$,..., $\xi_{n+k,.}$ for all $k \geq 2$. More importantly a crucial property is that they depend $\xi_{n+1,.}$ if and only if they depend on say $\xi_{n+2,.}$, in view of the expression of the operators $\mathcal{F}(\xi_{j,.}, \epsilon_j)$ and the expression (A.61). Thus, the righthandside of (A.60) in particular depends on $\xi_{n+1,.}$ if and only if it depends on $\xi_{n+2,.}$, which is a contradiction with the expression on the lefthandside which only depends on $\xi_{n+1,.}$ (in addition to the Y_k, $k \leq n$, which are independent from the the $\xi_{k,.}$, $k \geq n+1$). Thus, (A.58) is violated and we have

$$\lim_{m\to\infty} -C_{-1}^{(1)} + C_{-(k_0-2)}^{(2)(m)} Y_{n-(k_0-2)} + C_{-(k_0-1)}^{(3)(m)} Y_{n-(k_0-1)} = 0 \quad \text{a.s.} \tag{A.62}$$

Let us now notice that the sequences $(C_{-(k_0-2)}^{(2)(m)})_{m\in\mathbb{N}}$ and $(C_{-(k_0-1)}^{(3)(m)})_{m\in\mathbb{N}}$ are then necessarily bounded. Indeed, if say $(C_{-(k_0-2)}^{(2)(m)})_{m\in\mathbb{N}}$ was not bounded, then we may suppose that $\|C_{-(k_0-2)}^{(2)(m)}\| \longrightarrow \infty$ up to a subsequence. Then (A.62) implies that $\|C_{-(k_0-1)}^{(3)(m)}\| \longrightarrow \infty$, and, dividing on both sides of (A.62) by $\|C_{-(k_0-1)}^{(3)(m)}\|$ and letting $m \to \infty$ yields that the (deterministic) limit $L := \lim_{m\to\infty} C_{-(k_0-2)}^{(2)(m)} / \|C_{-(k_0-1)}^{(3)(m)}\|$ exists and that $Y_{n-(k_0-1)} = L Y_{n-(k_0-2)}$, which is not possible by an independence argument. Hence the boundedness of the two sequences, from which one gets that $(C_{-(k_0-2)}^{(2)(m)})_{m\in\mathbb{N}}$ and $(C_{-(k_0-1)}^{(3)(m)})_{m\in\mathbb{N}}$ converge (up to a subsequence) towards some limits L_2 and L_3, which in turn yields from (A.62) that $-C_{-1}^{(1)} + L_2 Y_{n-(k_0-2)} + L_3 Y_{n-(k_0-1)} = 0$ a.s. An indenpendence argument again yields that L_2, L_3 and, more importantly, that $C_{-1}^{(1)}$ are equal to zero. Similar arguments yield that $C_{-1}^{(2)} = C_{-1}^{(3)} = 0$ i.e. $\phi_{-1} = 0$. Similarly, one proves that $\phi_{-j} = 0$ for all $j \geq 2$.

We now prove that $\|\Phi_k\| = o(1/k^2)$ as $k \to \infty$. We define, as in (A.44), the function

$$g_{\mathcal{Y}}(z) := \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} C(h) z^h, \quad C(h) := \mathrm{Cov}(\mathcal{Y}_0, \mathcal{Y}_h),$$

which is convergent when $\lambda_0^{1/2} < |z| < 1/\lambda_0^{1/2}$ thanks to arguments similar to the ones leading to the convergence of (A.43). As in (A.45), one has the relation

$$g_{\mathcal{Y}}(z) = \Psi(1/z)\Sigma_\varepsilon \Psi(z)', \quad \lambda_0^{1/2} < |z| < 1/\lambda_0^{1/2}. \tag{A.63}$$

Note now that the assumption that $f_{\underline{\epsilon}}$ is invertible on $[0, 2\pi]$ implies that there exists $\delta > 0$ small enough such that $\mathcal{C}_\delta := \{z \in \mathbb{C} |\ 1 - \delta \le |z| \le 1 + \delta\} \subset \left\{z \in \mathbb{C} |\ \lambda_0^{1/2} < |z| < 1/\lambda_0^{1/2}\right\}$ and the function

$$g_{\underline{\epsilon}} : z \mapsto \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \mathrm{Cov}(\underline{\epsilon}_0, \underline{\epsilon}_h) z^h$$

is invertible on $\mathcal{C}_\delta$. As in (A.47), one may specify the dependence of $g_{\mathcal{Y}}$ w.r.t. λ_0 and prove that $\sup_{z \in \mathcal{C}_\delta} \|g_{\mathcal{Y}}(\lambda_0, z) - g_{\underline{\epsilon}}(z)\| \longrightarrow 0$ as $\lambda_0 \to 0$ so that, if λ_0 is small enough, $g_{\mathcal{Y}}$ is invertible on $\mathcal{C}_\delta$. Thus, for such a λ_0, one may invert (A.63) and deduce in particular that $\Phi(z)$ in (31) is a convergent series on $\mathcal{C}_\delta$. This in particular implies that $\sum_{k=1}^{\infty} \|\Phi_k\|(1+\delta)^k < \infty$, which in turn implies $\|\Phi_k\| = o(1/k^2)$. □

Proof of Theorem 5.2. As announced after the statement of the theorem, the scheme of the proof is quite standard and is similar to Boubacar Mainassara et al. (2012, Section 3.3.1). Therefore, the following series of Lemmas will be proved only when their proofs are significantly different from the corresponding ones in the latter paper. We start by introducing some notation. If $r \ge 1$, we let $\Phi_r(z) = I_3 - \sum_{k=1}^{r} \Phi_{r,k} z^k$, where $\Phi_{r,1}, \ldots, \Phi_{r,r}$ denote the coefficients of the least squares regression of $\mathcal{Y}_n$ on $\mathcal{Y}_{n-1}, \ldots, \mathcal{Y}_{n-r}$ as defined in (33), and we recall that $\varepsilon_{r,n}$ is the corresponding residual. We let $\mathcal{Y}_{r,t} := (\mathcal{Y}_{t-1}', \ldots, \mathcal{Y}_{t-r}')' \in \mathbb{R}^{3r}$, and introduce the following (real or empirical) covariance matrices

$$\begin{array}{rcl} \Sigma_{\mathcal{Y}} & := & \mathbb{E}(\mathcal{Y}_0 \mathcal{Y}_0') \in \mathbb{R}^{3\times 3}, \\ \Sigma_{\mathcal{Y}_r} & := & \mathbb{E}(\mathcal{Y}_{r,0} \mathcal{Y}_{r,0}') \in \mathbb{R}^{3r\times 3r}, \\ \Sigma_{\mathcal{Y},\mathcal{Y}_r} & := & \mathbb{E}(\mathcal{Y}_0 \mathcal{Y}_{r,0}') \in \mathbb{R}^{3\times 3r}, \end{array} \tag{A.64}$$

$$\begin{array}{rcl} \hat{\Sigma}_{\mathcal{Y}} & := & \frac{1}{n}\sum_{t=1}^{n} \mathcal{Y}_t \mathcal{Y}_t' \in \mathbb{R}^{3\times 3}, \\ \hat{\Sigma}_{\mathcal{Y}_r} & := & \frac{1}{n}\sum_{t=1}^{n} \mathcal{Y}_{r,t} \mathcal{Y}_{r,t}' \in \mathbb{R}^{3r\times 3r}, \\ \hat{\Sigma}_{\mathcal{Y},\mathcal{Y}_r} & := & \frac{1}{n}\sum_{t=1}^{n} \mathcal{Y}_t \mathcal{Y}_{r,t}' \in \mathbb{R}^{3\times 3r}, \end{array} \tag{A.65}$$

$$\begin{array}{rcl} \hat{\Sigma}_{\hat{\mathcal{Y}}} & := & \frac{1}{n}\sum_{t=1}^{n} \hat{\mathcal{Y}}_t \hat{\mathcal{Y}}_t' \in \mathbb{R}^{3\times 3}, \\ \hat{\Sigma}_{\hat{\mathcal{Y}}_r} & := & \frac{1}{n}\sum_{t=1}^{n} \hat{\mathcal{Y}}_{r,t} \hat{\mathcal{Y}}_{r,t}' \in \mathbb{R}^{3r\times 3r}, \\ \hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} & := & \frac{1}{n}\sum_{t=1}^{n} \hat{\mathcal{Y}}_t \hat{\mathcal{Y}}_{r,t}' \in \mathbb{R}^{3\times 3r}, \end{array} \tag{A.66}$$

Then we have the matrix expressions for the least square projections:

$$\underline{\Phi_r} := (\Phi_{r,1}, \ldots, \Phi_{r,r})' = \Sigma_{\mathcal{Y},\mathcal{Y}_r} \Sigma_{\mathcal{Y}_r}^{-1}, \quad \underline{\hat{\Phi}_r} := (\hat{\Phi}_{r,1}, \ldots, \hat{\Phi}_{r,r})' = \hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} \hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1}. \tag{A.67}$$

We use here the matrix norm $\|A\| := \sup_{\|x\| \le 1} \|Ax\|$ for all matrix $A = (a_{ij})_{i=1,\ldots,d_1, j=1,\ldots,d_2} \in \mathbb{R}^{d_1 \times d_2}$, associated to the usual euclidean norm, and recall that it is submultiplicative and

verifies

$$\|A\|^2 \leq \sum_{i,j} a_{ij}^2. \tag{A.68}$$

We now state the following lemmas that lead to the proof of the theorem.

Lemma Appendix A.3. One has that $\sup_{r\geq 1} \|\Sigma_{\mathcal{Y}_r}\|$, $\sup_{r\geq 1} \|\Sigma_{\mathcal{Y},\mathcal{Y}_r}\|$ and $\sup_{r\geq 1} \|\Sigma_{\mathcal{Y}_r}^{-1}\|$ are finite.

Proof. We follow the proof of Boubacar Mainassara et al. (2012, Lemma 1). We first observe that for all column vector $x \in \mathbb{R}^{3r}$ we have the inequalities $\left\|\Sigma_{\mathcal{Y}_{r+1}} x\right\| \leq \left\|\Sigma_{\mathcal{Y}_{r+1}}(x', 0_3')\right\|$ and $\|\Sigma_{\mathcal{Y},\mathcal{Y}_r}(0_3', x')\|$, where $0_3 := (0,0,0)$, so that $(\|\Sigma_{\mathcal{Y}_r}\|)_{r\geq 1}$ is an increasing sequence and $\|\Sigma_{\mathcal{Y},\mathcal{Y}_r}\| \leq \|\Sigma_{\mathcal{Y}_{r+1}}\|$ for all $r \geq 1$. Thus it suffices to prove that $\sup_{r\geq 1} \|\Sigma_{\mathcal{Y}_r}\| < \infty$ in order for $\sup_{r\geq 1} \|\Sigma_{\mathcal{Y},\mathcal{Y}_r}\|$ to be finite.
For this, we recall that the spectral density $f_{\mathcal{Y}}$ associated to the process $\{\mathcal{Y}_n,\ n \in \mathbb{Z}\}$ defined by (A.44) is absolutely convergent for all $x \in \mathbb{R}$, thus verifies

$$\|f_{\mathcal{Y}}(x)\| \leq K_{max}, \quad x \in \mathbb{R}, \tag{A.69}$$

for some constant $K_{max} > 0$. Also, $\Sigma_{\mathcal{Y}_r}$ is symmetric matrix with real entries, so that the matrix norm entails that $\|\Sigma_{\mathcal{Y}_r}\|$ is its largest (real) eigenvalue, that we denote by $\rho_{max}(r)$. Hence, there exists an eigenvector $\delta^{(r)} = (\delta_1^{(r)'}, ..., \delta_r^{(r)'})' \in \mathbb{R}^{3r}$, where $\delta_i \in \mathbb{R}^3$, such that

$$\|\Sigma_{\mathcal{Y}_r}\| = \rho_{max}(r) = \delta^{(r)'} \Sigma_{\mathcal{Y}_r} \delta^{(r)} \quad \text{and} \quad \|\delta^{(r)}\| = 1. \tag{A.70}$$

Besides, one obtains from (A.44) and the inversion formula that $C(h) = \int_{-\pi}^{\pi} f_{\mathcal{Y}}(x) e^{-ihx} dx$ for all $h \in \mathbb{Z}$, so that we obtain from (A.70) that

$$\begin{aligned}\|\Sigma_{Y_r}\| &= \sum_{j,s=1}^{r} \delta_j^{(r)'} \left[\int_{-\pi}^{\pi} f_{\mathcal{Y}}(x) e^{i(s-j)x} dx\right] \delta_s^{(r)} \\ &= \int_{-\pi}^{\pi} \left(\sum_{s=1}^{r} \delta_s^{(r)} e^{i(s-1)x}\right)' f_{\mathcal{Y}}(x) \overline{\left(\sum_{s=1}^{r} \delta_s^{(r)} e^{i(s-1)x}\right)} dx. \end{aligned} \tag{A.71}$$

Since $f_{\mathcal{Y}}(x)$ is an hermitian matrix for all $x \in [0, 2\pi]$, one has $X' f_{\mathcal{Y}}(x) \bar{X} \leq \|f_{\mathcal{Y}}(x)\| X' \bar{X}$ for all column vector $X \in \mathbb{C}^{3r}$. Coupled with (A.69), (A.71) as well as some easy computation, this yields the bound

$$\|\Sigma_{\mathcal{Y}_r}\| \leq K_{max} \int_{-\pi}^{\pi} \left(\sum_{s=1}^{r} \delta_s^{(r)} e^{i(s-1)x}\right)' \overline{\left(\sum_{s=1}^{r} \delta_s^{(r)} e^{i(s-1)x}\right)} dx = K_{max} \|\delta^{(r)}\|^2 = K_{max},$$

a constant which is independent from $r \geq 1$, which proves $\sup_{r\geq 1} \|\Sigma_{\mathcal{Y}_r}\| < \infty$.
We now turn to $\sup_{r\geq 1} \|\Sigma_{\mathcal{Y}_r}^{-1}\|$. We start from the observation that $\|\Sigma_{\mathcal{Y}_r}^{-1}\|$ is the inverse of the smallest eigenvalue $\rho_{min}(r)$ (which is positive) of $\Sigma_{\mathcal{Y}_r}$. As for $\rho_{max}(r)$, there thus exists an eigenvector $\iota^{(r)} = (\iota_1^{(r)'}, ..., \iota_r^{(r)'})' \in \mathbb{R}^{3r}$, where $\iota_i \in \mathbb{R}^3$, such that

$$\|\Sigma_{\mathcal{Y}_r}^{-1}\|^{-1} = \rho_{min}(r) = \iota^{(r)'} \Sigma_{\mathcal{Y}_r} \iota^{(r)} \quad \text{and} \quad \|\iota^{(r)}\| = 1. \tag{A.72}$$

Now, (**A4**) and (A.45) imply that $f_{\mathcal{Y}}(x)$ is invertible for all $x \in [-\pi, \pi]$, so that $X' f_{\mathcal{Y}}(x) \bar{X} > 0$ for all $X \in S(\mathbb{C}^{3r}) := \{V \in \mathbb{C}^{3r} |\ \|V\|^2 = V'\bar{V} = 1\}$. Since $[-\pi, \pi] \times S(\mathbb{C}^{3r})$ is a compact set of $\mathbb{R} \times \mathbb{C}^{3r}$, we deduce that there exists some constant $K_{min} > 0$ such that

$$X' f_{\mathcal{Y}}(x) \bar{X} \geq K_{min}, \quad \forall (x, X) \in [-\pi, \pi] \times S(\mathbb{C}^{3r}),$$

from which an analysis similar to (A.71) yields from (A.72) that

$$\begin{aligned}
\|\Sigma_{\mathcal{Y}_r}^{-1}\|^{-1} &= \int_{-\pi}^{\pi} \left(\sum_{s=1}^{r} \iota_s^{(r)} e^{i(s-1)x} \right)' f_{\mathcal{Y}}(x) \overline{\left(\sum_{s=1}^{r} \iota_s^{(r)} e^{i(s-1)x} \right)} dx \\
&\geq K_{min} \int_{-\pi}^{\pi} \left(\sum_{s=1}^{r} \iota_s^{(r)} e^{i(s-1)x} \right)' \overline{\left(\sum_{s=1}^{r} \iota_s^{(r)} e^{i(s-1)x} \right)} dx = K_{min} \|\iota^{(r)}\|^2 = K_{min},
\end{aligned}$$

so that $\|\Sigma_{\mathcal{Y}_r}^{-1}\| \leq K_{min}^{-1}$, a constant independent from $r \geq 1$, proving that $\sup_{r \geq 1} \|\Sigma_{\mathcal{Y}_r}^{-1}\| < \infty$. □

Lemma Appendix A.4. For all $s \in \mathbb{N}$, the series $\sum_{h=-\infty}^{\infty} \|\mathrm{Cov}(\mathcal{Y}_0 \mathcal{Y}_s', \mathcal{Y}_h \mathcal{Y}_{h+s}')\|$ is convergent.

