



**HAL**  
open science

# Generalized least squares estimation of fractional autoregressive models

Youssef Esstafa

► **To cite this version:**

Youssef Esstafa. Generalized least squares estimation of fractional autoregressive models. 2023. hal-04300303

**HAL Id: hal-04300303**

**<https://hal.science/hal-04300303>**

Preprint submitted on 22 Nov 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Generalized least squares estimation of fractional autoregressive models

Youssef Esstafa

*Le Mans Université,  
Laboratoire Manceau de Mathématiques,  
Avenue Olivier Messiaen, 72085 Le Mans Cedex 9, France.*

*e-mail: [Youssef.Esstafa@univ-lemans.fr](mailto:Youssef.Esstafa@univ-lemans.fr)*

**Abstract:** In this paper we derive the asymptotic properties of the generalized least squares estimator (GLSE) of autoregressive models endowed with fractional Gaussian noise (the so-called fractional autoregressive models). We establish the consistency and the asymptotic normality of the GLSE. Some simulation studies are presented to corroborate our theoretical work.

**AMS 2000 subject classifications:** Primary 62M10; secondary 91B84.

**Keywords and phrases:** Fractional Brownian motion; Fractional Gaussian noise; FrAR models; Generalized least squares estimator; Consistency; Asymptotic normality.

## 1. Introduction

The error term in the classical autoregressive models is usually assumed to be a strong white noise (namely an independent and identically distributed (iid for short) sequence of random variables with mean 0 and common variance), a martingale difference or a stationary process with a bounded spectral density (see [Anderson and Taylor \[1976\]](#), [Anderson and Taylor \[1979\]](#), [Lai and Robbins \[1977\]](#), [Lai et al. \[1978\]](#), [Lai et al. \[1979\]](#) and [Solo \[1981\]](#)). In the late 1990s, Francq, C. and Zakoïan, J.-M. have released the assumption of the independence on the noise and have considered the problem of estimating ARMA models with uncorrelated but non-independent error term, the so-called weak ARMA models (see [Francq and Zakoïan \[1998\]](#)). This extension allowed ARMA models to be less restrictive and to cover a large class of general nonlinear models. Nevertheless, the uncorrelatedness of the noise assumed in these classes of models considerably restricts their application to short-memory stochastic processes.

Yet, in many scientific disciplines and applied fields such as hydrology, climatology, economics, finance, to name a few, time series exhibit long-range dependence (see for instance [Granger and Joyeux \[1980\]](#), [Fox and Taqqu \[1986\]](#), [Dahlhaus \[1989\]](#), [Hosking \[1981\]](#), [Beran et al. \[2013\]](#), [Palma \[2007\]](#), among others), that is dependence between distant observations is so strong that classical or weak ARMA models are unable to identify and then to take into consideration.

The fractional Gaussian noise (see [Mandelbrot and Van Ness \[1968\]](#)) and the fractional autoregressive integrated moving average (FARIMA, for short) models (see [Granger and Joyeux \[1980\]](#) and [Hosking \[1981\]](#)) are widely used to model the long memory phenomenon. We shall consider in this work the problem of generalized least squares estimation of autoregressive models endowed with fractional Gaussian noise. More explicitly, we shall study the asymptotic properties of the generalized least squares estimator of the parameters of the centered stationary fractional autoregressive

process of order 1 (FrAR(1))  $(X_t)_{t \in \mathbb{Z}}$  defined, for all  $t \in \mathbb{Z}$ , by

$$X_t = a_0 X_{t-1} + \epsilon_t^H, \quad (1)$$

where  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is a fractional Gaussian noise and where the autoregressive parameter  $a_0$  and the Hurst index or Hurst parameter  $H$  associated with the fractional Gaussian noise  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  are assumed to satisfy the following standard assumption:

**(A0):**  $a_0 \in (-1, 1)$  and  $H \in (0, 1)$ .

The condition  $a_0 \in (-1, 1)$  ensures the causality and then the second-order stationarity of the process  $(X_t)_{t \in \mathbb{Z}}$ . In fact, for AR( $p$ ) models a causal weak-sense stationary solution is obtained once the roots of the corresponding autoregressive polynomial lie outside the closed unit disk.

The process  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is the increment process of the fractional Brownian motion  $(B_H(t))_{t \in \mathbb{R}}$  of Hurst index  $H$ , namely, for any  $t \in \mathbb{Z}$ , it holds

$$\epsilon_t^H = B_H(t+1) - B_H(t).$$

Based on the definition and the properties of the fractional Brownian motion  $(B_H(t))_{t \in \mathbb{R}}$ , the process  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is the unique continuous stationary Gaussian centered process with autocovariance function defined for  $k \in \mathbb{Z}$  by

$$\gamma_{\epsilon^H}(k) = \frac{1}{2} \left( |k-1|^{2H} - 2|k|^{2H} + |k+1|^{2H} \right). \quad (2)$$

A Taylor expansion of the function  $\ell : x \mapsto (1-x)^{2H} - 2 + (1+x)^{2H}$  around 0 implies that for large  $k$ ,

$$\gamma_{\epsilon^H}(k) = \frac{1}{2} k^{2H} \ell(1/k) = H(2H-1)k^{2H-2} + o(k^{2H-2}).$$

The last result implies that  $\gamma_{\epsilon^H}(k) \rightarrow 0$  as  $k \rightarrow \infty$ . So the process  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is mixing (see [Ito \[1944\]](#)). Mixing is a stronger property than ergodicity, this result is known as the Khinchin's theorem (see [Khinchin \[1949\]](#)). Hence  $\lim_{k \rightarrow \infty} \gamma_{\epsilon^H}(k) = 0$  implies that the process  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is ergodic.

When  $1/2 < H < 1$ , the series of the autocovariances of  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is absolutely divergent and the process  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  exhibits long-range dependence.

In the literature, many approaches have been proposed to estimate the parameters of strongly dependent Gaussian processes. We cite for example the R/S method, periodogram estimation and maximum likelihood estimation. [Mandelbrot \[1975\]](#) and [Mandelbrot and Taqqu \[1979\]](#) have established the theoretical properties of the R/S estimates. [Mohr \[1981\]](#), [Graf \[1980\]](#), [Geweke and Porter-Hudak \[1983\]](#), [Fox and Taqqu \[1986\]](#) and [Dahlhaus \[1989\]](#) have considered the problem of periodogram estimation. [McLeod and Hipel \[1978\]](#) have discussed computational considerations involved in the application of maximum likelihood estimation. [Sweeting \[1980\]](#) has proposed assumptions to obtain the consistency and the asymptotic normality of maximum likelihood estimator. [Dahlhaus \[1989\]](#) has established the asymptotic properties of the exact maximum likelihood estimation of the parameters of strongly dependent Gaussian processes.

In the particular case of autoregressive models endowed with a strongly dependent noise, [Brouste et al. \[2014\]](#) have proposed a new approach based on a transformation of the initial model to establish the asymptotic properties of the maximum likelihood estimator. See also [Yajima \[1988\]](#) for

the case of a regression problem with nonstochastic regressors.

We study in this paper the asymptotic properties of the generalized least squares estimator of autoregressive models endowed with fractional Gaussian noise. We propose a simple and natural transformation of the initial model and we establish the convergence and asymptotic normality of the generalized least squares estimator based on this transformation. These properties are obtained using only the standard hypothesis **(A0)**. We propose simple proofs of the asymptotic properties of the generalized least squares estimator of the autoregressive parameters. We use the closed form of the estimator to provide very intuitive proofs. We also extend the support of the memory parameter by considering that  $0 < H < 1$ . Contrary therefore to [Dahlhaus \[1989\]](#) and [Brouste et al. \[2014\]](#), we take into consideration not only long memory dynamics but also short memory fluctuations. We also show numerically that our estimator is preferable (in terms of quadratic risk) to the Whittle estimator proposed by [Dahlhaus \[1989\]](#) when the sample size is relatively moderate.

The paper is organized as follows. In [Section 2](#), we introduce our model and we study the asymptotic properties of our estimator. In [Section 3](#), we show that our results obtained in the framework of fractional autoregressive models of order 1 can easily be extended to the case of autoregressive models of order  $p \geq 1$  induced by fractional Gaussian noise. We present in [Section 4](#) some simulation studies to illustrate our theoretical results. We also offer comparisons in terms of the quadratic risk of our estimator with that of Whittle (see for example [Dahlhaus \[1989\]](#)). Proofs of our two main results are set out to [Section 5](#). All figures are collected in the last section of the paper.

In all this work,  $x^T$  designates the transpose of the vector  $x$ . We shall also use the matrix norm defined by  $\|A\| = \sup_{\|x\|_{\mathbb{R}^{k_2}} \leq 1} \|Ax\|_{\mathbb{R}^{k_1}} = \rho^{1/2}(A^T A)$ , when  $A$  is a  $\mathbb{R}^{k_1 \times k_2}$  matrix,  $\|x\|_{\mathbb{R}^{k_2}}^2 = x^T x$  is the Euclidean norm of the vector  $x \in \mathbb{R}^{k_2}$ , and  $\rho(\cdot)$  denotes the spectral radius.