Proof. Let us put $w_h(s) := \mathrm{Cov}(Y_0 Y_{k_0} Y_s Y_{s+k_0}, Y_h Y_{h+k_0} Y_{s+h} Y_{h+s+k_0})$, which is one of the entries of $\mathrm{Cov}(\mathcal{Y}_0 \mathcal{Y}_s', \mathcal{Y}_h \mathcal{Y}_{h+s}')$. We proceed by establishing estimates similar to the ones in Lemmas Appendix A.1 and Appendix A.2. As in (A.5), we have

$$\begin{aligned}
w_h(s) \quad = \quad & \sum_{s'=0}^{k_0-1} \lambda_0^{s'} \left[\sum_{t=0}^{s-k_0-1} \lambda_0^t \mathrm{Cov}(Y_0 Y_{k_0} Y_s Y_{s+k_0}, Y_h Y_{h+k_0} \epsilon_{s+h} \epsilon_{s+h+k_0-s'}) \right. \\
& \left. + \lambda_0^{s-k_0} \mathrm{Cov}(Y_0 Y_{k_0} Y_s Y_{s+k_0}, Y_h Y_{h+k_0}^2 \epsilon_{s+h+k_0-s'}) \right] \\
& + \lambda_0^{k_0} \mathrm{Cov}(Y_0 Y_{k_0} Y_s Y_{s+k_0}, Y_h Y_{h+k_0} Y_{s+h}^2),
\end{aligned}$$

Similarly to the bound (A.9), and since ϵ has moments of order 8 thanks to Assumption $(\mathbf{A2})_\mathbf{1}$, both quantities $\sum_{t=0}^{s-k_0-1} |\lambda_0^t \mathrm{Cov}(Y_0 Y_{k_0} Y_s Y_{s+k_0}, Y_h Y_{h+k_0} \epsilon_{s+h} \epsilon_{s+h+k_0-s'})|$ and $|\lambda_0^{s-k_0} \mathrm{Cov}(Y_0 Y_{k_0} Y_s Y_{s+k_0}, Y_h Y_{h+k_0}^2 \epsilon_{s+h+k_0-s'})|$ are less than

$$K \left[\alpha_\epsilon \left(\left\lfloor \frac{h-s-k_0-1}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\left\lfloor \frac{h-s-k_0-1}{2} \right\rfloor} \right]$$

for $h \geq s + k_0 + 1$ and some constant $K > 0$. Also, estimations similar to (A.6) and (A.7) leading to (A.9) may be used to obtain that $|\mathrm{Cov}(Y_0 Y_{k_0} Y_s Y_{s+k_0}, Y_h Y_{h+k_0} Y_{s+h}^2)|$ has the same upper bound, so that $\sum_{h=0}^{\infty} |w_h(s)| < \infty$. We prove similarly that $\sum_{h=-\infty}^{0} |w_h(s)|$ and, more generally, that $\sum_{h=-\infty}^{\infty} \|\mathrm{Cov}(\mathcal{Y}_0 \mathcal{Y}_s', \mathcal{Y}_h \mathcal{Y}_{h+s}')\|$ is convergent. □

Lemma Appendix A.5. The r.v. $\sqrt{r}\|\hat{\Sigma}_{\mathcal{Y}_r} - \Sigma_{\mathcal{Y}_r}\|$, $\sqrt{r}\|\hat{\Sigma}_{\mathcal{Y}} - \Sigma_{\mathcal{Y}}\|$ and $\sqrt{r}\|\hat{\Sigma}_{\mathcal{Y},\mathcal{Y}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\|$ converge to 0 in probability as $n \to \infty$ when $r = r(n) = \mathrm{o}(n^{1/3})$.

Proof. Remember from (A.64) and (A.65) that $\hat{\Sigma}_{\mathcal{Y}_r} - \Sigma_{\mathcal{Y}_r}$ has $(3r)^2$ coefficients. Each entry of $\hat{\Sigma}_{\mathcal{Y}_r}$ is of the form

$$\hat{Z}_{r_1,r_2}^{i,j} := \frac{1}{n} \sum_{t=1}^{n} \mathcal{Y}_{t-r_1}^i \mathcal{Y}_{t-r_2}^j, \quad i,j = 1,2,3,\ 1 \leq r_1, r_2 \leq r,$$

of which variance verifies $\mathrm{Var}(\hat{Z}^{i,j}_{r_1,r_2}) = \frac{1}{n^2}\sum_{h=-n+1}^{n-1}(n-|h|)\mathrm{Cov}(\mathcal{Y}^i_{-r_1}\mathcal{Y}^j_{-r_2}, \mathcal{Y}^i_{h-r_1}\mathcal{Y}^j_{h-r_2})$ which is less than C_1/n for some constant C_1 thanks to Lemma Appendix A.4. We deduce that

$$\mathbb{E}\left(\left[\sqrt{r}\|\hat{\Sigma}_{\mathcal{Y}_r} - \Sigma_{\mathcal{Y}_r}\|\right]^2\right) = r\,\mathbb{E}\left(\|\hat{\Sigma}_{\mathcal{Y}_r} - \Sigma_{\mathcal{Y}_r}\|^2\right) \le r\sum_{i,j=1}^{3}\sum_{r_1,r_2=1}^{r}\mathrm{Var}(\hat{Z}^{i,j}_{r_1,r_2})$$
$$\le \frac{3r^3}{n}C_1 \longrightarrow 0, \quad n\to\infty,\ r = r(n) = \mathrm{o}(n^{1/3}),$$

which proves that $\sqrt{r}\|\hat{\Sigma}_{\mathcal{Y}_r} - \Sigma_{\mathcal{Y}_r}\|$ converges in $\mathbb{L}^2$, hence in probability, to 0. We prove the same convergence for $\sqrt{r}\|\hat{\Sigma}_{\mathcal{Y}} - \Sigma_{\mathcal{Y}}\|$ and $\sqrt{r}\|\hat{\Sigma}_{\mathcal{Y},\mathcal{Y}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\|$ in an analogous manner. □

Lemma Appendix A.6. The r.v. $\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y}_r}\|$, $\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}}} - \Sigma_{\mathcal{Y}}\|$ and $\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\|$ converge to 0 in probability as $n\to\infty$ when $r = r(n) = \mathrm{o}(n^{1/3})$.

Proof. We are going to prove that $\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}}_r} - \hat{\Sigma}_{\mathcal{Y}_r}\| \xrightarrow{\mathbb{P}} 0$ as $n\to\infty$ when $r = r(n) = \mathrm{o}(n^{1/3})$ so as to conclude the convergence of $\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y}_r}\|$ thanks to Lemma Appendix A.5 and the triangular inequality. We note that the entries of $\hat{\Sigma}_{\hat{\mathcal{Y}}_r} - \hat{\Sigma}_{\mathcal{Y}_r}$ are of the form

$$\frac{1}{n}\sum_{t=1}^{n}\left[\mathcal{Y}^i_{t-r_1}\mathcal{Y}^j_{t-r_2} - \hat{\mathcal{Y}}^i_{t-r_1}\hat{\mathcal{Y}}^j_{t-r_2}\right], \quad i,j = 1,2,3,\ 1\le r_1,r_2\le r,$$

of which most complex term is the following:

$$\begin{aligned}
&\frac{1}{n}\sum_{t=1}^{n}\{(Y_{t-r_1}Y_{t+k_0-r_1} - u_{k_0-1})(Y_{t-r_2}Y_{t+k_0-r_2} - u_{k_0-1})\\
&-(Y_{t-r_1}Y_{t+k_0-r_1} - \bar{Y}_{k_0,t-r_1})(Y_{t-r_2}Y_{t+k_0-r_2} - \bar{Y}_{k_0,t-r_2})\}\\
=\ &\frac{1}{n}\sum_{t=1}^{n}\{u^2_{k_0-1} - \bar{Y}_{k_0,t-r_1}\bar{Y}_{k_0,t-r_2} + Y_{t-r_1}Y_{t+k_0-r_1}(\bar{Y}_{k_0,t-r_2} - u_{k_0-1})\\
&+Y_{t-r_2}Y_{t+k_0-r_2}(\bar{Y}_{k_0,t-r_1} - u_{k_0-1})\}. \qquad (A.73)
\end{aligned}$$

We study each terms on the righthandside of the above equality and start with

$$\frac{1}{n}\sum_{t=1}^{n}Y_{t-r_1}Y_{t+k_0-r_1}(\bar{Y}_{k_0,t-r_2} - u_{k_0-1}). \qquad (A.74)$$

We first observe that a consequence of one of the steps in the proof of Theorem 22 (see convergence (A.39) in Step 1 of the proof of this theorem) is $\sqrt{n}\|\bar{Y}_{k_0,n} - u_{k_0-1}\|_2$ is upper bounded by some constant K for all $n\ge 1$, of which consequence is obviously that $\|\bar{Y}_{k_0,n} - u_{k_0-1}\|_2$ is obviously upper bounded in $n\in\mathbb{N}$. Hence, using the classical inequality $\sqrt{a+b}\le\sqrt{a}+\sqrt{b}$ for all $a,b\ge 0$, we have, for all $r_2\le r$,

$$\begin{aligned}
\sqrt{t}\|\bar{Y}_{k_0,t-r_2} - u_{k_0-1}\|_2 &\le K\sqrt{r}, \quad \forall t = 1,\dots r_2,\\
\sqrt{t}\|\bar{Y}_{k_0,t-r_2} - u_{k_0-1}\|_2 &\le \sqrt{t-r_2}\|\bar{Y}_{k_0,t-r_2} - u_{k_0-1}\|_2 + \sqrt{r_2}\|\bar{Y}_{k_0,t-r_2} - u_{k_0-1}\|_2\\
&\le K(1+\sqrt{r}), \quad \forall t\ge r_2,
\end{aligned}$$

from which one has the global bound $\sqrt{t}\|\bar{Y}_{k_0,t-r_2} - u_{k_0-1}\|_2 \leq K(1+\sqrt{r})$ for some constant K and for all $t \geq 0$. Coupled with the fact that $\|Y_n\|_4$ is uniformly bounded by some constant C as well as Cauchy Scwharz inequality, we thus deduce that (A.74) is bounded as follows

$$\left\| \frac{1}{n}\sum_{t=1}^{n} Y_{t-r_1} Y_{t+k_0-r_1}(\bar{Y}_{k_0,t-r_2} - u_{k_0-1}) \right\|_1 \leq \frac{1}{n}\sum_{t=1}^{n} C^2 \frac{K(1+\sqrt{r})}{\sqrt{t}} \leq K' \frac{1+\sqrt{r}}{\sqrt{n}}$$

for some contant K', where we used that $\sum_{t=1}^{n} \frac{1}{\sqrt{t}}$ is equivalent to $\sqrt{n}$ up to a factor. We thus deduce that

$$\sqrt{r} \left\| \frac{1}{n}\sum_{t=1}^{n} Y_{t-r_1} Y_{t+k_0-r_1}(\bar{Y}_{k_0,t-r_2} - u_{k_0-1}) \right\|_1 \longrightarrow 0, \quad n \to \infty,\ r = r(n) = \mathrm{o}(n^{1/3}). \quad \text{(A.75)}$$

A similar approach yields the following limit that concerns the second term in (A.73):

$$\sqrt{r} \left\| \frac{1}{n}\sum_{t=1}^{n} Y_{t-r_2} Y_{t+k_0-r_2}(\bar{Y}_{k_0,t-r_1} - u_{k_0-1}) \right\|_1 \longrightarrow 0, \quad n \to \infty,\ r = r(n) = \mathrm{o}(n^{1/3}). \quad \text{(A.76)}$$

We now turn to the last first term in (A.73) namely $n^{-1}\sum_{t=1}^{n}(u_{k_0-1}^2 - \bar{Y}_{k_0,t-r_1}\bar{Y}_{k_0,t-r_2})$. For this, we write $u_{k_0-1}^2 - \bar{Y}_{k_0,t-r_1}\bar{Y}_{k_0,t-r_2} = u_{k_0-1}(u_{k_0-1} - \bar{Y}_{k_0,t-r_1}) + \bar{Y}_{k_0,t-r_1}(u_{k_0-1} - \bar{Y}_{k_0,t-r_1})$ and conclude in a similar manner that

$$\sqrt{r} \left\| \frac{1}{n}\sum_{t=1}^{n}(u_{k_0-1}^2 - \bar{Y}_{k_0,t-r_1}\bar{Y}_{k_0,t-r_2}) \right\|_1 \longrightarrow 0, \quad n \to \infty,\ r = r(n) = \mathrm{o}(n^{1/3}). \quad \text{(A.77)}$$

Gathering (A.75), (A.76) and (A.77) we deduce that (A.73) multiplied by $\sqrt{r}$ tends to 0 in $\mathbb{L}^1$ hence in probability as $n \to \infty$ when $r = r(n) = \mathrm{o}(n^{1/3})$. As announced earlier, we prove similarly that all other entries of $\sqrt{r}[\hat{\Sigma}_{\hat{\mathcal{Y}}_r} - \hat{\Sigma}_{\mathcal{Y}_r}]$ also converges in probability. This concludes the proof. □

Lemma Appendix A.7. Let us define $\Phi_r^* := (\Phi_1, ..., \Phi_r)'$. Then $\sqrt{r}\|\Phi_r^* - \underline{\Phi_r}\| \longrightarrow 0$ as $r \to \infty$, where we recall that $\underline{\Phi_r}$ is defined in (A.67).

Proof. We follow the proof of Boubacar Mainassara et al. (2012, Lemma 5). We recall from (30) and the definition of the least square regression coefficients $\Phi_{r,1}, \ldots, \Phi_{r,r}$ with corresponding residual $\varepsilon_{n,r}$ that $\mathcal{Y}_n = \sum_{k=1}^{r} \Phi_k \mathcal{Y}_{n-k} + \sum_{k=r+1}^{\infty} \Phi_k \mathcal{Y}_{n-k} + \varepsilon_n = \sum_{k=1}^{r} \Phi_{r,k} \mathcal{Y}_{n-k} + \varepsilon_{n,r}$ which entails the matrix relation

$$(\underline{\Phi_r} - \Phi_r^*)\mathcal{Y}_{r,n} = \sum_{k=r+1}^{\infty} \Phi_k \mathcal{Y}_{n-k} + \varepsilon_n := \varepsilon_{n,r}^*, \quad \text{(A.78)}$$

from which we deduce that

$$\underline{\Phi_r} - \Phi_r^* = \Sigma_{\varepsilon_r^*}\Sigma_{\mathcal{Y}_r}^{-1}, \quad \text{where } \Sigma_{\varepsilon_r^*} := \mathbb{E}(\varepsilon_{n,r}^* \mathcal{Y}_{r,n}'). \quad \text{(A.79)}$$

Let us note that ε_n is orthogonal to $\mathcal{Y}_{r,n}$ i.e. $\mathbb{E}(\varepsilon_n \mathcal{Y}_{r,n}') = 0$, as this vector depends on $\mathcal{Y}_t$ for $t = n-r, ..., n-1$. So that we obtain from (A.78) and (A.79) that

$$\Sigma_{\varepsilon_r^*} = \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}(\mathcal{Y}_{n-k} \mathcal{Y}_{r,n}'). \quad \text{(A.80)}$$

We proved in Step 1 of the proof of Theorem (22) that $\mathfrak{S}_{k_0}$ in (23) is an absolute convergent series: a byproduct of this is that $\|\mathcal{Y}_0\|_2$ is finite. Also, since $\mathcal{Y}_{r,n}$ is a $3r \times 3$ matrix, the matrix norm property (A.68) implies that $\|\mathcal{Y}_{r,n}\|_2 \leq \sqrt{Kr}$ for some constant $K > 0$. Thus, we deduce from the Cauchy-Schwarz inequality that $\|\mathbb{E}(\mathcal{Y}_{n-k}\mathcal{Y}'_{r,n})\| \leq \mathbb{E}(\|\mathcal{Y}_{n-k}\mathcal{Y}'_{r,n}\|) \leq \mathbb{E}(|\mathcal{Y}_{n-k}\| \, \|\mathcal{Y}'_{r,n}\|) \leq |\mathcal{Y}_{n-k}\|_2 \, \|\mathcal{Y}'_{r,n}\|_2 = \mathrm{O}(\sqrt{r})$. So that, by (A.80):

$$\|\Sigma_{\varepsilon_r^*}\| \leq \sum_{k=r+1}^{\infty} \|\Phi_k\| \, \|\mathbb{E}(\mathcal{Y}_{n-k}\mathcal{Y}'_{r,n})\| = \mathrm{O}(\sqrt{r}) \sum_{k=r+1}^{\infty} \|\Phi_k\|,$$

hence we obtain from (A.79) that $\sqrt{r}\|\Phi_r^* - \underline{\Phi_r}\| = \mathrm{O}(r)\sum_{k=r+1}^{\infty}\|\Phi_k\| \longrightarrow 0$ thanks to the assumption $\|\Phi_k\| = \mathrm{o}(1/k^2)$. □

Lemma Appendix A.8. The r.v. $\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y}_r}^{-1}\|$ converge to 0 in probability as $n \to \infty$ when $r = r(n) = \mathrm{o}(n^{1/3})$.

Proof. The proof is exactly the same as the one in Boubacar Mainassara et al. (2012, Lemma 6), up to a change of notation. □

Lemma Appendix A.9. The r.v. $\sqrt{r}\|\underline{\hat{\Phi}_r} - \underline{\Phi_r}\|$ converge to 0 in probability as $n \to \infty$ when $r = r(n) = \mathrm{o}(n^{1/3})$.

Proof. We follow the proof of Boubacar Mainassara et al. (2012, Lemma 7). Similarly to (A.79), and thanks to (A.67), one has that

$$\begin{aligned}\underline{\hat{\Phi}_r} - \underline{\Phi_r} = \hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r}\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\Sigma_{\mathcal{Y}_r}^{-1} &= (\hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r})(\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y}_r}^{-1}) \\ &\quad + (\hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r})\Sigma_{\mathcal{Y}_r}^{-1} + \Sigma_{\mathcal{Y},\mathcal{Y}_r}(\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y}_r}^{-1}),\end{aligned}$$

whence one obtains for all $a > 0$ that

$$\begin{aligned}\mathbb{P}(\sqrt{r}\|\underline{\hat{\Phi}_r} - \underline{\Phi_r}\| > a) &\leq \mathbb{P}\Big(\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\| \, \|(\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y}_r}^{-1})\| \\ &\quad + \sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\| \, \|\Sigma_{\mathcal{Y}_r}^{-1}\| + \sqrt{r}\|\Sigma_{\mathcal{Y},\mathcal{Y}_r}\| \, \|\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y}_r}^{-1}\| > a\Big) \\ &\leq \mathbb{P}\Big(\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\| \, \|(\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y}_r}^{-1})\| > a/3\Big) + \mathbb{P}\left(\sqrt{r}\|\hat{\Sigma}_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma_{\mathcal{Y},\mathcal{Y}_r}\| > a/(3\sup_{j\geq 1}\|\Sigma_{\mathcal{Y}_j}^{-1}\|)\right) \\ &\quad + \mathbb{P}\left(\sqrt{r}\|\Sigma_{\mathcal{Y},\mathcal{Y}_r}\|\|\hat{\Sigma}_{\hat{\mathcal{Y}}_r}^{-1} - \Sigma_{\mathcal{Y}_r}^{-1}\| > a/(3\sup_{j\geq 1}\|\Sigma_{\mathcal{Y},\mathcal{Y}_j}\|)\right). \quad \text{(A.81)}\end{aligned}$$

Now, each of the terms in (A.81) tends to 0 thanks to Lemmas Appendix A.6 and Appendix A.3, and the proof is complete. □

The end of proof of Theorem 5.2 relies on the above Lemmas Appendix A.3 to Appendix A.8, and is the same as the end of proof as Boubacar Mainassara et al. (2012, Theorem A.1, page 7 of Supplementary material). We give it below for the sake of presentation. Let $\mathbb{E}_r := \mathbf{I}_3 \otimes \mathbf{1}_r$ where we recall that $\mathbf{I}_3$ is the 3×3 identity matrix, $\mathbf{1}_r = (1, ..., 1)'$ of size r. Then by the submultiplicativity of the matrix norm and Lemma Appendix A.8: $\|\sum_{i=1}^{r}\left(\hat{\Phi}_{r,i} - \Phi_{r,i}\right)\| = \|(\hat{\Phi}_r - \Phi_r)\mathbb{E}_r\| \leq \sqrt{3r}\|\hat{\Phi}_r - \Phi_r\| \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$ when $r = r(n) =$

$o(n^{1/3})$. Similarly using Lemma Appendix A.7, we have that $\| \sum_{i=1}^{r} (\Phi_{r,i} - \Phi_i) \|\| \longrightarrow 0$. We then deduce by triangular inequality that

$$\|\hat{\Phi}_r(1) - \Phi(1)\| \leq \left\| \sum_{i=1}^{r} \left(\hat{\Phi}_{r,i} - \Phi_{r,i} \right) \right\| + \left\| \sum_{i=1}^{r} (\Phi_{r,i} - \Phi_i) \right\| + \left\| \sum_{i=r+1}^{\infty} \Phi_i \right\| \xrightarrow{\mathbb{P}} 0, \quad n \to \infty, \ r = r(n) = o(n^{1/3}). \quad \text{(A.82)}$$

We now pass on to the convergence of the estimator $\hat{\Sigma}_{\hat{\varepsilon}_r}$. (33) implies that $\hat{\Sigma}_{\hat{\varepsilon}_r} = \hat{\Sigma}_{\hat{\mathcal{Y}}} - \underline{\hat{\Phi}_r} \hat{\Sigma}'_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r}$. Also, the expression (28), the orthogonality of ε_0 and $P_{-1}\mathcal{Y}_0$ and (30) imply that

$$\begin{aligned} \Sigma_\varepsilon = \mathbb{E}(\varepsilon_0 \varepsilon_0') = \mathbb{E}(\varepsilon_0 \mathcal{Y}_0') &= \mathbb{E}\left[\left(\mathcal{Y}_0 - \sum_{k=1}^{\infty} \Phi_k \mathcal{Y}_{-k} \right) \mathcal{Y}_0' \right] \\ &= \Sigma_{\mathcal{Y}} - \sum_{k=1}^{\infty} \Phi_k \mathbb{E}(\mathcal{Y}_{-k} \mathcal{Y}_0') = \Sigma_{\mathcal{Y}} - \Phi_r^* \Sigma'_{\mathcal{Y},\mathcal{Y}_r} - \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}(\mathcal{Y}_{-k} \mathcal{Y}_0'). \end{aligned}$$

We deduce that

$$\begin{aligned} \|\hat{\Sigma}_{\hat{\varepsilon}_r} - \Sigma_{\varepsilon_r}\| &= \Big\| \hat{\Sigma}_{\hat{\mathcal{Y}}} - \Sigma_{\mathcal{Y}} - \left(\underline{\hat{\Phi}_r} - \Phi_r^* \right) \hat{\Sigma}'_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} \\ &\quad - \Phi_r^* \left(\hat{\Sigma}'_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma'_{\mathcal{Y},\mathcal{Y}_r} \right) + \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}(\mathcal{Y}_{-k} \mathcal{Y}_0') \Big\| \\ &\leq \|\hat{\Sigma}_{\hat{\mathcal{Y}}} - \Sigma_{\mathcal{Y}}\| + \left\| \left(\underline{\hat{\Phi}_r} - \Phi_r^* \right) \left(\hat{\Sigma}'_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma'_{\mathcal{Y},\mathcal{Y}_r} \right) \right\| + \left\| \left(\underline{\hat{\Phi}_r} - \Phi_r^* \right) \Sigma'_{\mathcal{Y},\mathcal{Y}_r} \right\| \\ &\quad + \left\| \Phi_r^* \left(\hat{\Sigma}'_{\hat{\mathcal{Y}},\hat{\mathcal{Y}}_r} - \Sigma'_{\mathcal{Y},\mathcal{Y}_r} \right) \right\| + \left\| \sum_{k=r+1}^{\infty} \Phi_k \mathbb{E}(\mathcal{Y}_{-k} \mathcal{Y}_0') \right\|. \end{aligned} \quad \text{(A.83)}$$

Lemma Appendix A.5 entails that the first term on the right hand side of (A.83) converges in probability to 0. The combination of Lemmas Appendix A.7, Appendix A.9 and Appendix A.6 yield that the second term converges to 0, and likewise the fourth term tends to 0 because the $\{\Phi_i, \ i \in \mathbb{N}\}$ is a bounded sequence thanks for example to the assumption $\|\Phi_k\| = o(1/k^2)$. Lemma Appendix A.3 also entails that the third term tends to 0. Finally, the last term tends to 0 since it is the remainder of a convergent series. All in all, we thus deduce from (A.83) that $\hat{\Sigma}_{\hat{\varepsilon}_r} \xrightarrow{\mathbb{P}} 0$ as $n \to \infty$, $r = r(n) = o(n^{1/3})$, which combined with (A.82), proves (35). □

Appendix A.7. Proof of Proposition 6.3

The following lemma will be needed.

Lemma Appendix A.10. There exist constants M_ϖ, M_H^1 and M_H^2 such that

$$|\varpi_k'(|x|) \ln(|x|)| \leq M_\varpi \frac{k}{\lambda_-^k}, \quad \text{(A.84)}$$

$$|H_k'(x)| \leq M_H^1 \frac{1}{\lambda_-^k}, \quad \text{(A.85)}$$

$$|H_k''(x)| \quad \leq \quad M_H^2 \frac{1}{\lambda_-^k}, \tag{A.86}$$

for all $x \in \mathbb{R}$. Besides, for each $k \in \mathbb{N}$ there exists $M_1(k)$ such that

$$||\nabla \psi_k(a,b)|| \leq M_1(k), \quad \forall (a,b) \in \mathbb{R}^2, \tag{A.87}$$
$$||D^2 \psi_k(a,b)|| \leq M_2(k), \quad \forall (a,b) \in \mathbb{R}^2, \tag{A.88}$$

with $M_1(k)$ and $M_2(k)$ being $\mathrm{O}\left(\frac{1}{\lambda_-^k}\right)$.