## 2. Generalized least squares estimation of FrAR models of order 1

In this section we present the parametrization that is used in the sequel and we state the asymptotic properties of the GLSE of FrAR(1) when the Hurst exponent  $H$  is known. An application of these asymptotic properties can be found in [Ben Hariz et al. \[2021\]](#) where a two steps procedure is performed when  $H$  is unknown.

### 2.1. Statement of the problem and notations

First, note that the ordinary least squares estimator of the parameter  $a_0$  in [\(1\)](#) is biased. To see that, let  $\hat{a}_n^{(OLS)}$  be the standard least squares estimator of  $a_0$ . It is well known that

$$\hat{a}_n^{(OLS)} = \frac{\sum_{t=1}^n X_t X_{t-1}}{\sum_{t=1}^n X_{t-1}^2}.$$

Thanks to [\(1\)](#), one has

$$\hat{a}_n^{(OLS)} - a_0 = \frac{\sum_{t=1}^n \epsilon_t^H X_{t-1}}{\sum_{t=1}^n X_{t-1}^2}.$$

The ergodicity of  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  and the stationarity of  $(X_t)_{t \in \mathbb{Z}}$  imply that

$$\hat{a}_n^{(OLS)} - a_0 \xrightarrow[n \rightarrow \infty]{\text{a.s.}} \frac{\mathbb{E}[\epsilon_2^H X_1]}{\text{Var}(X_1)}.$$

From (2) and (16), one can easily obtain that

$$\mathbb{E} \left[ \epsilon_2^H X_1 \right] = \frac{a_0 - 2}{2} + \frac{(a_0 - 1)^2}{2} \sum_{j \geq 2} a_0^{j-2} j^{2H}.$$

Since the function  $H \in (0, 1) \rightarrow \sum_{j \geq 2} a_0^{j-2} j^{2H}$  is bijective, we deduce that  $\mathbb{E}[\epsilon_2^H X_1] = 0$  if and only if  $H = 1/2$ . This leads to the standard case of an autoregressive model where the noise is an iid sequence with standard normal distribution.

Due to the bias problem of the ordinary least squares estimator (problem caused by the correlation of the noise), we consider here the generalized least squares estimation procedure. This technique consists in transforming Model (1) to another autoregressive model, with the same unknown autoregressive parameter  $a_0$ , endowed with a standard white noise and consider then the ordinary least squares estimation on the new model.

From (1), one can write

$$\begin{pmatrix} X_2 \\ \vdots \\ X_{n+1} \end{pmatrix} = a_0 \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} + \begin{pmatrix} \epsilon_2^H \\ \vdots \\ \epsilon_{n+1}^H \end{pmatrix}. \quad (3)$$

The vector  $(\epsilon_2^H, \dots, \epsilon_{n+1}^H)^T$  is a centered normal random vector with covariance matrix  $\Omega_{n,H} := [\gamma_{\epsilon^H}(j-i)]_{1 \leq i, j \leq n} = [\text{Cov}(\epsilon_i^H, \epsilon_j^H)]_{1 \leq i, j \leq n}$ .

**Remark 1.** The covariance matrix  $\Omega_{n,H}$  depends only on the Hurst parameter  $H$ . As a symmetric real matrix,  $\Omega_{n,H}$  can be diagonalized by an orthogonal matrix. More explicitly, there exists a real orthogonal matrix  $P_{n,H}$  such that  $D_{n,H} = P_{n,H}^T \Omega_{n,H} P_{n,H}$  is a diagonal matrix. Since  $\Omega_{n,H}$  is positive definite (see Proposition 1 hereafter), we have  $\Omega_{n,H}^{1/2} := P_{n,H} D_{n,H}^{1/2} P_{n,H}^T$  where  $D_{n,H}^{1/2} = \text{diag}((\lambda_{1,H}^{(n)})^{1/2}, \dots, (\lambda_{n,H}^{(n)})^{1/2})$  and the  $\lambda_{i,H}^{(n)}$  are the eigenvalues of  $\Omega_{n,H}$ . The matrix  $\Omega_{n,H}^{1/2}$  is the square root of  $\Omega_{n,H}$  since it satisfies  $\Omega_{n,H}^{1/2} \Omega_{n,H}^{1/2} = \Omega_{n,H}$ .

In view of this last remark,  $\Omega_{n,H}^{-1/2}(\epsilon_2^H, \dots, \epsilon_{n+1}^H)^T$  is a standard normal random vector. Thus, a rather natural and simple transformation that can be proposed is the following:

$$Y_{n,H} = a_0 Z_{n,H} + U_{n,H}, \quad (4)$$

where

$$Y_{n,H} = \Omega_{n,H}^{-1/2} \begin{pmatrix} X_2 \\ \vdots \\ X_{n+1} \end{pmatrix}, \quad Z_{n,H} = \Omega_{n,H}^{-1/2} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \quad \text{and} \quad U_{n,H} = \Omega_{n,H}^{-1/2} \begin{pmatrix} \epsilon_2^H \\ \vdots \\ \epsilon_{n+1}^H \end{pmatrix}. \quad (5)$$

The random variable  $\hat{a}_n$  is called generalized least squares estimator if it satisfies, almost surely,

$$\hat{a}_n = \underset{a \in (-1,1)}{\text{argmin}} \frac{1}{n} \|U_{n,H}(a)\|_{\mathbb{R}^n}^2, \quad (6)$$

where

$$U_{n,H}(a) = \Omega_{n,H}^{-1/2} \begin{pmatrix} e_2^H(a) \\ \vdots \\ e_{n+1}^H(a) \end{pmatrix} \quad (7)$$

and

$$e_t^H(a) = X_t - aX_{t-1} \quad (8)$$

with  $e_t^H(a) = X_t = 0$  if  $t \leq 0$ . Note that  $U_{n,H}(a_0) = U_{n,H}$ .

The resolution of the optimization problem presented in (6) leads to an explicit expression of  $\hat{a}_n$ . One can easily prove that

$$\hat{a}_n = \frac{Z_{n,H}^T Y_{n,H}}{\|Z_{n,H}\|_{\mathbb{R}^n}^2}. \quad (9)$$

Moreover, in view of this last equality and Equation (4), one can deduce that

$$\hat{a}_n - a_0 = \frac{Z_{n,H}^T U_{n,H}(a_0)}{\|Z_{n,H}\|_{\mathbb{R}^n}^2}. \quad (10)$$

## 2.2. Asymptotic properties

The asymptotic properties of the generalized least squares estimator of the fractional autoregressive model of order 1 are stated in the following two theorems.

**Theorem 1.** (Consistency). *We assume that  $(X_t)_{t \in \mathbb{Z}}$  satisfies (1). Under (A0), we have*

$$\hat{a}_n \xrightarrow[n \rightarrow \infty]{\mathbb{P}} a_0,$$

where  $(\hat{a}_n)_{n \geq 1}$  is a sequence of generalized least squares estimators as defined in (9).

The proof of this theorem is given in Subsection 5.2.

**Theorem 2.** (Asymptotic normality). *We assume that  $(X_t)_{t \in \mathbb{Z}}$  satisfies (1). Under (A0), the sequence  $(\sqrt{n}(\hat{a}_n - a_0))_{n \geq 1}$  has a limiting centered normal distribution with variance  $1 - a_0^2$ .*

The proof of this theorem is set out to Subsection 5.3.