Proof. The assumption (41) on the regularity at 0 of the function ϖ imply that $\varpi'(x) = \mathrm{O}(x)$ as $x \to^+ 0$, hence the existence of some positive ϑ^1 (that we may also suppose less than 1 w.l.o.g.) and C^1_ϖ such that $0 \leq \varpi'(x) \leq C^1_\varpi x$ when $x \in [0, \vartheta^1]$. Furthermore, the regularity conditions implies the existence of some $C_\varpi > 0$ such that $0 \leq \varpi'(x) \leq C_\varpi$, $x \in [0,1]$; we also recall that $\varpi'(x) = 0$ if $x < 0$ or $x > 1$. Consequently, ϖ_k defined in (42) verifies

$$|\varpi_k'(|x|)\ln(|x|)| \leq |\ln(|x|)| C^1_\varpi \frac{2}{K_m \lambda_-^k} x \quad \text{if } \frac{2}{K_m\lambda_-^k}|x| \leq \vartheta^1 \iff |x| \leq \frac{K_m \lambda_-^k}{2}\vartheta^1. \tag{A.89}$$

We now observe that the function $x \mapsto x \ln x$ is decreasing and negative on the interval $[0, e^{-1}]$, and in particular on the interval $\left[0, K_m\lambda_-^k \vartheta^1/2\right]$ for k large enough. We thus deduce from (A.89) that $|\varpi_k'(|x|)\ln(|x|)| \leq C^1_\chi \left|\ln\left(K_m\lambda_-^k\vartheta^1/2\right)\right| = \mathrm{O}(k)$. Since $\lambda_- < 1$, an $\mathrm{O}(k)$ is also an $\mathrm{O}(k/\lambda_-^k)$ so that we obtain

$$|\varpi_k'(|x|)\ln(|x|)| = \mathrm{O}(k/\lambda_-^k), \quad x \in \left[0, \frac{K_m\lambda_-^k}{2}\vartheta^1\right]. \tag{A.90}$$

Now, since $x \mapsto \ln x$ is increasing and negative on $(0,1]$, and picking k large enough such that $(K_m\lambda_-^k)/2 < 1$, we have

$$\begin{aligned}|\varpi_k'(|x|)\ln(|x|)| \leq C^1_\varpi |\ln(|x|)| \leq C^1_\varpi \left|\ln\left(\frac{K_m\lambda_-^k}{2}\vartheta^1\right)\right| \\ = \mathrm{O}(k/\lambda_-^k), \quad x \in \left[\frac{K_m\lambda_-^k}{2}\vartheta^1, \frac{K_m\lambda_-^k}{2}\right].\end{aligned} \tag{A.91}$$

And, since $\varpi_k'(|x|)\ln(|x|) = 0$ when $|x| > (K_m\lambda_-^k)/2$, (A.90) and (A.91) yield (A.84).
We now compute

$$H_k'(x) = \frac{1}{k}\frac{1}{|x|}\varpi_k(|x|) + \frac{1}{k}\varpi_k'(|x|)\ln|x|, \quad x \in \mathbb{R}, \tag{A.92}$$

observing incidentally that $H_k'(x)$ is continuous at $x = 0$ thanks to the fact that $\varpi(x) = \mathrm{o}(x^2)$ and $\varpi'(x) = \mathrm{o}(x)$ when $x \to 0$. We first note that (A.84) implies that the second term on the righthandside of (A.92) verifies $\frac{1}{k}\varpi_k'(x)\ln|x| = \mathrm{O}(1/\lambda_-^k)$. Furthermore, ϖ has the property that $\varpi(x) = \mathrm{O}(x)$ so that there exists $\vartheta^0 > 0$ and C^0_ϖ such that $0 \leq \varpi(x) \leq C^0_\varpi x$ when $x \in [0, \vartheta^0]$. Hence the following upper bound

$$\left|\frac{1}{k}\frac{1}{|x|}\varpi_k(|x|)\right| \leq \frac{1}{k}\frac{2}{K_m\lambda_-^k}C^0_\chi = \mathrm{O}(1/\lambda_-^k) \quad \text{if } \left|\frac{2}{K_m\lambda_-^k}x\right| \leq \vartheta^0 \iff |x| \leq \frac{K_m\lambda_-^k}{2}\vartheta^0.$$

If $|x| \geq (K_m \lambda_-^k \vartheta^0)/2$ then $|\varpi_k(|x|)/(k|x|)| \leq |1/(k|x|)|\, C_\varpi \leq \left|2/(kK_m\lambda_-^k\vartheta^0)\right| C_\varpi = \mathrm{O}(1/\lambda_-^k)$ which, coupled with the above inequality, yields that the first term on the righthandside of (A.92) is an $\mathrm{O}(1/\lambda_-^k)$. All in all, we thus obtain the inequality (A.85). With similar arguments, (A.86) is obtained by differentiating $H_k'(x)$ and using $\varpi(x) = \mathrm{o}(x^3)$.
We now prove (A.87) from the two afore established upper bounds (A.84) and (A.85). We observe that ψ is twice differentiable, and compute its first derivative w.r.t. a as

$$\partial_a \psi_k(a,b) = G'(a)G(b)H_k(a^2-b) + 2aG(a)G(b)H_k'(a^2-b). \tag{A.93}$$

Since $\varpi(x) = \mathrm{O}(x)$ as $x \to 0$, a similar approach leading to (A.90) yields that $|\varpi_k(|x|)\ln(|x|)| = \mathrm{O}(k/\lambda_-^k)$ on $|x| \leq (K_m\lambda_-^k\vartheta^0)/2$. Consequently we have that

$$|G'(a)G(b)H_k(a^2-b)| = \mathrm{O}(1/\lambda_-^k), \quad |a^2-b| \in \left[0, \frac{K_m\lambda_-^k}{2}\vartheta^0\right]. \tag{A.94}$$

Since G has a finite support and is thus zero outside of some compact interval say I, $(a,b) \mapsto G'(a)G(b)H_k(a^2-b)$ is 0 when $(a,b) \notin I^2$, and is continuous hence uniformly bounded on the compact set $I^2 \cap \left\{(a,b) \in \mathbb{R}^2 |\ |a^2-b| > (K_m\lambda_-^k\vartheta^0)/2\right\}$. Thanks to the factor $1/k$ in the definition (43) of H_k, it is easily verified that this uniform bound on that set is in fact an $\mathrm{O}(1/k)$, and so is also an $\mathrm{O}(1/\lambda_-^k)$. This latter fact, coupled with (A.94), yields that the first term on the righthandside of (A.93) is an $\mathrm{O}(1/\lambda_-^k)$ uniformly in $(a,b) \in \mathbb{R}^2$. Similarly, a similar analysis along with the inequality (A.85) proved previously yields that second term on the righthandside of (A.93) is an $\mathrm{O}(1/\lambda_-^k)$ uniformly in $(a,b) \in \mathbb{R}^2$. Again, a similar analysis yields similar upper bounds for $\partial_b\psi_k(a,b)$, so that (A.87) holds. The upper bound (A.88) is obtained by similar arguments thanks to (A.86). This concludes the proof of Lemma Appendix A.10 □

We now turn to the proof of Proposition 3.2. Let us set $U(k) := (C_1, u_k)$, $k \in \mathbb{N}$. The definition (48) as well as the finite increment theorem yields the existence of some (random) $c_{k,n} \in (0,1)$ such that

$$\begin{aligned}\psi_k(\bar{Y}_n, \bar{Y}_{n,k}) - \psi_k(U(k)) &= \psi_k\left(U(k) + \frac{S_n(k)}{n}\right) - \psi_k(U(k)) \\ &= \nabla\psi_k\left(U(k) + c_{k,n}\frac{S_n(k)}{n}\right).\frac{S_n(k)}{n},\end{aligned}$$

so that, thanks to (A.87), we obtain the L^2 bound

$$||\psi_k(\bar{Y}_n, \bar{Y}_{n,k}) - \psi_k(U(k))||_2 \leq M_1(k)\left|\left|\frac{S_n(k)}{n}\right|\right|_2. \tag{A.95}$$

We now observe that, for k large enough we have from (13) as well as Assumption (**A6**) that $|C_1^2 - u_k| = \lambda_0^k\left|\sum_{j=0}^k \lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right| \geq \lambda_0^k K_m/2$, so that the definition of χ_k entails that $\varpi_k(|C_1^2 - u_k|) = 1$. Now, by Assumption (**A7**) we have $0 \leq C_1 = ||Y_0||_1 \leq C_Y$ and $0 \leq u_k = \mathbb{E}(Y_0Y_{k+1}) \leq ||Y_0||_2||Y_{k+1}||_2 = ||Y_0||_2^2 \leq C_Y^2$ (by the Cauchy Scwharz inequality) so that the definition of function G entails that $G(C_1) = G(u_k) = 1$. Thus $\psi_k(U(k)) =$

$G(C_1)G(u_k)H_k(C_1^2-u_k)=\frac{1}{k}\ln|C_1^2-u_k|=S_k$ for k large enough. All in all, we thus have, from (45) and (13), that

$$|\psi_k(U(k))-\ln\lambda_0|=\frac{1}{k}\left|\sum_{j=0}^{k}\lambda_0^{-j}\chi_j+\lambda_0(C_1^2-C_2)\right|=\mathrm{O}\left(\frac{1}{k}\right)$$

for k large enough. The upper bound (56) is deduced from the above inequality, Minkowski's inequality, (A.95) as well as the estimate $M_1(k)=\mathrm{O}\left(1/\lambda_-^k\right)$ proved in Lemma Appendix A.10. □

Appendix A.8. Proof of Proposition 6.4

The scheme somewhat has some common ideas with the proof of Theorem 4.2 in Section Appendix A.5 (with the notable difference that k_0 in this section is now replaced with k_n with $\lim_{n\to\infty}k_n=\infty$) combined with a triangular central limit theorem for stationary sequences of random variables proved in Francq and Zakoïan (2005), and is decomposed in several steps. Let $(k_n)_{n\in\mathbb{N}}$ be a sequence of integers verifying (58).

Appendix A.8.0.1. ⋄ *Step 1:.* we first prove the existence of the limit $\mathfrak{S}$ in (57). We first observe that

$$\frac{1}{n}\mathbb{E}(S_n(k_n)S_n(k_n)')=\frac{1}{n}\sum_{h=-n}^{n}(n-|h|)\mathbb{E}\left(X_0(k_n)X_h(k_n)'\right) \tag{A.96}$$

where we recall that $S_n(.)$ and $X_h(.)$ are defined in (20). Since

$$\mathbb{E}\left(X_0(k_n)X_h(k_n)'\right)=\begin{bmatrix}\mathrm{Cov}(Y_0,Y_h) & \mathrm{Cov}(Y_0,Y_hY_{h+k_n+1})\\ \mathrm{Cov}(Y_0Y_{k_n+1},Y_h) & \mathrm{Cov}(Y_0Y_{k_n+1},Y_hY_{h+k_n+1})\end{bmatrix}, \tag{A.97}$$

we then see that (A.96) converges if we prove that $\sum_{h=1}^{n}\mathbb{E}\left(X_0(k_n)X_h(k_n)'\right)$ and $\sum_{h=1}^{n}h\,\mathbb{E}\left(X_0(k_n)X_h(k_n)'\right)$ converge by the dominant convergence theorem, in which case the limit of (A.96) is related to the limit of $\sum_{h=1}^{n}\mathbb{E}\left(X_0(k_n)X_h(k_n)'\right)$ as $n\to\infty$. For this we study the corresponding terms in (A.97). The simplest term in the latter expression is $\mathrm{Cov}(Y_0,Y_h)$, of which corresponding series converges thanks to a direct application of the bound (A.9) that yields $|\mathrm{Cov}(Y_0,Y_h)|\leq K\left[\alpha_\epsilon\left(\lfloor(h-1)/2\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{h-1}{2}\rfloor}\right]$ for some constant $K>0$. In that case the corresponding limit may also be written as

$$\frac{1}{n}\sum_{h=-n}^{n}(n-|h|)\mathrm{Cov}(Y_0,Y_h)\longrightarrow 2\sum_{h=1}^{\infty}(C_1^2-u_{h-1}),\quad n\to\infty. \tag{A.98}$$

We next study $\mathrm{Cov}(Y_0,Y_hY_{h+k_n+1})$. We get from (A.5) that

$$\begin{aligned}\mathrm{Cov}(Y_0,Y_hY_{h+k_n+1}) &= \sum_{s=0}^{k_n}\lambda_0^s\left[\sum_{t=0}^{h-1}\lambda_0^t\mathrm{Cov}(Y_0,\epsilon_{h-t}\epsilon_{k_n+1+h-s})+\lambda_0^h\mathrm{Cov}(Y_0,Y_0\epsilon_{k_n+1+h-s})\right]\\ &\quad+\lambda_0^{k_n+1}\mathrm{Cov}(Y_0,Y_h^2).\end{aligned} \tag{A.99}$$

We wish to apply the dominant convergence theorem to establish the limit of $\sum_{h=1}^{n} \mathrm{Cov}(Y_0, Y_h Y_{h+k_n+1})$ as $n \to \infty$. For this we observe that, similarly to (A.14):

$$|\mathrm{Cov}(Y_0, Y_h Y_{h+k_n+1})| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{h-1}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{h-1}{2} \rfloor} + \lambda_0^h \right]$$

for some constant $K > 0$, of which righthandside upper bound is independent from n and summable in $h \geq 1$. Thus, we now need to study the convergence of the terms on the righthandside of (A.99) as $n \to \infty$ for fixed $h \geq 1$. We start with $\sum_{s=0}^{k_n} \lambda_0^s \sum_{t=0}^{h-1} \lambda_0^t \mathrm{Cov}(Y_0, \epsilon_{h-t}\epsilon_{k_n+1+h-s})$. We write that

$$\begin{aligned} \mathrm{Cov}(Y_0, \epsilon_{h-t}\epsilon_{k_n+1+h-s}) &= \mathrm{Cov}(Y_0\epsilon_{h-t}, \epsilon_{k_n+1+h-s}) + \mathbb{E}(Y_0\epsilon_{h-t})\mathbb{E}(\epsilon_{k_n+1+h-s}) \\ &\quad -\mathbb{E}(Y_0)\mathbb{E}(\epsilon_{h-t}\epsilon_{k_n+1+h-s}) \\ &= \mathrm{Cov}(Y_0\epsilon_{h-t}, \epsilon_{k_n+1+h-s}) + \mathbb{E}(Y_0\epsilon_{h-t})m_0 - C_1\mathbb{E}(\epsilon_{h-t}\epsilon_{k_n+1+h-s}), \end{aligned}$$

of which terms are easier to analyse than $\mathrm{Cov}(Y_0, \epsilon_{h-t}\epsilon_{k_n+1+h-s})$ directly. For fixed $h, s, t \geq 1$, the Davydov inequality entails that $\lambda_0^s\lambda_0^t \mathrm{Cov}(Y_0\epsilon_{h-t}, \epsilon_{k_n+h-s}) \longrightarrow 0$ as $n \to \infty$. Furthermore, the Cauchy Schwarz inequality yields

$$|\lambda_0^s \mathrm{Cov}(Y_0\epsilon_{h-t}, \epsilon_{k_n+1+h-s})| \leq \lambda_0^s ||Y_0||_4 ||\epsilon||_4 ||\epsilon||_2$$

of which upper bound is summable w.r.t. $s \in \mathbb{N}$, so that a domination convergence theorem argument yields that

$$\sum_{s=0}^{k_n} \lambda_0^s \sum_{t=0}^{h-1} \lambda_0^t \mathrm{Cov}(Y_0\epsilon_{h-t}, \epsilon_{k_n+1+h-s}) \longrightarrow 0, \quad n \to \infty,$$

for fixed $h \geq 1$. Moreover, the following limits are easily obtained as $n \to \infty$:

$$\begin{aligned} \sum_{s=0}^{k_n} \lambda_0^s \sum_{t=0}^{h-1} \lambda_0^t \mathbb{E}(Y_0\epsilon_{h-t}) m_0 &\longrightarrow \sum_{t=0}^{h-1} \lambda_0^t \mathbb{E}(Y_0\epsilon_{h-t}) \frac{m_0}{1-\lambda_0} = \sum_{t=0}^{h-1} \lambda_0^t v_{h-t-1} \frac{m_0}{1-\lambda_0}, \\ \sum_{s=0}^{k_n} \lambda_0^s \sum_{t=0}^{h-1} \lambda_0^t C_1 \mathbb{E}(\epsilon_{h-t}\epsilon_{k_n+1+h-s}) &\longrightarrow \sum_{t=0}^{h-1} \lambda_0^t C_1 m_0^2 \frac{1}{1-\lambda_0} \end{aligned}$$

so that the following limit is established:

$$\begin{aligned} \sum_{s=0}^{k_n} \lambda_0^s \sum_{t=0}^{h-1} \lambda_0^t \mathrm{Cov}(Y_0, \epsilon_{h-t}\epsilon_{k_n+1+h-s}) &\longrightarrow \frac{m_0}{1-\lambda_0} \sum_{t=0}^{h-1} \lambda_0^t \left[v_{h-t-1} - C_1 m_0 \right] \\ &= -\frac{m_0}{1-\lambda_0} \sum_{t=0}^{h-1} \lambda_0^t \chi_{h-t-1}, \quad n \to \infty, \quad \text{(A.100)} \end{aligned}$$

thanks to Relations (11) and (7). A similar analysis yields

$$\sum_{s=0}^{k_n} \lambda_0^s \lambda_0^h \mathrm{Cov}(Y_0, Y_0\epsilon_{k_n+1+h-s}) = \sum_{s=0}^{k_n} \lambda_0^s \lambda_0^h \left[\mathrm{Cov}(Y_0^2, \epsilon_{k_n+1+h-s}) + C_2 m_0 - C_1 \mathbb{E}(Y_0\epsilon_{k_n+1+h-s}) \right]$$

$$\longrightarrow \sum_{s=0}^{\infty}\lambda_0^s\lambda_0^h\times 0+\sum_{s=0}^{\infty}\lambda_0^s\lambda_0^h C_2 m_0-\sum_{s=0}^{\infty}\lambda_0^s\lambda_0^h C_1^2 m_0$$

$$= \lambda_0^h m_0\frac{C_2-C_1^2}{1-\lambda_0} \tag{A.101}$$

$$\lambda_0^{k_n+1}\mathrm{Cov}(Y_0,Y_h^2) \longrightarrow 0 \tag{A.102}$$

as $n\to\infty$, for fixed $h\geq 1$. Combining (A.100), (A.101), (A.102) and (A.99), we then obtain

$$\sum_{h=1}^{n}\mathrm{Cov}(Y_0,Y_hY_{h+k_n+1}) \longrightarrow -\frac{m_0}{1-\lambda_0}\sum_{h=1}^{\infty}\sum_{t=0}^{h-1}\lambda_0^t\chi_{h-t-1}+m_0\frac{C_2-C_1^2}{1-\lambda_0}\sum_{h=1}^{\infty}\lambda_0^h$$
$$=-\frac{m_0}{1-\lambda_0}\frac{1}{1-\lambda_0}\sum_{h=0}^{\infty}\chi_h+m_0\frac{C_2-C_1^2}{1-\lambda_0}\frac{\lambda_0}{1-\lambda_0}=\frac{m_0}{(1-\lambda_0)^2}\left[-\sum_{h=0}^{\infty}\chi_h+\lambda_0(C_2-C_1^2)\right],\quad n\to\infty. \tag{A.103}$$

A similar analysis yields that $\sum_{h=1}^{n}h\ \mathrm{Cov}(Y_0,Y_hY_{h+k_n+1})$ converges as $n\to\infty$ so that $\lim_{n\to\infty}\sum_{h=1}^{n}\frac{h}{n}\ \mathrm{Cov}(Y_0,Y_hY_{h+k_n})=0$, from which, along with (A.103) and the fact that $\lim_{n\to\infty}\mathrm{Cov}(Y_0,Y_0Y_{k_n+1})=0$, gives that

$$\frac{1}{n}\sum_{h=1}^{n}(n-|h|)\mathrm{Cov}(Y_0,Y_hY_{h+k_n+1})\longrightarrow 2\frac{m_0}{(1-\lambda_0)^2}\left[-\sum_{h=0}^{\infty}\chi_h+\lambda_0(C_2-C_1^2)\right],\quad n\to\infty. \tag{A.104}$$

We then study the term $\sum_{h=-1}^{-n}\mathrm{Cov}(Y_0,Y_hY_{h+k_n+1})$. A change of index $h:=-h$ as well as the stationarity of the process $\{Y_n,\ n\in\mathbb{Z}\}$ yields

$$\sum_{h=-n}^{-1}\mathrm{Cov}(Y_0,Y_hY_{h+k_n+1})=\sum_{h=1}^{n}\mathrm{Cov}(Y_h,Y_0Y_{k_n+1}). \tag{A.105}$$

We split the righthandside of (A.105) into the two sums $\sum_{h=1}^{k_n}$ and $\sum_{h=k_n+1}^{n}$. Writing for $h=1,...,k_n$ that

$$\begin{aligned}\mathrm{Cov}(Y_h,Y_0Y_{k_n+1}) &= \mathrm{Cov}(Y_0,Y_hY_{k_n+1})+\mathbb{E}(Y_0)\mathbb{E}(Y_hY_{k_n+1})-\mathbb{E}(Y_h)\mathbb{E}(Y_0Y_{k_n+1})\\ &= \mathrm{Cov}(Y_0,Y_hY_{k_n+1})+C_1u_{k_n-h}-C_1u_{k_n}\\ &= \mathrm{Cov}(Y_0,Y_hY_{k_n+1})+C_1(u_{k_n-h}-C_1^2)-C_1(u_{k_n}-C_1^2),\end{aligned} \tag{A.106}$$

we notice that, similarly to (A.14), $|\mathrm{Cov}(Y_0,Y_hY_{k_n+1})|\leq K\left[\alpha_\epsilon\left(\lfloor h/2\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{h}{2}\rfloor}+\lambda_0^h\right]$ which, since the mixing coefficient function $\alpha_\epsilon(\cdot)$ is decreasing and $\lfloor h/2\rfloor\geq\lfloor(k_n/4)-1\rfloor$, $h=\lfloor k_n/2\rfloor,...,k_n$, yields that

$$\left|\sum_{h=\lfloor\frac{k_n}{2}\rfloor}^{k_n}\mathrm{Cov}(Y_0,Y_hY_{k_n})\right|\leq k_nK\left[\alpha_\epsilon\left(\left\lfloor\frac{k_n}{4}-1\right\rfloor\right)^{1-2/\beta}+\lambda_0^{\lfloor\frac{k_n}{4}-1\rfloor}+\lambda_0^{\lfloor\frac{k_n}{2}\rfloor}\right]\longrightarrow 0, \tag{A.107}$$

Next, writing $\mathrm{Cov}(Y_0,Y_hY_{k_n+1})=\mathrm{Cov}(Y_0Y_h,Y_{k_n+1})+\mathbb{E}(Y_0Y_h)\mathbb{E}(Y_{k_n+1})-\mathbb{E}(Y_0)\mathbb{E}(Y_hY_{k_n+1})=\mathrm{Cov}(Y_0Y_h,Y_{k_n+1})+C_1(u_{h-1}-C_1^2)+C_1(C_1^2-u_{k_n-h})$, and since, by an argument similar

to (A.107), $\sum_{h=1}^{\lfloor \frac{k_n}{2} \rfloor - 1} \mathrm{Cov}(Y_0 Y_h, Y_{k_n+1}) \longrightarrow 0$ as $n \to \infty$, and $\sum_{h=1}^{\lfloor \frac{k_n}{2} \rfloor - 1} (C_1^2 - u_{k_n - h}) = \sum_{h=k_n - \lfloor \frac{k_n}{2} \rfloor + 1}^{k_n - 1} (C_1^2 - u_h) \longrightarrow 0$ as $n \to \infty$, we thus get that

$$\sum_{h=1}^{\lfloor \frac{k_n}{2} \rfloor - 1} \mathrm{Cov}(Y_0, Y_h Y_{k_n+1}) \longrightarrow C_1 \sum_{h=1}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty,$$

which, coupled with (A.107), yields that

$$\sum_{h=1}^{k_n} \mathrm{Cov}(Y_0, Y_h Y_{k_n+1}) \longrightarrow C_1 \sum_{h=1}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty. \tag{A.108}$$

Pluged into (A.106), and since $k_n(u_{k_n} - C_1^2) = \mathrm{O}(k_n \lambda_0^{k_n}) \longrightarrow 0$ as $n \to \infty$ thanks to Corollary 3.2 and $\sum_{h=1}^{k_n} (u_{k_n - h} - C_1^2) = \sum_{h=0}^{k_n - 1} (u_h - C_1^2) \longrightarrow \sum_{h=0}^{\infty} (u_{h-1} - C_1^2)$, we thus get

$$\sum_{h=1}^{k_n} \mathrm{Cov}(Y_h, Y_0 Y_{k_n+1}) \longrightarrow 2C_1 \sum_{h=1}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty. \tag{A.109}$$

We now consider the $\sum_{h=k_n+1}^{n}$ sum in the splitting of righthandside sum of (A.105). Using the stationarity of $\{Y_n,\ n \in \mathbb{Z}\}$ and a change of index $h := h - k_n$, we first write

$$\sum_{h=k_n+1}^{n} \mathrm{Cov}(Y_h, Y_0 Y_{k_n+1}) = \sum_{h=0}^{n-k_n-1} \mathrm{Cov}(Y_{-k_n-1} Y_0, Y_h). \tag{A.110}$$

We use a dominated convergence argument to determine the limit of (A.110) as $n \to \infty$. Similarly to (A.14), the following inequality holds

$$|\mathrm{Cov}(Y_{-k_n-1} Y_0, Y_h)| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{h}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{h}{2} \rfloor} + \lambda_0^h \right], \tag{A.111}$$

the upper bound being summable. Next, we decompose the summand in the righthandside of (A.110) in the same spirit as (A.106) as

$$\mathrm{Cov}(Y_{-k_n-1} Y_0, Y_h) = \mathrm{Cov}(Y_{-k_n-1}, Y_0 Y_h) + C_1(u_{h-1} - C_1^2) - C_1(u_{k_n} - C_1^2). \tag{A.112}$$

For fixed $h \in \mathbb{N}$, a standard argument yields that

$$|\mathrm{Cov}(Y_{-k_n-1}, Y_0 Y_h)| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{k_n + 1}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{k_n+1}{2} \rfloor} + \lambda_0^h \right] \longrightarrow 0$$

as $n \to \infty$. We already saw that $\lim_{n \to \infty} u_{k_n} - C_1^2 = 0$, so that from (A.112) the following limit holds

$$\mathrm{Cov}(Y_{-k_n-1} Y_0, Y_h) \longrightarrow C_1(u_{h-1} - C_1^2), \quad n \to \infty,\ \forall h \in \mathbb{N}. \tag{A.113}$$

Hence, (A.110) along with (A.111) and (A.113) yield by the dominated convergence theorem that

$$\sum_{h=k_n+1}^{n} \mathrm{Cov}(Y_h, Y_0 Y_{k_n+1}) \longrightarrow \sum_{h=0}^{\infty} C_1(u_{h-1} - C_1^2) = C_1(C_2 - C_1^2) + C_1 \sum_{h=1}^{\infty} (u_{h-1} - C_1^2),$$

which, coupled with (A.109) as well as (A.105), yields

$$\sum_{h=-n}^{-1} \mathrm{Cov}(Y_0, Y_h Y_{h+k_n+1}) \longrightarrow C_1(C_2 - C_1^2) + 3C_1 \sum_{h=1}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty,$$

which in turn gives

$$\frac{1}{n} \sum_{h=-n}^{-1} (n - |h|)\mathrm{Cov}(Y_0, Y_h Y_{h+k_n+1}) \longrightarrow 2C_1(C_2 - C_1^2) + 3C_1 \sum_{h=1}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty.$$

Gathering (A.104) and the above limit we eventually arrive after some easy calculation at

$$\begin{aligned} \frac{1}{n} \sum_{h=-n}^{n} (n - |h|)\mathrm{Cov}(Y_0, Y_h Y_{h+k_n+1}) \longrightarrow & -2\frac{m_0}{(1-\lambda_0)^2} \sum_{h=0}^{\infty} \chi_h \\ & + \frac{C_1}{1-\lambda_0}(C_2 - C_1^2) + 3C_1 \sum_{h=1}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty. \end{aligned} \tag{A.114}$$

We then study $\mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1})$. The limit of $\sum_{h=1}^{n} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1})$ is obtained by splitting the sum in three blocks, namely $\sum_{h=1}^{\lfloor k_n/2 \rfloor}$, $\sum_{\lfloor k_n/2 \rfloor + 1}^{k_n+1}$ and $\sum_{h=k_n+2}^{n}$. We start by the summation over $h = 1, ..., \lfloor k_n/2 \rfloor$. We exploit here the fact that h is "far" from $k_n + 1$, and write that

$$\begin{aligned} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1}) &= \mathrm{Cov}(Y_0 Y_h, Y_{k_n+1} Y_{h+k_n+1}) + \mathbb{E}(Y_0 Y_h)\mathbb{E}(Y_{k_n+1} Y_{k_n+1+h}) \\ &\quad - \mathbb{E}(Y_0 Y_{k_n+1})\mathbb{E}(Y_h Y_{k_n+1+h}) \\ &= \mathrm{Cov}(Y_0 Y_h, Y_{k_n+1} Y_{h+k_n+1}) + u_{h-1}^2 - u_{k_n}^2. \end{aligned} \tag{A.115}$$

Similarly to (A.14) we have

$$|\mathrm{Cov}(Y_0 Y_h, Y_{k_n+1} Y_{h+k_n+1})| \leq K \left[\alpha_\epsilon \left(\left\lfloor \frac{k_n - h}{2} \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{k_n - h}{2} \rfloor} + \lambda_0^{k_n+1-h} \right]$$

which, since the mixing coefficient function $\alpha_\epsilon(\cdot)$ is decreasing and $\lfloor (k_n - h)/2 \rfloor \geq \lfloor (k_n/4) - 1 \rfloor$, $h = 1, ..., \lfloor k_n/2 \rfloor$, entails that

$$\begin{aligned} \sum_{h=1}^{\lfloor k_n/2 \rfloor} |\mathrm{Cov}(Y_0 Y_h, Y_{k_n+1} Y_{h+k_n+1})| \leq K \lfloor k_n/2 \rfloor \left[\alpha_\epsilon \left(\left\lfloor \frac{k_n}{4} - 1 \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{k_n}{4} - 1 \rfloor} + \lambda_0^{(k_n+1)/2} \right] \\ \longrightarrow 0, \quad n \to \infty. \end{aligned} \tag{A.116}$$

Remembering that $C_1^2 - u_k = \mathrm{O}(\lambda_0^k)$ from Corollary 3.2, so that $\lim_{k \to \infty} k(C_1^2 - u_k) = 0$, we also have, writing $u_{h-1}^2 - u_{k_n-1}^2 = (u_{h-1}^2 - C_1^4) + (C_1^4 - u_{k_n-1}^2)$:

$$\sum_{h=1}^{\lfloor k_n/2 \rfloor} (u_{h-1}^2 - u_{k_n-1}^2) = \sum_{h=1}^{\lfloor k_n/2 \rfloor} (u_{h-1}^2 - C_1^4) + \lfloor k_n/2 \rfloor (C_1^4 - u_{k_n-1}^2) \longrightarrow \sum_{h=1}^{\infty} (u_{h-1}^2 - C_1^4), \quad n \to \infty, \tag{A.117}$$

as indeed $\lfloor k_n/2 \rfloor (C_1^4 - u_{k_n-1}^2) = \mathrm{O}\left(\lfloor k_n/2 \rfloor (C_1^2 - u_{k_n-1})\right) \longrightarrow 0$ as $n \to \infty$. Gathering (A.116) and (A.117) in (A.115) thus yields

$$\sum_{h=1}^{\lfloor k_n/2 \rfloor} \mathrm{Cov}(Y_0 Y_{k_n}, Y_h Y_{h+k_n}) \longrightarrow \sum_{h=1}^{\infty} (u_{h-1}^2 - C_1^4), \quad n \to \infty. \tag{A.118}$$