## 3. Generalized least squares estimation of FrAR models of order $p$

The results obtained in the previous section can be extended to fractional autoregressive models of order  $p$ . To be more precise, suppose that the process  $(X_t)_{t \in \mathbb{Z}}$  admits an autoregressive representation of the form

$$X_t = a_{01}X_{t-1} + \cdots + a_{0p}X_{t-p} + \epsilon_t^H, \quad (11)$$

where  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is a fractional Gaussian noise with  $0 < H < 1$ . To ensure the stability of the process  $(X_t)_{t \in \mathbb{Z}}$  in this case, we suppose that its reverse characteristic polynomial has no root in and on the complex unit circle, i.e. we assume that

$$1 - a_{01}z - \cdots - a_{0p}z^p \neq 0 \text{ for } |z| \leq 1. \quad (12)$$

In view of (11), it can readily be shown that

$$\begin{pmatrix} X_2 \\ \vdots \\ X_{n+1} \end{pmatrix} = \begin{pmatrix} X_1 & \cdots & X_{2-p} \\ \vdots & & \vdots \\ X_n & \cdots & X_{n+1-p} \end{pmatrix} \begin{pmatrix} a_{01} \\ \vdots \\ a_{0p} \end{pmatrix} + \begin{pmatrix} \epsilon_2^H \\ \vdots \\ \epsilon_{n+1}^H \end{pmatrix}. \quad (13)$$

We readopt the same transformation as the one used in the previous section to obtain a vector representation similar to the one introduced in (4). Multiplying all the terms of Equation (13) by the appropriate matrix  $\Omega_{n,H}^{-1/2}$ , one has

$$Y_{n,H} = \underline{\mathbf{Z}}_{n,H} \underline{\mathbf{A}}_0 + U_{n,H}(\underline{\mathbf{A}}_0),$$

where  $Y_{n,H}$  and  $U_{n,H}(\cdot)$  are defined in (5) and (7) and where

$$\underline{\mathbf{Z}}_{n,H} = \Omega_{n,H}^{-1/2} \begin{pmatrix} X_1 & \cdots & X_{2-p} \\ \vdots & & \vdots \\ X_n & \cdots & X_{n+1-p} \end{pmatrix} \text{ and } \underline{\mathbf{A}}_0 = \begin{pmatrix} a_{01} \\ \vdots \\ a_{0p} \end{pmatrix}.$$

Similarly as before, we again show that the generalized least squares estimator  $\hat{\underline{\mathbf{A}}}_n$  of  $\underline{\mathbf{A}}_0$  takes a closed form and it is explicitly given by:

$$\begin{aligned} \hat{\underline{\mathbf{A}}}_n &= \underset{\underline{\mathbf{A}}}{\operatorname{argmin}} \frac{1}{n} \|U_{n,H}(\underline{\mathbf{A}})\|_{\mathbb{R}^n}^2 \\ &= \left( \underline{\mathbf{Z}}_{n,H}^T \underline{\mathbf{Z}}_{n,H} \right)^{-1} \left( \underline{\mathbf{Z}}_{n,H}^T Y_{n,H} \right) \\ &= \underline{\mathbf{A}}_0 + \left( \underline{\mathbf{Z}}_{n,H}^T \underline{\mathbf{Z}}_{n,H} \right)^{-1} \left( \underline{\mathbf{Z}}_{n,H}^T U_{n,H}(\underline{\mathbf{A}}_0) \right), \end{aligned}$$

where the minimum is taken over all vectors  $\underline{\mathbf{A}} = (a_1, \dots, a_p)^T$  of  $\mathbb{R}^{p-1} \times \mathbb{R}^*$  such that (12) is satisfied.

The following theorem gives the asymptotic properties of the generalized least squares estimator of autoregressive models of order  $p$  induced by fractional Gaussian noise. The proof of this theorem is similar to the one of Theorems 1 and 2 detailed in Section 5.

**Theorem 3.** *We assume that  $(X_t)_{t \in \mathbb{Z}}$  satisfies (11). Suppose that  $0 < H < 1$  and that  $\underline{\mathbf{A}}_0 = (a_{01}, \dots, a_{0p})^T$  is such that Condition (12) is verified. Then, the sequence  $(\sqrt{n}(\hat{\underline{\mathbf{A}}}_n - \underline{\mathbf{A}}_0))_{n \geq 1}$  has a limiting centered normal distribution with variance  $\Gamma^{-1}(\underline{\mathbf{A}}_0)$ , where  $\Gamma(\underline{\mathbf{A}}_0)$  is the limit in probability of  $\underline{\mathbf{Z}}_{n,H}^T \underline{\mathbf{Z}}_{n,H} / n$ .*

#### 4. Numerical illustrations

In this section, we investigate the finite sample properties of the asymptotic results that we introduced in this work. For that, we use Monte Carlo experiments. The numerical illustrations of this section are made with the open source statistical software R (see R Development Core Team, 2017) or (see <http://cran.r-project.org/>).

#### 4.1. Simulation studies

We first study numerically the behavior of the generalized least squares estimator for fractional autoregressive models of the form

$$X_t = aX_{t-1} + \epsilon_t^H, \quad (14)$$

where the unknown autoregressive parameter is taken as  $a = 0.3$ , and where  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is a fractional Gaussian noise with Hurst index  $H = 0.7$ .

We simulate  $N = 2,000$  independent trajectories of sizes  $n = 30$ ,  $n = 500$  and  $n = 2,000$  of Model (14).

Figures 1, 2 and 3 illustrate the distribution of the generalized least squares estimator  $\hat{a}_n$  of the autoregressive parameter  $a$ . We can notice that these numerical results are consistent with the convergence established in Theorem 2.

In Figures 4, 5 and 6, we corroborate the convergence of the estimator  $1 - \hat{a}_n^2$  of the asymptotic variance of the generalized least squares estimator  $\hat{a}_n$  given in Theorem 2. We can remark that even for a small sample size the estimator is very close to the true value of the parameter. This suggests that our estimator is not greedy in terms of sample size. This phenomenon is confirmed by Figures 8-12. For different values of  $n$  and for this same set of parameters (*i.e.*  $a = 0.3$  and  $H = 0.7$ ), we compare in Figure 7 the quadratic risk of the generalized least squares estimator  $\hat{a}_n$  (in black) with that of the Whittle estimator considered for example by Dahlhaus [1989] (in green) based on  $N = 2,000$  independent replications for each value of  $n$ . It is clear in this case that when the sample size is relatively small our estimator makes less error than the Whittle estimator. It is also worth noting that this closed form of our estimator makes it preferable even when the sample size is large.

To have a more global view on the behavior of our estimator when the sample size is relatively small, we consider the cases where  $a = -0.9, -0.5, 0, 0.5, 0.9$  in Model (14). For each value of  $a$ , we vary  $H$  over the interval  $[0.01, 0.99]$  with a step of 0.01. For each value of the autoregressive parameter  $a$  and the memory parameter  $H$ , we simulate  $N = 1,000$  independent trajectories of (14). We compare in Figures 8-12, the logarithm of the quadratic risk of the generalized least squares estimator  $\hat{a}_n$  (in black) with that of the Whittle estimator (in green) of  $N = 1,000$  replications of Model (14) with different values of  $a$  and  $H$ . In each figure, four sample sizes were considered. From left to right and top to bottom, the size takes the values  $n = 50, 100, 150, 200$ . Clearly, for the different values of  $a$  and  $n$ , our estimator shows a certain stability unlike that of Whittle where the quadratic risk is characterized by a certain instability especially when  $H$  is close to 0 and 1. This observation highlights the fact that the generalized least squares estimator  $\hat{a}_n$  gives very satisfactory results in the different memory scenarios presented by the process. Note that for all the illustrations given in Figures 8-12,  $H$  sweeps the entire memory interval  $(0, 1)$ . The Whittle estimator seems to be less precise for several values of  $a$  when the process is short memory (*i.e.* when  $H \in (0, 1/2]$ ).

## 5. Proofs

In all our proofs,  $K$  is a positive constant that may vary from line to line.



### 5.1. Preliminary results

In this subsection, we state some results that will be useful for the proofs of the asymptotic properties of the generalized least squares estimator introduced in Section 2. We use **(A0)** to write  $Z_{n,H}$  as a function of  $U_{n,H}(a_0)$ . This expression will provide very intuitive proofs of the asymptotic properties of the GLSE.

We also give the asymptotic behavior of the components of the inverse variance matrix of the fractional Gaussian noise  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  based on the behavior in a neighbourhood of zero of the spectral density of this process.

#### 5.1.1. The expression of $Z_{n,H}$ as a function of $U_{n,H}(a_0)$

We recall that, when  $|a| < 1$ , the polynomial  $a(z) = 1 - az$  is invertible and we have

$$a^{-1}(z) = \sum_{j \geq 0} a^j z^j. \quad (15)$$

This implies that under **(A0)** one may write the process  $(X_t)_{t \in \mathbb{Z}}$  as a linear combination of the infinite past  $(\epsilon_{t-j}^H)_{j \geq 0}$ . More explicitly, we have

$$X_t = (1 - a_0 L)^{-1} \epsilon_t^H = \sum_{j \geq 0} a_0^j \epsilon_{t-j}^H, \quad (16)$$

where  $L$  stands for the back-shift operator.