We now consider the summation over $h = \lfloor k_n/2 \rfloor + 1, ..., k_n + 1$. We this time exploit the fact that h and $k_n + 1 + h$ are respectively "far" from 0 and $k_n + 1$, and write

$$\begin{aligned} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1}) &= \mathrm{Cov}(Y_0 Y_h Y_{k_n+1}, Y_{h+k_n+1}) + [\mathrm{Cov}(Y_0, Y_h Y_{k_n+1}) \\ &\quad + \mathbb{E}(Y_0)\mathbb{E}(Y_h Y_{k_n+1})]\mathbb{E}(Y_{h+k_n+1}) - \mathbb{E}(Y_0 Y_{k_n+1})\mathbb{E}(Y_h Y_{h+k_n+1}) \\ &= \mathrm{Cov}(Y_0 Y_h Y_{k_n+1}, Y_{h+k_n+1}) + [\mathrm{Cov}(Y_0, Y_h Y_{k_n+1}) \\ &\quad + C_1 u_{k_n-h}] C_1 - u_{k_n}^2. \end{aligned} \tag{A.119}$$

Similarly to (A.116):

$$\begin{aligned} \sum_{\lfloor k_n/2 \rfloor + 1}^{k_n+1} |\mathrm{Cov}(Y_0 Y_h Y_{k_n+1}, Y_{h+k_n} + 1)| &\leq K \lfloor k_n/2 \rfloor \left[\alpha_\epsilon \left(\left\lfloor \frac{k_n}{4} - 1 \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{k_n}{4} - 1 \rfloor} + \lambda_0^{\frac{k_n+1}{2}} \right] \\ &\longrightarrow 0, \quad n \to \infty, \\ \sum_{\lfloor k_n/2 \rfloor + 1}^{k_n+1} |\mathrm{Cov}(Y_0, Y_h Y_{k_n+1})| &\leq K \lfloor k_n/2 \rfloor \left[\alpha_\epsilon \left(\left\lfloor \frac{k_n}{4} - 1 \right\rfloor \right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{k_n}{4} - 1 \rfloor} + \lambda_0^{\frac{k_n+1}{2}} \right] \\ &\longrightarrow 0, \quad n \to \infty. \end{aligned}$$

And, writing $C_1^2 u_{k_n-h} - u_{k_n}^2 = C_1^2 (u_{k_n-h} - C_1^2) + C_1^4 - u_{k_n}^2$, we obtain, similarly to (A.117):

$$\begin{aligned} \sum_{h=\lfloor k_n/2 \rfloor + 1}^{k_n+1} C_1^2 u_{k_n-h} - u_{k_n}^2 = C_1^2 \sum_{h=0}^{k_n - \lfloor k_n/2 \rfloor - 1} (u_{h-1} - C_1^2) + (k_n - \lfloor k_n/2 \rfloor)(C_1^4 - u_{k_n}^2) \\ \longrightarrow C_1^2 \sum_{h=0}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty, \end{aligned}$$

so that we obtain from (A.119) the following limit

$$\sum_{h=\lfloor k_n/2 \rfloor + 1}^{k_n+1} \mathrm{Cov}(Y_0 Y_{k_n}, Y_h Y_{h+k_n}) \longrightarrow C_1^2 \sum_{h=0}^{\infty} (u_{h-1} - C_1^2), \quad n \to \infty. \tag{A.120}$$

Then, we consider the summation over $h = k_n + 2, ..., n$. The expansion established in (A.13) reads here

$$\begin{aligned} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1}) = \sum_{s=0}^{k_n} \lambda_0^s \left[\sum_{t=0}^{h-k_n-2} \lambda_0^t \mathrm{Cov}(Y_0 Y_{k_n+1}, \epsilon_{h-t} \epsilon_{h+k_n+1-s}) \right. \\ \left. + \lambda_0^{h-k_n-1} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_{k_n+1} \epsilon_{h+k_n+1-s}) \right] + \lambda_0^{k_n+1} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h^2). \end{aligned} \tag{A.121}$$

We study the summation over $h = k_n + 2, ..., n$ of each term in (A.121), and start with the first term, which reads, thanks to a change of variable $r := h - k_n - 1$,

$$\sum_{h=k_n+2}^{n} \sum_{s=0}^{k_n} \sum_{t=0}^{h-k_n-2} \lambda_0^{s+t} \mathrm{Cov}(Y_0 Y_{k_n+1}, \epsilon_{h-t}\epsilon_{h+k_n+1-s}) = \sum_{r=1}^{n-k_n-1} J_1(r, k_n), \tag{A.122}$$

$$J_1(r, k_n) := \sum_{s=0}^{k_n} \sum_{t=0}^{r-1} \lambda_0^{s+t} \mathrm{Cov}(Y_0 Y_{k_n+1}, \epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s}).$$

Studying the limit of (A.122) is done by applying the dominant convergence theorem. We see from (A.8) that $\left|\sum_{t=0}^{r-1} \lambda_0^t \mathrm{Cov}(Y_0 Y_{k_n+1}, \epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s})\right| \leq K\left[\alpha_\epsilon \left(\lfloor (r-1)/2 \rfloor\right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{r-1}{2} \rfloor}\right]$ for all $s = 0, ..., k_n$, so that $|J_1(r, k_n)|$ is dominated as follows

$$|J_1(r, k_n)| \leq K\left[\alpha_\epsilon \left(\left\lfloor \frac{r-1}{2} \right\rfloor\right)^{1-2/\beta} + \lambda_0^{\lfloor \frac{r-1}{2} \rfloor}\right]$$

of which righhandside is summable. As to the pointwise convergence of $J_1(r, k_n)$ for fixed r as $n \to \infty$, we use again a domination argument in the variables s and t. Indeed we have that $|\lambda_0^{s+t} \mathrm{Cov}(Y_0 Y_{k_n+1}, \epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s})| \leq K\lambda_0^{s+t}$ for some constant K thanks to the Cauchy Schwarz inequality, and we have

$$\begin{aligned}
&\mathrm{Cov}(Y_0 Y_{k_n+1}, \epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s}) = \mathrm{Cov}(Y_0, Y_{k_n+1}\epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s}) \\
&+\mathbb{E}(Y_0)\mathbb{E}(Y_{k_n+1}\epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s}) - \mathbb{E}(Y_0 Y_{k_n+1})\mathbb{E}(\epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s}) \\
&= \mathrm{Cov}(Y_0, Y_{k_n+1}\epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s}) \\
&+C_1\mathbb{E}(Y_0\epsilon_{r-t}\epsilon_{r+k_n+1-s}) - u_{k_n}\mathbb{E}(\epsilon_{r-t}\epsilon_{r+k_n+1-s})
\end{aligned} \tag{A.123}$$

Standard arguments yields the following limits as $n \to \infty$:

$$\begin{aligned}
\mathrm{Cov}(Y_0, Y_{k_n+1}\epsilon_{r+k_n+1-t}\epsilon_{r+2k_n+2-s}) &\longrightarrow 0, \\
\mathbb{E}(Y_0\epsilon_{r-t}\epsilon_{r+k_n+1-s}) &\longrightarrow \mathbb{E}(Y_0\epsilon_{r-t})m_0, \\
u_{k_n}\mathbb{E}(\epsilon_{r-t}\epsilon_{r+k_n+1-s}) &\longrightarrow C_1^2 m_0^2,
\end{aligned}$$

which, plugged into (A.123), yields the following thanks to Relation (11)

$$\begin{aligned}
J_1(r, k_n) &\longrightarrow \sum_{s=0}^{\infty} \sum_{t=0}^{r-1} \lambda_0^{s+t}[C_1\mathbb{E}(Y_0\epsilon_{r-t})m_0 - C_1^2 m_0^2] \\
&= \frac{C_1 m_0}{1-\lambda_0} \sum_{t=0}^{r-1} \lambda_0^t [v_{r-t-1} - C_1 m_0] = -\frac{C_1 m_0}{1-\lambda_0} \sum_{t=0}^{r-1} \lambda_0^t \chi_{r-t-1} \\
&= -\frac{C_1 m_0}{1-\lambda_0} \lambda_0^r \sum_{t=1}^{r} \lambda_0^{-t} \chi_{t-1}, \quad n \to \infty.
\end{aligned}$$

Hence the following limit from (A.122):

$$\sum_{h=k_n+2}^{n} \sum_{s=0}^{k_n} \sum_{t=0}^{h-k_n-2} \lambda_0^{s+t} \mathrm{Cov}(Y_0 Y_{k_n+1}, \epsilon_{h-t}\epsilon_{h+k_n+1-s}) \longrightarrow -\frac{C_1 m_0}{1-\lambda_0} \sum_{r=1}^{\infty} \lambda_0^r \sum_{t=1}^{r} \lambda_0^{-t} \chi_{t-1}$$

$$= -\frac{C_1 m_0}{(1-\lambda_0)^2} \sum_{t=1}^{\infty} \lambda_0^{-t} \chi_{t-1}. \quad \text{(A.124)}$$

We now pass on to the second term in (A.121) and we write

$$\sum_{h=k_n+2}^{n} \sum_{s=0}^{k_n} \lambda_0^{s+h} \mathrm{Cov}(Y_0, Y_0 \epsilon_{k_n+1+h-s}) = \sum_{r=1}^{n-k_n-1} J_2(r, k_n), \quad \text{(A.125)}$$

$$J_2(r, k_n) = \sum_{s=0}^{k_n} \lambda_0^{s+r+k_n+1} \mathrm{Cov}(Y_0, Y_0 \epsilon_{r+2k_n+2-s}).$$

From the Cauchy Schwarz inequality as well as crude bounding we get that $|J_2(r, k_n)|$ is upper bounded by $K\lambda_0^{r+k_n}$ for some constant K, so that from (A.124) we obtain that

$$\left| \sum_{h=k_n+2}^{n} \sum_{s=0}^{k_n} \lambda_0^{s+h} \mathrm{Cov}(Y_0, Y_0 \epsilon_{k_n+1+h-s}) \right| = \mathrm{O}(\lambda_0^{k_n}) \longrightarrow 0, \quad n \to \infty. \quad \text{(A.126)}$$

As to the third term in (A.121), standard estimates yield that $\sum_{h=1}^{\infty} \mathrm{Cov}(Y_0, Y_h^2)$ is a convergent series, so that we get

$$\sum_{h=k_n+2}^{n} \lambda_0^{k_n+1} \mathrm{Cov}(Y_0, Y_h^2) = \mathrm{O}(\lambda_0^{k_n}) \longrightarrow 0, \quad n \to \infty, \quad \text{(A.127)}$$

so that, gathering (A.124), (A.126) and (A.127), we get from (A.121) that

$$\sum_{h=k_n+2}^{n} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1}) \longrightarrow -\frac{C_1 m_0}{(1-\lambda_0)^2} \sum_{t=1}^{\infty} \lambda_0^{-t} \chi_{t-1}, \quad n \to \infty. \quad \text{(A.128)}$$

Finally, gathering (A.118), (A.120) and (A.128) we arrive at the following limit as $n \to \infty$:

$$\sum_{h=1}^{n} \mathrm{Cov}(Y_0 Y_{k_n}, Y_h Y_{h+k_n+1}) \longrightarrow \sum_{h=1}^{\infty} (u_{h-1}^2 - C_1^4) + C_1^2 \sum_{h=0}^{\infty} (u_{h-1} - C_1^2) - \frac{C_1 m_0}{(1-\lambda_0)^2} \sum_{t=1}^{\infty} \lambda_0^{-t} \chi_{t-1}. \quad \text{(A.129)}$$

A similar analysis yields that $\sum_{h=1}^{n} h \ \mathrm{Cov}(Y_0 Y_{k_n}, Y_h Y_{h+k_n})$ converges as $n \to \infty$ so that $\lim_{n\to\infty} \sum_{h=1}^{n} \frac{h}{n} \mathrm{Cov}(Y_0 Y_{k_n}, Y_h Y_{h+k_n}) = 0$. Also, it is not difficult to check that $\lim_{n\to\infty} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_0 Y_{k_n+1}) = \lim_{n\to\infty} \mathrm{Cov}(Y_0^2, Y_{k_n+1}^2) + C_2^2 - C_1^4 + C_1^2 - u_{k_n}^2 = C_2^2 - C_1^4$.

As to $\sum_{h=-n}^{-1} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1})$, we use again by a stationarity argument and a change of index that it is equal to $\sum_{h=1}^{n} \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1})$, so that its limit as $n \to \infty$ is also given by the righthandside of (A.129). Thus, we obtain the following limit

$$\frac{1}{n} \sum_{h=-n}^{n} (n-|h|) \mathrm{Cov}(Y_0 Y_{k_n+1}, Y_h Y_{h+k_n+1}) \longrightarrow C_2^2 - C_1^4 + 2 \sum_{h=1}^{\infty} (u_{h-1}^2 - C_1^4) + 2C_1^2 \sum_{h=0}^{\infty} (u_{h-1} - C_1^2)$$
$$- 2 \frac{C_1 m_0}{(1-\lambda_0)^2} \sum_{t=1}^{\infty} \lambda_0^{-t} \chi_{t-1}, \quad n \to \infty. \quad \text{(A.130)}$$

Hence, in view of (A.96), (A.97) and the limits (A.98), (A.114) and (A.130), we thus proved the expression (53) for $\mathfrak{S}$.