In view of the definition of the process  $(e_t^H(a))_{t \in \mathbb{Z}}$  introduced in (8) and (15), one can express  $X_t$  when  $1 \leq t \leq n+1$  as a function of the finite past  $(e_{t-j}^H(a_0))_{0 \leq j \leq t-1}$ . Indeed, Equation (8) and Assumption **(A0)** imply that

$$X_t = (1 - a_0 L)^{-1} e_t^H(a_0) = \sum_{j=0}^{t-1} a_0^j e_{t-j}^H(a_0), \quad \forall t \in \{1, \dots, n+1\}. \quad (17)$$

From (17), we deduce that the vector  $(X_1, \dots, X_n)^T$  can be rewritten in the form

$$\begin{aligned} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} &= \begin{pmatrix} \sum_{j=0}^{1-1} a_0^j e_{1-j}^H(a_0) \\ \sum_{j=0}^{2-1} a_0^j e_{2-j}^H(a_0) \\ \vdots \\ \sum_{j=0}^{n-1} a_0^j e_{n-j}^H(a_0) \end{pmatrix} = \begin{pmatrix} e_1^H(a_0) \\ e_2^H(a_0) + a_0 e_1^H(a_0) \\ \vdots \\ e_n^H(a_0) + \dots + a_0^{n-1} e_1^H(a_0) \end{pmatrix} \\ &= \begin{pmatrix} e_1^H(a_0) \\ e_2^H(a_0) \\ \vdots \\ e_n^H(a_0) \end{pmatrix} + a_0 \begin{pmatrix} e_0^H(a_0) \\ e_1^H(a_0) \\ \vdots \\ e_{n-1}^H(a_0) \end{pmatrix} + \dots + a_0^{n-1} \begin{pmatrix} e_{2-n}^H(a_0) \\ e_{3-n}^H(a_0) \\ \vdots \\ e_1^H(a_0) \end{pmatrix} \\ &= \sum_{j=0}^{n-1} a_0^j \begin{pmatrix} e_{1-j}^H(a_0) \\ e_{2-j}^H(a_0) \\ \vdots \\ e_{n-j}^H(a_0) \end{pmatrix} = \sum_{j=0}^{n-1} a_0^j \left[ L^{j+1} \begin{pmatrix} e_2^H(a_0) \\ e_3^H(a_0) \\ \vdots \\ e_{n+1}^H(a_0) \end{pmatrix} \right], \end{aligned} \quad (18)$$

where the shift operator  $L^{j+1}$  acts on all the components of the vector  $(e_2^H(a_0), \dots, e_{n+1}^H(a_0))^T$ , i.e.

$$L^{j+1} \begin{pmatrix} e_2^H(a_0) \\ \vdots \\ e_{n+1}^H(a_0) \end{pmatrix} = \begin{pmatrix} e_{1-j}^H(a_0) \\ \vdots \\ e_{n-j}^H(a_0) \end{pmatrix}.$$

In view of (18), it follows that the vector  $Z_{n,H}$  defined in (7) can be expressed as a function of  $U_{n,H}(a_0)$ :

$$Z_{n,H} = \Omega_{n,H}^{-1/2} \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix} = \sum_{j=0}^{n-1} a_0^j \left[ L^{j+1} \Omega_{n,H}^{-1/2} \begin{pmatrix} e_2^H(a_0) \\ e_3^H(a_0) \\ \vdots \\ e_{n+1}^H(a_0) \end{pmatrix} \right] = \sum_{j=1}^n a_0^{j-1} (L^j U_{n,H}(a_0)). \quad (19)$$

### 5.1.2. Asymptotic properties of the components of the inverse covariance matrix of $(\epsilon_t^H)_{1 \leq t \leq n}$

By Fox and Taqqu [1986], Sinai [1976], the spectral density  $f_{\epsilon^H}(\cdot)$  of the fractional Gaussian noise  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  is defined for all  $\omega \in [-\pi, \pi]$  by:

$$\begin{aligned} f_{\epsilon^H}(\omega) &= \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma_{\epsilon^H}(h) e^{-i\omega h} \\ &= C_H (1 - \cos(\omega)) \sum_{j \in \mathbb{Z}} \frac{1}{|\omega + 2j\pi|^{1+2H}} \\ &= C_H (1 - \cos(\omega)) |\omega|^{-1-2H} + C_H (1 - \cos(\omega)) \sum_{\substack{j \in \mathbb{Z} \\ j \neq 0}} \frac{1}{|\omega + 2j\pi|^{1+2H}}, \end{aligned}$$

where  $C_H = 2\Gamma(2H+1)\sin(\pi H)$  and  $\Gamma(\cdot)$  denotes the Gamma function.

As  $\omega \rightarrow 0$ , we have

$$f_{\epsilon^H}(\omega) \sim \frac{C_H}{2} |\omega|^{1-2H}. \quad (20)$$

**Proposition 1.** Under Assumption (A0), the following three assertions hold true.

1. The covariance matrix  $\Omega_{n,H}$  of the centered Gaussian random vector  $(\epsilon_1^H, \dots, \epsilon_n^H)^T$  is a symmetric Toeplitz positive definite matrix.
2. For large  $k$ , the behavior of the autocovariance  $\gamma_{\epsilon^H}(k)$  is given by:

$$(\Omega_{n,H})_{1,k} = \gamma_{\epsilon^H}(k) = H(2H-1)k^{2H-2} + o(k^{2H-2}). \quad (21)$$

3. The inverse autocovariances of  $(\epsilon_t^H)_{1 \leq t \leq n}$  satisfy

$$\left( \Omega_{n,H}^{-1} \right)_{j,k} \sim \left( \frac{1}{k-j} \right)^{2H} \frac{\Gamma(2H)e^{i\pi H}}{C_H \pi^2} \text{ when } |k-j| \rightarrow \infty. \quad (22)$$

*Proof.* For the first point, observe that  $\Omega_{n,H}$  can be considered as a self-adjoint endomorphism of the vector space  $\mathbb{C}^n$ . Note now that for any  $C \in \mathbb{C}^n \setminus \{0_{\mathbb{C}^n}\}$ , the spectral representation of  $\Omega_{n,H}$  implies that

$$C^T \Omega_{n,H} C = \sum_{j,k=1}^n C_k (\Omega_{n,H})_{k,j} C_j = \sum_{j,k=1}^n C_k \left( \int_{-\pi}^{\pi} f_{\epsilon^H}(\omega) e^{i(j-k)\omega} d\omega \right) C_j = \int_{-\pi}^{\pi} \left| \sum_{j=1}^n C_j e^{ij\omega} \right|^2 f_{\epsilon^H}(\omega) d\omega.$$

The function  $\omega \mapsto \sum_{j=1}^n C_j e^{ij\omega}$  is an analytic function on the disc and is not identically zero. Since  $f_{\epsilon^H}(\omega) > 0$  for  $\omega \neq 0$  the last integral is positive. Thus, we conclude that the symmetric Toeplitz matrix  $\Omega_{n,H}$  is positive definite.

In view of (2), a Taylor expansion of the function  $\ell : x \mapsto (1-x)^{2H} - 2 + (1+x)^{2H}$  around 0 implies that for large  $k$ ,

$$(\Omega_{n,H})_{1,k} = \gamma_{\epsilon^H}(k) = \frac{1}{2} k^{2H} \ell(1/k) = H(2H-1)k^{2H-2} + o(k^{2H-2}).$$

This gives the second point of the proposition.

The components of the inverse matrix of  $\Omega_{n,H}$  can also be expressed as a function of the spectral density of  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  (see Fox and Taqqu [1986]). More precisely, we have for all  $j, k = 1, \dots, n$ ,

$$\left(\Omega_{n,H}^{-1}\right)_{j,k} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{1}{f_{\epsilon^H}(\omega)} e^{i(k-j)\omega} d\omega. \quad (23)$$

Using the parity of the function  $f_{\epsilon^H}(\cdot)$ , observe that when  $j = k$  we have

$$\left(\Omega_{n,H}^{-1}\right)_{j,j} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{1}{f_{\epsilon^H}(\omega)} d\omega = \frac{1}{2\pi^2} \int_0^{\pi} \frac{1}{f_{\epsilon^H}(\omega)} d\omega. \quad (24)$$

From equivalence in (20), one has, when  $\omega \rightarrow 0$ ,

$$\frac{1}{f_{\epsilon^H}(\omega)} = \frac{2}{C_H} |\omega|^{2H-1} + o\left(\frac{2}{C_H} |\omega|^{2H-1}\right).$$

This implies that for  $\kappa > 0$  there exists  $\delta_\kappa > 0$  such that for any  $\omega \in ]-\delta_\kappa, \delta_\kappa[$  we have

$$(1-\kappa) \frac{2}{C_H} |\omega|^{2H-1} \leq \frac{1}{f_{\epsilon^H}(\omega)} \leq (1+\kappa) \frac{2}{C_H} |\omega|^{2H-1}. \quad (25)$$

Thanks to (24), (25) and the boundedness of the function  $\omega \rightarrow 1/f_{\epsilon^H}(\omega)$  on  $]\delta_\kappa, \pi]$ , one can show that there exists a positive constant  $K$  such that for any  $j = 1, \dots, n$ ,

$$\begin{aligned} \left| \left(\Omega_{n,H}^{-1}\right)_{j,j} \right| &\leq \frac{1+\kappa}{C_H \pi^2} \int_0^{\delta_\kappa} \omega^{2H-1} d\omega + \frac{1}{2\pi^2} \int_{\delta_\kappa}^{\pi} \frac{1}{f_{\epsilon^H}(\omega)} d\omega \\ &\leq \frac{\delta_\kappa^{2H}(1+\kappa)}{2HC_H \pi^2} + \frac{\pi - \delta_\kappa}{2\pi^2} \sup_{\omega \in ]\delta_\kappa, \pi]} \frac{1}{f_{\epsilon^H}(\omega)} \\ &\leq K. \end{aligned} \quad (26)$$

When  $j \neq k$  (suppose without loss of generality that  $k > j$ ), one can use Lemma 6.3.2 (page 442) of Ablowitz and Fokas [2003] which is an analog version of Watson's Lemma (see for instance Watson [1918] for the original proof or Miller [2006] for a more recent development) to obtain the asymptotic equivalent of the  $(j, k)$ -th component of  $\Omega_{n,H}^{-1}$  when  $|k-j| \rightarrow \infty$  stated in the third point of the proposition. □

**Remark 2.** From (21), for any  $\rho > 0$  there exists  $N_\rho \in \mathbb{N}$  such that for any  $k \geq N_\rho$  we have

$$H \{(2H-1) - \rho |2H-1|\} k^{2H-2} \leq (\Omega_{n,H})_{1,k} \leq H \{\rho |2H-1| + (2H-1)\} k^{2H-2}.$$

Since  $-1 < 2H - 1 < 1$  we obtain that

$$|(\Omega_{n,H})_{1,k}| \leq H \{\rho |2H - 1| + 1\} k^{2H-2}. \quad (27)$$

Observe that for  $k = 1, \dots, N_\rho$ , we have in view of (2) that the autocovariances  $(\Omega_{n,H})_{1,k}$  are bounded, i.e. there exists  $K_{H,\rho} > 0$  such that

$$|(\Omega_{n,H})_{1,k}| \leq K_{H,\rho}.$$

Since  $k^{2H-2} \leq 1$  for  $k = 1, \dots, N_\rho$ , it follows that

$$|(\Omega_{n,H})_{1,k}| \leq K_{H,\rho} k^{2H-2}, \quad k = 1, \dots, N_\rho. \quad (28)$$

Consequently, we obtain from (27) and (28) that for any  $k \geq 1$ ,

$$|(\Omega_{n,H})_{1,k}| \leq K k^{2H-2}, \quad (29)$$

where  $K = \max(H \{\rho |2H - 1| + 1\}, K_{H,\rho})$ .

A similar calculation can be done to show that there exists a positive constant  $K$  such that for any  $j, k = 1, 2, \dots$ , we have

$$\left| (\Omega_{n,H}^{-1})_{j,k} \right| \leq K \left( \frac{1}{k-j} \right)^{2H}. \quad (30)$$

## 5.2. Proof of Theorem 1

The proof of convergence in probability of  $(\hat{a}_n)_{n \geq 1}$  to  $a_0$  will be done in two steps. We first show that  $(\|Z_{n,H}\|_{\mathbb{R}^n}^2/n)_{n \geq 1}$  converges in mean square to  $1/(1 - a_0^2)$  and we prove in a second time that  $(Z_{n,H}^T U_{n,H}(a_0)/n)_{n \geq 1}$  converges in mean square to 0.

### 5.2.1. Proof of the convergence in mean square of $(\|Z_{n,H}\|_{\mathbb{R}^n}^2/n)_{n \geq 1}$ to $1/(1 - a_0^2)$

To establish the mean square convergence of  $(\|Z_{n,H}\|_{\mathbb{R}^n}^2/n)_{n \geq 1}$  to  $1/(1 - a_0^2)$ , we consider the following lemmas and remark.

**Lemma 1.** Under Assumption (A0), we have

$$\mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] \xrightarrow{n \rightarrow \infty} \frac{1}{1 - a_0^2}.$$

*Proof.* In view of (19), one has

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] &= \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^n a_0^{j_1+j_2-2} \mathbb{E} \left[ \left\{ L^{j_1} U_{n,H}^T(a_0) \right\} \left\{ L^{j_2} U_{n,H}(a_0) \right\} \right] \\ &= \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^n a_0^{j_1+j_2-2} \sum_{k=1}^n \mathbb{E} \left[ \left\{ L^{j_1}(U_{n,H}(a_0))_k \right\} \left\{ L^{j_2}(U_{n,H}(a_0))_k \right\} \right] \\ &= \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^n a_0^{j_1+j_2-2} \sum_{k=1}^n \sum_{r_1=1}^n \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1/2} \right)_{k,r_1} \left( \Omega_{n,H}^{-1/2} \right)_{k,r_2} \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1-j_2}^H(a_0) \right] \\ &= \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=j_2}^n \left( \Omega_{n,H}^{-1} \right)_{r_1,r_2} \left( \Omega_{n,H} \right)_{r_1+1-j_1,r_2+1-j_2}. \end{aligned}$$

The stationarity of  $(e_t^H)_{t \in \mathbb{Z}}$  implies that

- for  $j_2 < j_1$ :

$$(\Omega_{n,H})_{r_1+1-j_1, r_2+1-j_2} = (\Omega_{n,H})_{r_2, r_1+j_2-j_1},$$

- for  $j_2 \geq j_1$ :

$$(\Omega_{n,H})_{r_1+1-j_1, r_2+1-j_2} = (\Omega_{n,H})_{r_1, r_2+j_1-j_2}.$$

Thus, we obtain

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] &= \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=j_2}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H})_{r_2, r_1+j_2-j_1} \\ &\quad + \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=j_1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=j_2}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H})_{r_1, r_2+j_1-j_2}. \end{aligned}$$

This implies that

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] &= \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n (I_n)_{r_1, r_1+j_2-j_1} \\ &\quad - \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^{j_2-1} (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H})_{r_2, r_1+j_2-j_1} \\ &\quad + \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=j_1}^n a_0^{j_1+j_2-2} \sum_{r_2=j_2}^n (I_n)_{r_2, r_2+j_1-j_2} \\ &\quad - \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=j_1}^n a_0^{j_1+j_2-2} \sum_{r_1=1}^{j_1-1} \sum_{r_2=j_2}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H})_{r_1, r_2+j_1-j_2}. \end{aligned}$$

We have

$$\frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n (I_n)_{r_1, r_1+j_2-j_1} = 0$$

and

$$\frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=j_1}^n a_0^{j_1+j_2-2} \sum_{r_2=j_2}^n (I_n)_{r_2, r_2+j_1-j_2} = \frac{1}{n} \sum_{j=1}^n a_0^{2j-2} (n-j+1),$$

thus we may write

$$\mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] = \frac{1}{n} \sum_{j=1}^n a_0^{2j-2} (n-j+1) - W_{n,H}, \quad (31)$$

where

$$\begin{aligned} W_{n,H} &= \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^{j_2-1} (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H})_{r_2, r_1+j_2-j_1} \\ &\quad + \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=j_1}^n a_0^{j_1+j_2-2} \sum_{r_1=1}^{j_1-1} \sum_{r_2=j_2}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H})_{r_1, r_2+j_1-j_2}. \end{aligned} \quad (32)$$

We shall prove that  $W_{n,H} \xrightarrow[n \rightarrow \infty]{} 0$ . Consider the first term on the right side of Equation (32) (the second term can be treated in a similar way) and note that  $r_2 - r_1 \leq j_1 - r_1 - 1$  and  $r_2 - r_1 + j_1 - j_2 \leq j_1 - r_1 - 1$ . By the properties of the components of the matrix  $\Omega_{n,H}$  and its inverse given in (29) and (30), there exists a positive constant  $K$  such that

$$\begin{aligned} & \left| \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^{j_2-1} \left( \Omega_{n,H}^{-1} \right)_{r_1,r_2} \left( \Omega_{n,H} \right)_{r_2,r_1+j_2-j_1} \right| \\ & \leq \frac{K}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} |a_0|^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^{j_2-1} \frac{1}{(r_2 - r_1)^{2H}} \frac{1}{(r_2 - r_1 + j_1 - j_2)^{2-2H}} \\ & \leq \frac{K}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} |a_0|^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^{j_2-1} \frac{1}{(j_1 - r_1 - 1)^{2H}} \frac{1}{(j_1 - r_1 - 1)^{2-2H}} \\ & = \frac{K}{n} \sum_{j_1=1}^n \sum_{r_1=j_1}^n \frac{|a_0|^{j_1-1}}{(j_1 - r_1 - 1)^2} \sum_{j_2=1}^{j_1-1} (j_2 - 1) |a_0|^{j_2-1}. \end{aligned}$$

Since

$$\sum_{j_2=1}^{j_1-1} (j_2 - 1) |a_0|^{j_2-1} = \frac{1 - |a_0|^{j_1-1} \{(j_1 - 1)(1 - |a_0|) + 1\}}{(1 - |a_0|)^2} \leq \frac{1}{(1 - |a_0|)^2},$$

we deduce that

$$\left| \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^{j_2-1} \left( \Omega_{n,H}^{-1} \right)_{r_1,r_2} \left( \Omega_{n,H} \right)_{r_2,r_1+j_2-j_1} \right| \leq \frac{K}{n(1 - |a_0|)^2} \sum_{j_1=1}^n \sum_{r_1=j_1}^n \frac{|a_0|^{j_1-1}}{(j_1 - r_1 - 1)^2}.$$

Now we consider the change of indices  $r_1 = j_1 + k - 1$  to obtain that

$$\left| \frac{1}{n} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1-1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^{j_2-1} \left( \Omega_{n,H}^{-1} \right)_{r_1,r_2} \left( \Omega_{n,H} \right)_{r_2,r_1+j_2-j_1} \right| \leq \frac{K}{n(1 - |a_0|)^2} \sum_{j_1=1}^n |a_0|^{j_1-1} \sum_{k \geq 1} \frac{1}{k^2} \xrightarrow[n \rightarrow \infty]{} 0.$$

The second term on the right side of Equation (32) can be treated in the same way. The proof of the convergence of the sequence  $(W_{n,H})_{n \geq 1}$  towards 0 is then completed.