Appendix A.8.0.2. $\diamond$ *Step 2:.* Recalling for all $m \in \mathbb{N}$ the processes $\{Y_{n,m},\ n \in \mathbb{Z}\}$ and $\{Z_{n,m},\ n \in \mathbb{Z}\}$ in (A.15), as well as introducing

$$\begin{aligned} X_{j,m}(k_n) &= X^1_{j,m}(k_n) + X^2_{j,m}(k_n) \\ X^1_{j,m}(k_n) &= [Y_{j,m} - \mathbb{E}(Y_{j,m}), Y_{j+k_n+1,m} - \mathbb{E}(Y_{j+k_n+1,m})]' \\ X^2_{j,m}(k_n) &= [Z_{j,m} - \mathbb{E}(Z_{j,m}), Z_{j+k_n+1,m} - \mathbb{E}(Z_{j+k_n+1,m})]', \end{aligned} \tag{A.131}$$

we now follow the scheme in Step 2 of the proof of Theorem 4.2 with again the non negligible difference that k_0 is now replaced by k_n, and show that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} X^1_{i,m}(k_n) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \mathfrak{S}^m), \quad n \to \infty, \tag{A.132}$$

for some semi-definite positive matrix $\mathfrak{S}^m \in \mathbb{R}^{2\times 2}$. For this we need to verify Condition (2) in Francq and Zakoïan (2005) which says that the following limit exists and defines that matrix:

$$\begin{aligned} \frac{1}{n}\mathbb{E}\left(\left(\sum_{i=1}^{n} X^1_{i,m}(k_n)\right)\left(\sum_{i=1}^{n} X^1_{i,m}(k_n)\right)'\right) &= \frac{1}{n}\mathbb{E}\left(\sum_{1\le i,j\le n} X^1_{i,m}(k_n)' X^1_{j,m}(k_n)\right) \\ &= \frac{1}{n}\sum_{h=-n}^{n} (n-|h|)\mathbb{E}(X^1_{0,m}(k_n), X^1_{h,m}(k_n)) \longrightarrow \mathfrak{S}^m \end{aligned} \tag{A.133}$$

as $n \to \infty$, which we prove now. For this we first express the expectation above as covariances as follows

$$\mathbb{E}(X^1_{0,m}(k_n), X^1_{h,m}(k_n)) = \begin{bmatrix} \mathrm{Cov}(Y_{0,m}, Y_{h,m}) & \mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m}) \\ \mathrm{Cov}(Y_{k_n+1,m}, Y_{h,m}) & \mathrm{Cov}(Y_{k_n+1,m}, Y_{h+k_n+1,m}) \end{bmatrix}, \tag{A.134}$$

where $\mathrm{Cov}(Y_{k_n+1,m}, Y_{h+k_n+1,m}) = \mathrm{Cov}(Y_{0,m}, Y_{h,m})$ by stationarity. We then study the convergence of the series with general terms given by the entries of the above matrix, i.e. convergence of $\sum_{h=-n}^{n} \mathrm{Cov}(Y_{0,m}, Y_{h,m})$, $\sum_{h=-n}^{n} \mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m})$ and $\sum_{h=-n}^{n} \mathrm{Cov}(Y_{k_n+1,m}, Y_{h,m})$ as $n \to \infty$. Standard arguments yield the convergence of $\sum_{h=-n}^{n} \mathrm{Cov}(Y_{0,m}, Y_{h,m})$, as the summand does not depend on n, so that we are going to focus on the convergence of $\sum_{h=-n}^{n} \mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m})$, the one with the summand $\mathrm{Cov}(Y_{k_n+1,m}, Y_{h,m})$ being dealt with similarly. Remembering the definition (A.15) for $Y_{0,m}$ and $Y_{h,m}$, we have

$$\mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m}) = \sum_{r,r'=0}^{m} \mathrm{Cov}(\theta_0^{(r)} \circ \epsilon_{-r}, \theta_{h+k_n+1}^{(r')} \circ \epsilon_{h+k_n+1-r'}). \tag{A.135}$$

It thus suffices to prove that $\sum_{h=-n}^{n} \mathrm{Cov}(\theta_0^{(r)} \circ \epsilon_{-r}, \theta_{h+k_n+1}^{(r')} \circ \epsilon_{h+k_n+1-r'})$ admits a limit as $n \to \infty$ for all r and r' in $0, ..., m$ to be able to conclude as to the limit of $\sum_{h=-n}^{n} \mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m})$. We prove that the limit exists when r and r' are both 0 w.l.o.g. Remembering that $\theta_n^{(0)} = \mathrm{Id}$, we thus need to study the limit of $\sum_{h=-n}^{n} \mathrm{Cov}(\epsilon_0, \epsilon_{h+k_n+1})$ which we again split into $\sum_{h=0}^{n}$ and $\sum_{h=-n}^{-1}$. A dominated convervgence argument gives that

$$\sum_{h=0}^{n} \mathrm{Cov}(\epsilon_0, \epsilon_{h+k_n+1}) \longrightarrow 0, \quad n \to \infty. \tag{A.136}$$

Next, writing that

$$\sum_{h=-n}^{-1} \mathrm{Cov}(\epsilon_0, \epsilon_{h+k_n+1}) = \sum_{h=1}^{n} \mathrm{Cov}(\epsilon_0, \epsilon_{k_n+1-h}) = \sum_{h=1}^{k_n+1} + \sum_{h=k_n+2}^{n}, \tag{A.137}$$

of which two latter sums have the following limits (remember that $k_n = \mathrm{o}(n)$ by assumption, so that $n - k_n \longrightarrow \infty$ as $n \to \infty$):

$$\begin{aligned}
\sum_{h=1}^{k_n+1} \mathrm{Cov}(\epsilon_0, \epsilon_{k_n+1-h}) &= \sum_{h=0}^{k_n} \mathrm{Cov}(\epsilon_0, \epsilon_h) = \sum_{h=0}^{k_n} \nu_h \\
&\longrightarrow \sum_{h=0}^{\infty} \nu_h, \quad n \to \infty, \\
\sum_{h=k_n+2}^{n} \mathrm{Cov}(\epsilon_0, \epsilon_{k_n+1-h}) &= \sum_{h=k_n+2}^{n} \mathrm{Cov}(\epsilon_0, \epsilon_{h-k_n-1}) = \sum_{h=1}^{n-k_n-1} \mathrm{Cov}(\epsilon_0, \epsilon_h) \\
&= \sum_{h=1}^{n-k_n-1} \nu_h \longrightarrow \sum_{h=1}^{\infty} \nu_h, \quad n \to \infty
\end{aligned}$$

which, plugged into (A.137) and (A.136), yields the following limit

$$\sum_{h=-n}^{n} \mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m}) \longrightarrow \nu_0 + 2\sum_{h=1}^{\infty} \nu_h, \quad n \to \infty.$$

Note again that what we wanted was simply the existence of the limit, of which explicit expression is just a bonus. All in all, we thus have from (A.135) that $\sum_{h=-n}^{n} \mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m})$ admits a limit as $n \to \infty$. Similarly, one proves the existence of $\sum_{h=-n}^{n} |h\mathrm{Cov}(Y_{0,m}, Y_{h+k_n+1,m})|$ so that, by (A.134) and the dominated convergence theorem the limit $\mathfrak{S}^m$ indeed exists in (A.133).

We now conclude by saying that Condition (1) in Francq and Zakoïan (2005) is satisfied with $\nu^* = 2\beta - 2$. As discussed in the beginning of Francq and Zakoïan (2005, Section 2), and using the notation therein, Condition (3) is verified here with $T_n = 2h_n$, and thanks to the assumption (58). Hence the convergence in distribution (A.132) holds.

Appendix A.8.0.3. $\diamond$ *Step 3:.* As in Step 3 of the proof of Proposition 6.4, we prove that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\left|\left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{i,m}^2(k_n) \right|\right| > \epsilon \right) = 0, \quad \forall \epsilon > 0. \tag{A.138}$$

The equivalent of (A.36) is here

$$\begin{aligned}
\mathbb{P}\left(\left|\left| \frac{1}{\sqrt{n}} \sum_{i=1}^{n} X_{i,m}^2(k_n) \right|\right| > \epsilon \right) &\leq \frac{1}{n\epsilon^2} \mathbb{E}\left(\left|\left| \sum_{i=1}^{n} X_{i,m}^2(k_n) \right|\right|^2 \right) \\
&= \frac{1}{n\epsilon^2} \sum_{h=-n}^{n} (n - |h|) \mathbb{E}\left(X_{0,m}^2(k_n)' X_{h,m}^2(k_n) \right). \quad \text{(A.139)}
\end{aligned}$$

We see that it suffices to prove that

$$\sum_{h=-n}^{n} \|\mathbb{E}\left(X_{0,m}^2(k_n)'X_{h,m}^2(k_n)\right)\| \leq K\upsilon^m, \quad \sum_{h=-n}^{n} |h|.\|\mathbb{E}\left(X_{0,m}^2(k_n)'X_{h,m}^2(k_n)\right)\| \leq K\upsilon^m \tag{A.140}$$

in order to prove (A.138). We only prove this property for $\sum_{h=-n}^{n}\mathbb{E}\left(X_{0,m}^2(k_n)'X_{h,m}^2(k_n)\right)$, the proof being similar for $\sum_{h=-n}^{n} h\mathbb{E}\left(X_{0,m}^2(k_n)'X_{h,m}^2(k_n)\right)$. For this, and since, as in (A.134), one has the expression

$$\mathbb{E}(X_{0,m}^2(k_n), X_{h,m}^2(k_n)) = \begin{bmatrix} \mathrm{Cov}(Z_{0,m}, Z_{h,m}) & \mathrm{Cov}(Z_{0,m}, Z_{h+k_n+1,m}) \\ \mathrm{Cov}(Z_{k_n+1,m}, Z_{h,m}) & \mathrm{Cov}(Z_{0,m}, Z_{h,m}) \end{bmatrix}, \tag{A.141}$$

we saw that it suffices to prove that

$$\sum_{h=-n}^{n} |\mathrm{Cov}(Z_{0,m}, Z_{h+k_n+1})| \leq K\upsilon^m, \quad \forall n \in \mathbb{N}, \tag{A.142}$$

the other terms in (A.141) being dealt with similarly. Similarly to the decomposition (A.24), we write thanks to the definition (A.15), and for $h = 0, ..., n$:

$$\begin{aligned} \mathrm{Cov}(Z_{0,m}, Z_{h+k_n+1}) &= I_1(h,m,k_n) + I_2(h,m,k_n), \\ I_1(h,m,k_n) &:= \sum_{m+1\leq i<\lfloor h/2\rfloor+k_n+1} \mathrm{Cov}\left(Z_{0,m}, [\theta_{h+k_n+1}^{(i)} \circ \epsilon_{h+k_n+1-i}]\right), \\ I_2(h,m,k_n) &:= \sum_{\max(\lfloor h/2\rfloor+k_n+1, m+1)\leq i} \mathrm{Cov}\left(Z_{0,m}, [\theta_{h+k_n+1}^{(i)} \circ \epsilon_{h+k_n+1-i}]\right), \end{aligned} \tag{A.143}$$

with the convention that the sum over an empty set is equal to 0 (which is the case for example for $I_1(h,m,k_n)$ when $\lfloor h/2\rfloor \geq m+1$). An argument similar to (A.25) and (A.28) yields the following inequalities

$$\begin{aligned} |I_1(h,m,k_n)| &\leq K \sum_{m+1\leq i<\lfloor h/2\rfloor+k_n+1} \|Z_{0,m} - \mathbb{E}(Z_{0,m})\|_\beta \\ &\quad \cdot \left\|\theta_{h+k_n+1}^{(i)} \circ \epsilon_{h+k_n+1-i} - \mathbb{E}\left(\theta_{h+k_n+1}^{(i)} \circ \epsilon_{h+k_n+1-i}\right)\right\|_\beta \alpha_\epsilon(\lfloor h/2\rfloor)^{1-2/\beta}, \\ &\quad h = 0, ..., n, \\ \|Z_{0,m} - \mathbb{E}(Z_{0,m})\|_\beta &\leq K\nu^m, \end{aligned}$$

leading to the bound

$$|I_1(h,m,k_n)| \leq K\upsilon^m \alpha_\epsilon\left(\lfloor h/2\rfloor\right)^{1-2/\beta}, \quad h = 0, ..., n. \tag{A.144}$$

As to $I_2(h,m,k_n)$, we need to split the sum further according to whether $\lfloor h/2\rfloor + k_n + 1$ is larger or less than $m+1$. When it is larger than $m+1$ then an upper bound as in (A.30) holds, so that we obtain

$$\begin{aligned} |I_2(h,m,k_n)| &\leq K\nu^{\lfloor\lfloor h/2\rfloor+k_n+1}\nu^m \mathbb{1}_{[\lfloor h/2\rfloor+k_n+1\geq m+1]} + K\nu^m\nu^m \mathbb{1}_{[\lfloor h/2\rfloor+k_n+1<m+1]}, \\ &\leq K\nu^{\lfloor\lfloor h/2\rfloor}\nu^m \mathbb{1}_{[\lfloor h/2\rfloor+k_n+1\geq m+1]} + K\nu^{2m} \mathbb{1}_{[h<2(m+1)]}, \ h = 0, ..., n \end{aligned}$$

where we used the fact that $\lfloor h/2 \rfloor + k_n + 1 < m + 1$ implies that $h < 2(m+1)$. Summing (A.144) and the above inequality on $h = 0, ..., n$ leads to

$$\sum_{h=0}^{n} |\mathrm{Cov}(Z_{0,m}, Z_{h+k_n+1})| \leq K\nu^m \left[\sum_{h=0}^{\infty} \alpha_\epsilon \left(\lfloor h/2 \rfloor \right)^{1-2/\beta} \right] + K\nu^m \left[\sum_{h=0}^{\infty} \nu^{\lfloor h/2 \rfloor} \right] + 2(m+1)\nu^{2m} \leq K\nu^m. \quad \text{(A.145)}$$

We now consider $\sum_{h=-n}^{-1} |\mathrm{Cov}(Z_{0,m}, Z_{h+k_n+1})|$. We first observe that, by a change of index and stationarity,

$$\sum_{h=-n}^{-1} |\mathrm{Cov}(Z_{0,m}, Z_{h+k_n+1})| = \sum_{h=1}^{n} |\mathrm{Cov}(Z_{k_n+1,m}, Z_{h,m})| = \sum_{h=1}^{k_n+1} + \sum_{h=k_n+2}^{n}. \quad \text{(A.146)}$$

Let us first notice that the following bound holds thanks to standard arguments:

$$|\mathrm{Cov}(Z_{h,m}, Z_{0,m})| \leq K\nu^m \left[\alpha_\epsilon \left(\lfloor h/2 \rfloor \right)^{1-2/\beta} + \upsilon^{\lfloor h/2 \rfloor} \right],$$

so that the two sums on the righthandside of (A.146) verify the bounds

$$\begin{aligned}
\sum_{h=1}^{k_n+1} |\mathrm{Cov}(Z_{k_n+1,m}, Z_{h,m})| &= \sum_{h=1}^{k_n+1} |\mathrm{Cov}(Z_{k_n+1-h,m}, Z_{0,m})| = \sum_{h=0}^{k_n} |\mathrm{Cov}(Z_{h,m}, Z_{0,m})| \\
&\leq \sum_{h=0}^{\infty} |\mathrm{Cov}(Z_{h,m}, Z_{0,m})| \leq K\nu^m \sum_{h=0}^{\infty} \left[\alpha_\epsilon \left(\lfloor h/2 \rfloor \right)^{1-2/\beta} + \upsilon^{\lfloor h/2 \rfloor} \right] \\
\sum_{h=k_n+2}^{n} |\mathrm{Cov}(Z_{k_n+1,m}, Z_{h,m})| &= \sum_{h=k_n+2}^{n} |\mathrm{Cov}(Z_{0,m}, Z_{h-k_n+1,m})| = \sum_{h=1}^{n-k_n-1} |\mathrm{Cov}(Z_{0,m}, Z_{h,m})| \\
&\leq \sum_{h=0}^{\infty} |\mathrm{Cov}(Z_{h,m}, Z_{0,m})| \leq K\nu^m \sum_{h=0}^{\infty} \left[\alpha_\epsilon \left(\lfloor h/2 \rfloor \right)^{1-2/\beta} + \upsilon^{\lfloor h/2 \rfloor} \right]
\end{aligned}$$

which, plugged into (A.146), yields that $\sum_{h=-n}^{-1} |\mathrm{Cov}(Z_{0,m}, Z_{h+k_n+1})| \leq K\nu^m$. In turn, this latter bound associated to (A.145) yields (A.142) hence (A.140). Thus, all in all, (A.138) holds.