Finally, from (31) we conclude that

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] &= \left( \frac{n+1}{n} \right) \left( \frac{1 - a_0^{2n}}{1 - a_0^2} \right) - \frac{1}{n} \left( \frac{1 - a_0^{2n}}{1 - a_0^2} \right) - \frac{1}{n} \left( \frac{1 - a_0^{2n-2} \{n(1 - a_0^2) + a_0^2\}}{(1 - a_0^2)^2} \right) - W_{n,H} \\ &\xrightarrow[n \rightarrow \infty]{} \frac{1}{1 - a_0^2}, \end{aligned}$$

and the lemma is proved.  $\square$

**Remark 3.** Denote by  $\Sigma_{n,X,H}$  the variance matrix of the random vector  $(X_1, \dots, X_n)^\top$ , i.e. for any  $1 \leq i, j \leq n$ ,  $(\Sigma_{n,X,H})_{i,j} = \text{Cov}(X_i, X_j)$ , and let  $Z_{n,H}^{(1)}$  be the random vector defined by  $Z_{n,H}^{(1)} = \Sigma_{n,X,H}^{-1/2}(X_1, \dots, X_n)^\top$ . From (7), one has

$$\frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 = \frac{1}{n} Z_{n,H}^\top Z_{n,H} = \frac{1}{n} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}^\top \Omega_{n,H}^{-1} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \frac{1}{n} (Z_{n,H}^{(1)})^\top \Sigma_{n,X,H}^{1/2} \Omega_{n,H}^{-1} \Sigma_{n,X,H}^{1/2} Z_{n,H}^{(1)}.$$

The covariance matrix  $\Sigma_{n,X,H}^{1/2} \Omega_{n,H}^{-1} \Sigma_{n,X,H}^{1/2}$  is a symmetric real matrix. So there exists an orthogonal matrix  $Q_{n,H}$  such that  $\Lambda_{n,H} = Q_{n,H}^\top \Sigma_{n,X,H}^{1/2} \Omega_{n,H}^{-1} \Sigma_{n,X,H}^{1/2} Q_{n,H}$  is a diagonal matrix. Hence, the random variable  $\|Z_{n,H}\|_{\mathbb{R}^n}^2 / n$  can be rewritten as

$$\begin{aligned} \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 &= \frac{1}{n} (Q_{n,H}^\top Z_{n,H}^{(1)})^\top \Lambda_{n,H} (Q_{n,H}^\top Z_{n,H}^{(1)}) \\ &= \frac{1}{n} \sum_{j=1}^n \eta_{j,H}^{(n)} \left[ (Q_{n,H}^\top Z_{n,H}^{(1)})_j \right]^2, \end{aligned} \quad (33)$$

where  $\eta_{1,H}^{(n)}, \dots, \eta_{n,H}^{(n)}$  are the eigenvalues of the symmetric real matrix  $\Sigma_{n,X,H}^{1/2} \Omega_{n,H}^{-1} \Sigma_{n,X,H}^{1/2}$ . The expression of  $\|Z_{n,H}\|_{\mathbb{R}^n}^2 / n$  given in (33) has the advantage of involving the standard normal random vector  $Q_{n,H}^\top Z_{n,H}^{(1)}$ . This will considerably simplify the calculations in the following.

We show in the lemma below the boundedness of the eigenvalues  $\eta_{1,H}^{(n)}, \dots, \eta_{n,H}^{(n)}$  of the matrix  $\Sigma_{n,X,H}^{1/2} \Omega_{n,H}^{-1} \Sigma_{n,X,H}^{1/2}$  uniformly in  $n$ .

**Lemma 2.** Under Assumption (A0) and for any  $1 \leq j \leq n$ , we have

$$\frac{1}{(1 + |a_0|)^2} \leq \eta_{j,H}^{(n)} \leq \frac{1}{(1 - |a_0|)^2}.$$

*Proof.* Let  $f_X(\cdot)$  be the spectral density of  $(X_t)_{t \in \mathbb{Z}}$  and denote by  $f_{\epsilon^H}(\cdot)$  the spectral density of the stationary process  $(\epsilon_t^H)_{t \in \mathbb{Z}}$ . One can express  $f_X(\cdot)$  as a function of  $f_{\epsilon^H}(\cdot)$ . Indeed, using (16) and the stationarity of  $(X_t)_{t \in \mathbb{Z}}$  and  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  we have for any  $\omega \in \mathbb{R}$ ,

$$\begin{aligned} f_X(\omega) &= \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma_X(h) e^{-i\omega h} = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \text{Cov}(X_t, X_{t-h}) e^{-i\omega h} \\ &= \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \text{Cov} \left( \sum_{j \in \mathbb{N}} a_0^j \epsilon_{t-j}^H, \sum_{k \in \mathbb{N}} a_0^k \epsilon_{t-h-k}^H \right) e^{-i\omega h} \\ &= \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \sum_{j, k \in \mathbb{N}} a_0^{j+k} \gamma_{\epsilon^H}(h+k-j) e^{-i\omega(h+k-j)} e^{-i\omega(j-k)} \\ &= f_{\epsilon^H}(\omega) \left( \sum_{j \in \mathbb{N}} a_0^j e^{-i\omega j} \right) \left( \overline{\sum_{j \in \mathbb{N}} a_0^j e^{-i\omega j}} \right) = f_{\epsilon^H}(\omega) \left| \sum_{j \in \mathbb{N}} a_0^j e^{-i\omega j} \right|^2, \end{aligned}$$

where  $\gamma_X(\cdot)$  denotes the autocovariance function of the process  $(X_t)_{t \in \mathbb{Z}}$  and  $\bar{z}$  designates the complex conjugate of the complex number  $z$ . Furthermore, Assumption (A0) implies that the

modulus of the complex number  $a_0 e^{-i\omega}$  is strictly less than 1. Hence, we deduce that the spectral density of the stationary process  $(X_t)_{t \in \mathbb{Z}}$  is related to  $f_{\epsilon_H}(\cdot)$  as follows:

$$f_X(\omega) = f_{\epsilon_H}(\omega) \left| \frac{1}{1 - a_0 e^{-i\omega}} \right|^2 = \frac{f_{\epsilon_H}(\omega)}{1 - 2a_0 \cos(\omega) + a_0^2}, \quad \forall \omega \in \mathbb{R}. \quad (34)$$

From this last equation and taking into consideration the fact that  $(1 - |a_0|)^2 \leq 1 - 2a_0 \cos(\omega) + a_0^2 \leq (1 + |a_0|)^2$ , one obtains that for any  $\omega \in \mathbb{R}$ ,

$$(1 - |a_0|)^2 f_X(\omega) \leq f_{\epsilon_H}(\omega) \leq (1 + |a_0|)^2 f_X(\omega). \quad (35)$$

Note now that for any vector  $V$  of  $\mathbb{R}^{n \times 1}$ , we have

$$\begin{aligned} V^T (\Omega_{n,H} - (1 + |a_0|)^2 \Sigma_{n,X,H}) V &= \sum_{k=1}^n \sum_{j=1}^n V_j (\Omega_{n,H} - (1 + |a_0|)^2 \Sigma_{n,X,H})_{j,k} V_k \\ &= \sum_{k,j=1}^n V_j \int_{-\pi}^{\pi} \{f_{\epsilon_H}(\omega) - (1 + |a_0|)^2 f_X(\omega)\} e^{i\omega(k-j)} d\omega V_k \\ &= \int_{-\pi}^{\pi} \{f_{\epsilon_H}(\omega) - (1 + |a_0|)^2 f_X(\omega)\} V^T \gamma_n(\omega) V d\omega, \end{aligned} \quad (36)$$

where  $\gamma_n(\omega) = V_n(\omega) V_n^*(\omega)$  with  $V_n(\omega) = (e^{i\omega}, e^{2i\omega}, \dots, e^{ni\omega})^T$ . The notation  $V^*$  designates the conjugate transpose of the vector  $V$ .

Furthermore, the real  $V^T \gamma_n(\omega) V$  is a non-negative number. In fact,

$$V^T \gamma_n(\omega) V = V^T V_n(\omega) V_n^*(\omega) V = (V_n^*(\omega) V)^* (V_n^*(\omega) V) = |V_n^*(\omega) V|^2.$$

Then, from (35) and (36), one has

$$V^T (\Omega_{n,H} - (1 + |a_0|)^2 \Sigma_{n,X,H}) V \leq 0. \quad (37)$$

Now, let  $\zeta$  be an element of the spectrum of  $\Omega_{n,H}^{-1/2} \Sigma_{n,X,H} \Omega_{n,H}^{-1/2}$ . Then, there exists  $C_1 \in \mathbb{R}^{n \times 1} \setminus \{0_{\mathbb{R}^{n \times 1}}\}$  such that  $\Omega_{n,H}^{-1/2} \Sigma_{n,X,H} \Omega_{n,H}^{-1/2} C_1 = \zeta C_1$ . Hence, it follows that

$$C_1^T \Omega_{n,H}^{-1/2} \Sigma_{n,X,H} \Omega_{n,H}^{-1/2} C_1 = \zeta \|C_1\|_{\mathbb{R}^n}^2.$$

Taking  $C_2 = \Omega_{n,H}^{-1/2} C_1$ , we obtain from this last equation that

$$C_2^T \Sigma_{n,X,H} C_2 = \zeta \left\| \Omega_{n,H}^{1/2} C_2 \right\|_{\mathbb{R}^n}^2.$$

Using (37), we deduce that

$$C_2^T \Omega_{n,H} C_2 \leq (1 + |a_0|)^2 \zeta \left\| \Omega_{n,H}^{1/2} C_2 \right\|_{\mathbb{R}^n}^2.$$

Since  $C_1 = \Omega_{n,H}^{1/2} C_2 \neq 0_{\mathbb{R}^{n \times 1}}$  as an eigenvector, we conclude that

$$\zeta \geq \frac{C_2^T \Omega_{n,H} C_2}{\left\| \Omega_{n,H}^{1/2} C_2 \right\|_{\mathbb{R}^n}^2} \frac{1}{(1 + |a_0|)^2} = \frac{1}{(1 + |a_0|)^2}. \quad (38)$$



A similar calculation can be done to prove that, for any vector  $V$  of  $\mathbb{R}^{n \times 1}$ ,

$$V^T (\Omega_{n,H} - (1 - |a_0|)^2 \Sigma_{n,X,H}) V \geq 0.$$

Arguing as before, one obtains that

$$\zeta \leq \frac{C_2^T \Omega_{n,H} C_2}{\left\| \Omega_{n,H}^{1/2} C_2 \right\|_{\mathbb{R}^n}^2} \frac{1}{(1 - |a_0|)^2} = \frac{1}{(1 - |a_0|)^2}.$$

Since for any  $\alpha \in \mathbb{R}$ ,

$$\begin{aligned} \det \left( \Sigma_{n,X,H}^{-1/2} \Omega_{n,H}^{-1} \Sigma_{n,X,H}^{1/2} - \alpha I_n \right) &= \det \left( \Sigma_{n,X,H}^{1/2} \left\{ \Omega_{n,H}^{-1} \Sigma_{n,X,H} - \alpha I_n \right\} \Sigma_{n,X,H}^{-1/2} \right) = \det \left( \Omega_{n,H}^{-1} \Sigma_{n,X,H} - \alpha I_n \right) \\ &= \det \left( \Omega_{n,H}^{-1/2} \left\{ \Omega_{n,H}^{-1/2} \Sigma_{n,X,H} \Omega_{n,H}^{-1/2} - \alpha I_n \right\} \Omega_{n,H}^{1/2} \right) \\ &= \det \left( \Omega_{n,H}^{-1/2} \Sigma_{n,X,H} \Omega_{n,H}^{-1/2} - \alpha I_n \right), \end{aligned}$$

the spectrum of the matrix  $\Sigma_{n,X,H}^{1/2} \Omega_{n,H}^{-1} \Sigma_{n,X,H}^{1/2}$  is equal to the set of the eigenvalues of  $\Omega_{n,H}^{-1/2} \Sigma_{n,X,H} \Omega_{n,H}^{-1/2}$ , the lemma is then proved.  $\square$

**Lemma 3.** *Under Assumption (A0), we have*

$$\frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \xrightarrow[n \rightarrow \infty]{\mathbb{L}^2} \frac{1}{1 - a_0^2}.$$

*Proof.* In view of Lemma 2, Remark 3 and (33), we have

$$\begin{aligned} \text{Var} \left( \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right) &= \frac{1}{n^2} \sum_{j=1}^n \left( \eta_{j,H}^{(n)} \right)^2 \text{Var} \left( \left[ \left( Q_{n,H}^T Z_{n,H}^{(1)} \right)_j \right]^2 \right) = \frac{2}{n^2} \sum_{j=1}^n \left( \eta_{j,H}^{(n)} \right)^2 \\ &\leq \frac{2}{n(1 - |a_0|)^4} \xrightarrow[n \rightarrow \infty]{} 0. \end{aligned}$$

Now, observe that

$$\mathbb{E} \left[ \left| \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 - \frac{1}{1 - a_0^2} \right|^2 \right] = \text{Var} \left( \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right) + \left( \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] - \frac{1}{1 - a_0^2} \right)^2.$$

Thanks to Lemma 1 and the convergence stated just before, the expected result is obtained.  $\square$

### 5.2.2. Proof of the convergence in mean square of $(Z_{n,H}^T U_{n,H}(a_0)/n)_{n \geq 1}$ to 0

**Lemma 4.** *Under Assumption (A0), we have*

$$\frac{1}{n} \left[ Z_{n,H}^T U_{n,H}(a_0) \right] \xrightarrow[n \rightarrow \infty]{\mathbb{L}^2} 0.$$

*Proof.* Due to the fact that  $e_t^H(a_0)$  is centered for any  $t$ , we notice that for fixed  $(i, j, k)$  the joint cumulant of  $e_0^H(a_0)$ ,  $e_i^H(a_0)$ ,  $e_j^H(a_0)$  and  $e_k^H(a_0)$  is given by:

$$\begin{aligned} \text{cum}(e_0^H(a_0), e_i^H(a_0), e_j^H(a_0), e_k^H(a_0)) &= \mathbb{E} \left[ e_0^H(a_0) e_i^H(a_0) e_j^H(a_0) e_k^H(a_0) \right] \\ &\quad - \mathbb{E} \left[ e_0^H(a_0) e_i^H(a_0) \right] \mathbb{E} \left[ e_j^H(a_0) e_k^H(a_0) \right] - \mathbb{E} \left[ e_0^H(a_0) e_j^H(a_0) \right] \mathbb{E} \left[ e_i^H(a_0) e_k^H(a_0) \right] \\ &\quad - \mathbb{E} \left[ e_0^H(a_0) e_k^H(a_0) \right] \mathbb{E} \left[ e_i^H(a_0) e_j^H(a_0) \right]. \end{aligned}$$

From (7) and (19), the positive definiteness of  $\Omega_{n,H}$  and the fact that the process  $(e_t^H(a_0))_{t \in \mathbb{Z}}$  defined in (8) is centered, we obtain that

$$\begin{aligned} \mathbb{E} \left[ \left( \frac{1}{n} \left\{ Z_{n,H}^T U_{n,H}(a_0) \right\} \right)^2 \right] &= \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=1}^n a_0^{j_1+j_2-2} \mathbb{E} \left[ (L^{j_1} U_{n,H}(a_0))^T U_{n,H}(a_0) (L^{j_2} U_{n,H}(a_0))^T U_{n,H}(a_0) \right] \\ &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{k_1, k_2=1}^n \sum_{r_1, \dots, r_4=1}^n \left( \Omega_{n,H}^{-1/2} \right)_{k_1, r_1} \left( \Omega_{n,H}^{-1/2} \right)_{k_1, r_2} \left( \Omega_{n,H}^{-1/2} \right)_{k_2, r_3} \left( \Omega_{n,H}^{-1/2} \right)_{k_2, r_4} \\ &\quad \times \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1}^H(a_0) e_{r_3+1-j_2}^H(a_0) e_{r_4+1}^H(a_0) \right] \\ &= \Delta_{1,n,H} + \Delta_{2,n,H} + \Delta_{3,n,H} + \Delta_{4,n,H}, \end{aligned} \tag{39}$$

where

$$\begin{aligned} \Delta_{1,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1, r_2} \left( \Omega_{n,H}^{-1} \right)_{r_3, r_4} \\ &\quad \times \text{cum} \left( e_{r_1+1-j_1}^H(a_0), e_{r_2+1}^H(a_0), e_{r_3+1-j_2}^H(a_0), e_{r_4+1}^H(a_0) \right), \\ \Delta_{2,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1, r_2} \left( \Omega_{n,H}^{-1} \right)_{r_3, r_4} \\ &\quad \times \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_4+1}^H(a_0) \right], \\ \Delta_{3,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1, r_2} \left( \Omega_{n,H}^{-1} \right)_{r_3, r_4} \\ &\quad \times \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_4+1}^H(a_0) \right] \text{ and} \\ \Delta_{4,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1, r_2} \left( \Omega_{n,H}^{-1} \right)_{r_3, r_4} \\ &\quad \times \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right]. \end{aligned}$$