Appendix A.8.0.4. $\diamond$ *Step 4:.* As in the final step of the proof of Theorem 4.2, (A.132) and (A.138) yield that $\lim_{m\to\infty} \mathfrak{S}^m = \mathfrak{S}$ defined by (53), as well as the convergence in disrtibution (59). $\square$

Appendix A.9. Proof of Theorem 6.1

Setting $k = k_n$ in (56), and remembering from (48) that $\hat{S}_n = \psi_{k_n}(\bar{Y}_n, \bar{Y}_{k_n+1,n})$, we get

$$||\hat{S}_n - \ln \lambda_0||_2 \leq K_S \left(\frac{1}{\lambda_-^{k_n}} \left|\left| \frac{\hat{S}_n}{n} \right|\right|_2 + \frac{1}{k_n} \right) = K_S \left(\frac{1}{\lambda_-^{k_n}\sqrt{n}} \left|\left| \frac{\hat{S}_n}{\sqrt{n}} \right|\right|_2 + \frac{1}{k_n} \right).$$

The convergence (57) in Proposition 6.3 implies that $\left|\left|\hat{S}_n/\sqrt{n}\right|\right|_2 = \mathrm{O}(1)$. The trick here is to choose $k_n = \lfloor c\ln(n)\rfloor$ where $c < -1/(2\ln\lambda_-)$ which implies that $\lim_{n\to\infty}\lambda_-^{k_n}\sqrt{n}/k_n = \infty$, so that the upper bound on the righthanside of the above inequality is an $\mathrm{O}(1/k_n) = \mathrm{O}(1/\ln(n))$, proving the inequality (49). Finally, the convergence in probability of $\hat{N}_n$ comes from the a.s. convergence of $\bar{Y}_n$ towards C_1 and the convergence in probability of $e^{\hat{S}_n}$ (the latter stemming from the convergence in $\mathbb{L}^2$ of $\hat{S}_n$) towards λ_0, coupled with Relation (7). $\square$

Appendix A.10. Proof of Theorem 6.2

The increment theorem of function ψ_{k_n} at point $U(k_n) := (C_1, u_{k_n})$ gives

$$\hat{S}_n - \psi_{k_n}(U(k_n)) = \psi_{k_n}\left(U(k_n) + \frac{S_n(k_n)}{n}\right) - \psi_{k_n}(U(k_n)) = \nabla\psi_{k_n}\left(U(k_n) + c_n\frac{S_n(k_n)}{n}\right).\frac{S_n(k_n)}{n} \tag{A.147}$$

for some r.v. $c_n \in (0,1)$. Using again the increment theorem to $\nabla\psi_{k_n}$ at some point $U(k_n)$, we get, thanks to the upper bound (A.88) for $D^2\psi_{k_n}$ obtained in the upcoming Lemma Appendix A.10, that

$$\begin{aligned}\left|\left|k_n\lambda_0^{k_n}\left[\nabla\psi_{k_n}\left(U(k_n) + c_n\frac{S_n(k_n)}{n}\right) - \nabla\psi_{k_n}\left(U(k_n)\right)\right]\right|\right| &\\ \leq k_n\lambda_0^{k_n}\sup_{(a,b)\in\mathbb{R}^2}||D^2\psi_{k_n}(a,b)||.\left|\left|c_n\frac{S_n(k_n)}{n}\right|\right| &\\ \leq Kk_n\lambda_0^{k_n}\frac{1}{\lambda_-^{k_n}}\left|\left|c_n\frac{S_n(k_n)}{n}\right|\right| \leq Kk_n\lambda_0^{k_n}\frac{1}{\lambda_-^{k_n}\sqrt{n}}\left|\left|\frac{S_n(k_n)}{\sqrt{n}}\right|\right| &\end{aligned} \tag{A.148}$$

for some constant $K > 0$. Furthermore, for n large enough we have from (13) as well as Assumption (**A6**) that $|C_1^2 - u_{k_n}| = \lambda_0^{k_n}\left|\sum_{j=0}^{k_n}\lambda_0^{-j}\chi_j + \lambda_0(C_1^2 - C_2)\right| \geq \lambda_0^{k_n}K_m/2$, so that the definition (42) of ϖ_k entails that $\varpi_{k_n}(|C_1^2 - u_{k_n}|) = 1$ and $\varpi'_{k_n}(|C_1^2 - u_{k_n}|) = 0$. This in turn entails that $\nabla\psi_{k_n}(U(k_n)) = \left(2C_1/(k_n|C_1^2 - u_{k_n}|), -1/(k_n|C_1^2 - u_{k_n}|)\right)'$ (see the upcoming calculation of $\partial_a\psi_k$ in (A.93) and a similar calculation for $\partial_b\psi_k$ for all k in the proof of Lemma Appendix A.10 in Section Appendix A.7). Since k_n obviously satisfies (58), the central limit theorem (59) in Proposition 6.4 holds, and thanks to the expression (13) we thus have

$$k_n\lambda_0^{k_n}\nabla\psi_{k_n}(U(k_n))\frac{S_n(k_n)}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, V\mathfrak{S}V'\right), \quad n\to\infty \tag{A.149}$$

where $\mathfrak{S}$ is given by (53) and V is given by (52). Now, the choice of $k_n := \lfloor c\ln n\rfloor$ where $c < -1/(2\ln\lambda_-)$ implies that $\lim_{n\to\infty}\lambda_-^{k_n}\sqrt{n} = \infty$, and $||S_n(k_n)/\sqrt{n}||$ is bounded in $\mathbb{L}^2$ thanks to (57), so that (A.148) implies that

$$\begin{aligned}\left|\left|k_n\lambda_0^{k_n}\left[\nabla\psi_{k_n}\left(U(k_n) + c_n\frac{S_n(k_n)}{n}\right) - \nabla\psi_{k_n}\left(U(k_n)\right)\right].\frac{S_n(k_n)}{\sqrt{n}}\right|\right| &\\ \leq \left|\left|k_n\lambda_0^{k_n}\left[\nabla\psi_{k_n}\left(U(k_n) + c_n\frac{S_n(k_n)}{n}\right) - \nabla\psi_{k_n}\left(U(k_n)\right)\right]\right|\right|.\left|\left|\frac{S_n(k_n)}{\sqrt{n}}\right|\right| &\\ \leq Kk_n\lambda_0^{k_n}\frac{1}{\lambda_-^{k_n}\sqrt{n}}\left|\left|\frac{S_n(k_n)}{\sqrt{n}}\right|\right|^2 \xrightarrow{\mathbb{L}^1} 0 &\end{aligned}$$

as $n \to \infty$ (in the above inequalities, it is convenient to take a multiplicative matrix norm $||\cdot||$ which verifies $||Mv|| \leq ||M||.||v||$ for all matrix M and vector v). Coupled with (A.149), we then get the following asymptotic normality

$$k_n \lambda_0^{k_n} \nabla \psi_{k_n} \left(U(k_n) + c_n \frac{S_n(k_n)}{n} \right) . \frac{S_n(k_n)}{\sqrt{n}} \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, V \mathfrak{S} V'\right), \quad n \to \infty. \tag{A.150}$$

Having considered the righthandside in (A.147), we now study its lefthandside. Since we saw that $\varpi_{k_n}(|C_1^2 - u_{k_n}|) = 1$ for n large enough, we thus have that $\psi_{k_n}(U(k_n)) = k_n^{-1} \ln |C_1^2 - u_{k_n}|$ so that, from (13):

$$\begin{aligned} k_n \lambda_0^{k_n} [\psi_{k_n}(U(k_n)) - \ln \lambda_0] &= \lambda_0^{k_n} \ln \left| \sum_{j=0}^{k_n} \lambda_0^{-j} \chi_j + \lambda_0 (C_1^2 - C_2) \right| \\ &= \lambda_0^{k_n} \ln \left| \sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0 (C_1^2 - C_2) \right| + \lambda_0^{k_n} \ln \left| 1 - \frac{\sum_{j=k_n+1}^{\infty} \lambda_0^{-j} \chi_j}{\left| \sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0 (C_1^2 - C_2) \right|} \right|. \end{aligned} \tag{A.151}$$

The assumption $\nu_h = \mathrm{O}(\zeta^h)$ entails from the definition (5) and $\zeta < \lambda_- \leq \lambda_0$ that $\chi_j = \mathrm{O}(\zeta^j)$ so that $\sum_{j=k_n+1}^{\infty} \lambda_0^{-j} \chi_j = \mathrm{O}\left((\zeta/\lambda_0)^{k_n} \right)$, a quantity that tends to 0 as $n \to \infty$. We thus deduce from (A.151) and the classical expansion $\ln(1+u) = \mathrm{O}(u)$ as $u \to 0$ that

$$k_n \lambda_0^{k_n} [\psi_{k_n}(U(k_n)) - \ln \lambda_0] = \lambda_0^{k_n} \ln \left| \sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0 (C_1^2 - C_2) \right| + \mathrm{O}(\zeta^{k_n}), \quad n \to \infty. \tag{A.152}$$

Gathering (A.147) and (A.152) yields

$$\begin{aligned} \hat{S}_n = \ln \lambda_0 + \frac{1}{k_n} \ln \left| \sum_{j=0}^{\infty} \lambda_0^{-j} \chi_j + \lambda_0 (C_1^2 - C_2) \right| \\ + \frac{1}{\sqrt{n} k_n \lambda_0^{k_n}} \left[k_n \lambda_0^{k_n} \nabla \psi_{k_n} \left(U(k_n) + c_n \frac{S_n(k_n)}{n} \right) . \frac{S_n(k_n)}{\sqrt{n}} + \mathrm{O}(\sqrt{n} \zeta^{k_n}) \right]. \end{aligned}$$

Now, the assumption $c > -1/(2 \ln \zeta)$ implies that $\lim_{n \to \infty} \sqrt{n} \zeta^{k_n} = 0$, so that the expansion (50) holds with $Z_n := k_n \lambda_0^{k_n} \nabla \psi_{k_n} \left(U(k_n) + c_n S_n(k_n)/n \right) . S_n(k_n)/\sqrt{n} + \mathrm{O}(\sqrt{n} \zeta^{k_n})$, which converges in distribution towards $\mathcal{N}(0, \sigma)$ thanks to (A.150), with σ defined as (51). Also note that the assumption $c < -1/(2 \ln \lambda_-)$ implies that $Z_n / \sqrt{n} k_n \lambda_0^{k_n}$ is indeed as second asymptotic term, as indeed we have $\lim_{n \to \infty} \sqrt{n} \lambda_0^{k_n} = \infty$. □

Acknowledgements

The authors thank the anonymous referee and the Associate Editor for the very careful reading of the paper.

References

Akutowicz, E. J. (1957). On an explicit formula in linear least squares prediction. *Math. Scand.*, 5:261–266.

Al-Osh, M. A. and Alzaid, A. A. (1987). First-order integer-valued autoregressive (INAR(1)) process. *J. Time Ser. Anal.*, 8(3):261–275.

Anderson, T. W. (1971). *The statistical analysis of time series*. John Wiley & Sons, Inc., New York-London-Sydney.

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.

Barczy, M., Basrak, B., Kevei, P., Pap, G., and Planinić, H. (2021). Statistical inference of subcritical strongly stationary galton–watson processes with regularly varying immigration. *Stochastic Processes and their Applications*, 132:33–75.

Berk, K. N. (1974). Consistent autoregressive spectral estimates. *Ann. Statist.*, 2:489–502. Collection of articles dedicated to Jerzy Neyman on his 80th birthday.

Boubacar Mainassara, Y., Carbon, M., and Francq, C. (2012). Computing and estimating information matrices of weak ARMA models. *Comput. Statist. Data Anal.*, 56(2):345–361.

Boubacar Mainassara, Y. and Francq, C. (2011). Estimating structural VARMA models with uncorrelated but non-independent error terms. *J. Multivariate Anal.*, 102(3):496–505.

Brockwell, P. J. and Davis, R. A. (1991). *Time series: theory and methods*. Springer Series in Statistics. Springer-Verlag, New York, second edition.

Daley, D. J., Vere-Jones, D., et al. (2003). *An introduction to the theory of point processes: volume I: elementary theory and methods*. Springer.

Davidson, J. (1994). *Stochastic limit theory*. Advanced Texts in Econometrics. The Clarendon Press, Oxford University Press, New York. An introduction for econometricians.

Davydov, J. A. (1968). The convergence of distributions which are generated by stationary random processes. *Teor. Verojatnost. i Primenen.*, 13:730–737.

den Haan, W. J. and Levin, A. T. (1997). A practitioner's guide to robust covariance matrix estimation. In *Robust inference*, volume 15 of *Handbook of Statist.*, pages 299–342. North-Holland, Amsterdam.

Francq, C. and Zakoïan, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference*, 68(1):145–165.

Francq, C. and Zakoïan, J.-M. (2005). A central limit theorem for mixing triangular arrays of variables whose dependence is allowed to grow with the sample size. *Econometric Theory*, 21(6):1165–1171.

Francq, C. and Zakoïan, J.-M. (2007). HAC estimation and strong linearity testing in weak ARMA models. *J. Multivariate Anal.*, 98(1):114–144.

Francq, C. and Zakoïan, J.-M. (2019). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley.

Hawkes, A. G. (1971a). Point spectra of some mutually exciting point processes. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 33(3):438–443.

Hawkes, A. G. (1971b). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.

Hawkes, A. G. and Oakes, D. (1974). A cluster process representation of a self-exciting process. *J. Appl. Probability*, 11:493–503.

Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Ann. Probab.*, 12(1):141–153.

Ispány, M., Körmendi, K., and Pap, G. (2014). Asymptotic behavior of cls estimators for 2-type doubly symmetric critical galton–watson processes with immigration. *Bernoulli*, 20(4):2247–2277.

Kevei, P. and Wiandt, P. (2021). Moments of the stationary distribution of subcritical multitype Galton-Watson processes with immigration. *Statist. Probab. Lett.*, 173:Paper No. 109067, 6.

Kirchner, M. (2016). Hawkes and INAR(∞) processes. *Stochastic Process. Appl.*, 126(8):2494–2525.

Kirchner, M. (2017). An estimation procedure for the Hawkes process. *Quant. Finance*, 17(4):571–595.

Klimko, L. A. and Nelson, P. I. (1978). On conditional least squares estimation for stochastic processes. *Ann. Statist.*, 6(3):629–642.

Körmendi, K. and Pap, G. (2018). Statistical inference of 2-type critical galton–watson processes with immigration. *Statistical Inference for Stochastic Processes*, 21:169–190.

McKenzie, E. (2003). Ch. 16. discrete variate time series. In *Stochastic Processes: Modelling and Simulation*, volume 21 of *Handbook of Statistics*, pages 573–606. Elsevier.

Newey, W. K. and West, K. D. (1987). A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55(3):703–708.

Romano, J. P. and Thombs, L. A. (1996). Inference for autocorrelations under weak assumptions. *Journal of the American Statistical Association*, 91:590–600.

Taniguchi, M. and Puri, M. L. (1996). Valid Edgeworth expansions of M-estimators in regression models with weakly dependent residuals. *Econometric Theory*, 12(2):331–346.

Venkataraman, K. (1982). A time series approach to the study of the simple subcritical galton–watson process with immigration. *Advances in Applied Probability*, 14(1):1–20.

Wei, C. Z. and Winnicki, J. (1990). Estimation of the means in the branching process with immigration. *Ann. Statist.*, 18(4):1757–1773.