Since the vector  $(\epsilon_1^H, \dots, \epsilon_{n+1}^H)^T$  is Gaussian random vector, any joint cumulant involving three (or more) elements of  $(\epsilon_1^H, \dots, \epsilon_{n+1}^H)^T$  is zero. The random variables  $e_{r_1+1-j_1}^H(a_0)$ ,  $e_{r_2+1}^H(a_0)$ ,  $e_{r_3+1-j_2}^H(a_0)$  and  $e_{r_4+1}^H(a_0)$  in (39) are elements of the Gaussian vector  $(\epsilon_1^H, \dots, \epsilon_{n+1}^H)^T$  so

$$\text{cum} \left( e_{r_1+1-j_1}^H(a_0), e_{r_2+1}^H(a_0), e_{r_3+1-j_2}^H(a_0), e_{r_4+1}^H(a_0) \right) = 0.$$

Thus, we have

$$\Delta_{1,n,H} = 0.$$

Let  $\Omega_{n+1,H}$  be the covariance matrix of the random vector  $(\epsilon_1^H, \dots, \epsilon_{n+1}^H)^\top$ . Thanks to the stationarity of the process  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  one has  $(\Omega_{n+1,H})_{r_3+1-j_2, r_4+1} = (\Omega_{n,H})_{r_3-j_2, r_4}$  when  $j_2, r_4 = 1, \dots, n$  and  $j_2 + 1 \leq r_3 \leq n$ . Thus  $\Delta_{2,n,H}$  can be rewritten as:

$$\begin{aligned} \Delta_{2,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{r_3, r_4} (\Omega_{n+1,H})_{r_1+1-j_1, r_2+1} (\Omega_{n+1,H})_{r_3+1-j_2, r_4+1} \\ &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2+1}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{r_3, r_4} (\Omega_{n+1,H})_{r_1+1-j_1, r_2+1} (\Omega_{n,H})_{r_3-j_2, r_4} \\ &\quad + \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{j_2, r_4} (\Omega_{n+1,H})_{r_1+1-j_1, r_2+1} (\Omega_{n+1,H})_{1, r_4+1}. \end{aligned}$$

This implies that

$$\begin{aligned} \Delta_{2,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2+1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n+1,H})_{r_1+1-j_1, r_2+1} (I_n)_{r_3, r_3-j_2} \\ &\quad + \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1+1}^n \sum_{r_2=1}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{j_2, r_4} (\Omega_{n,H})_{r_1-j_1, r_2} (\Omega_{n+1,H})_{1, r_4+1} \\ &\quad + \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_2=1}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{j_1, r_2} (\Omega_{n,H}^{-1})_{j_2, r_4} (\Omega_{n+1,H})_{1, r_2+1} (\Omega_{n+1,H})_{1, r_4+1} \\ &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2+1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n+1,H})_{r_1+1-j_1, r_2+1} (I_n)_{r_3, r_3-j_2} \\ &\quad + \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1+1}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{j_2, r_4} (\Omega_{n+1,H})_{1, r_4+1} (I_n)_{r_1, r_1-j_1} \\ &\quad + \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_2=1}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{j_1, r_2} (\Omega_{n,H}^{-1})_{j_2, r_4} (\Omega_{n+1,H})_{1, r_2+1} (\Omega_{n+1,H})_{1, r_4+1}. \end{aligned}$$

Since  $j_1, j_2 = 1, \dots, n$ ,  $(I_n)_{r_1, r_1-j_1} = (I_n)_{r_3, r_3-j_2} = 0$ , we obtain that

$$\begin{aligned} \Delta_{2,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_2=1}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{j_1, r_2} (\Omega_{n,H}^{-1})_{j_2, r_4} (\Omega_{n+1,H})_{1, r_2+1} (\Omega_{n+1,H})_{1, r_4+1} \\ &= \left( \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{r=1}^n (\Omega_{n,H}^{-1})_{j, r} (\Omega_{n+1,H})_{1, r+1} \right)^2. \end{aligned} \quad (40)$$

Using a similar approach as that given in the proof of Lemma 1, one can confirm that the term in the right side of Equation (40) is equal to the square of the expected value of the random variable  $Z_{n,H}^\top U_{n,H}(a_0)/n$ . Similar calculations to those of the proof of Lemma 1 imply that

$$\mathbb{E} \left[ \frac{1}{n} \left\{ Z_{n,H}^\top U_{n,H}(a_0) \right\} \right] \xrightarrow{n \rightarrow \infty} 0.$$

Let us be more precise. From (19), we have

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{n} \left\{ Z_{n,H}^T U_{n,H}(a_0) \right\} \right] &= \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{k=1}^n \mathbb{E} \left[ \left\{ L^j (U_{n,H}(a_0))_k \right\} (U_{n,H}(a_0))_k \right] \\ &= \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{k=1}^n \sum_{r_1, r_2=1}^n \left( \Omega_{n,H}^{-1/2} \right)_{k, r_1} \left( \Omega_{n,H}^{-1/2} \right)_{k, r_2} \mathbb{E} \left[ e_{r_1+1-j}^H(a_0) e_{r_2+1}^H(a_0) \right] \end{aligned} \quad (41)$$

$$= \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{r_1=j}^n \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1, r_2} \left( \Omega_{n+1,H} \right)_{r_1+1-j, r_2+1}. \quad (42)$$

As before, we proceed by cases to deduce that

$$\begin{aligned} \mathbb{E} \left[ \frac{1}{n} \left\{ Z_{n,H}^T U_{n,H}(a_0) \right\} \right] &= \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{r_1=j+1}^n \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1, r_2} \left( \Omega_{n,H} \right)_{r_1-j, r_2} \\ &\quad + \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1} \right)_{j, r_2} \left( \Omega_{n+1,H} \right)_{1, r_2+1} \\ &= \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{r_1=j+1}^n (I_n)_{r_1, r_1-j} + \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1} \right)_{j, r_2} \left( \Omega_{n+1,H} \right)_{1, r_2+1} \\ &= \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1} \right)_{j, r_2} \left( \Omega_{n+1,H} \right)_{1, r_2+1} \\ &= \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \left( \Omega_{n,H}^{-1} \right)_{j, j} \left( \Omega_{n+1,H} \right)_{1, j+1} + \frac{1}{n} \sum_{j=1}^n a_0^{j-1} \sum_{\substack{1 \leq r_2 \leq n \\ r_2 \neq j}} \left( \Omega_{n,H}^{-1} \right)_{j, r_2} \left( \Omega_{n+1,H} \right)_{1, r_2+1}. \end{aligned}$$

Finally, in view of (29) and (30) we conclude that

$$\left| \mathbb{E} \left[ \frac{1}{n} \left\{ Z_{n,H}^T U_{n,H}(a_0) \right\} \right] \right| \leq K \left( \frac{1}{n} \sum_{j=1}^n |a_0|^{j-1} + \frac{1}{n} \sum_{j=1}^n |a_0|^{j-1} \sum_{\substack{1 \leq r \leq n \\ r \neq j}} \frac{1}{(r-j)^{2H}} \frac{1}{r^{2-2H}} \right). \quad (43)$$

Using the fact that the autoregressive parameter  $a_0$  is assumed to belong to  $(-1, 1)$ , Cesàro's Lemma implies that the first term in the right hand side of (43) tends to 0.

Proceeding by a simple calculation and using Cesàro's Lemma, one can also prove that the second term in the right hand side of (43) converges to 0. More explicitly,

$$\begin{aligned} &\frac{1}{n} \sum_{j=1}^n |a_0|^{j-1} \sum_{\substack{1 \leq r \leq n \\ r \neq j}} \frac{1}{(r-j)^{2H}} \frac{1}{r^{2-2H}} \\ &= \frac{1}{n} \sum_{r=1}^n \frac{1}{r^{2-2H}} \sum_{\substack{1 \leq j \leq n \\ j \neq r}} |a_0|^{j-1} \frac{1}{(r-j)^{2H}} \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{n} \sum_{r=1}^n \frac{1}{r^{2-2H}} |a_0|^{r-1} \sum_{\substack{1-r \leq j \leq n-r \\ j \neq 0}} |a_0|^j \frac{1}{j^{2H}} \\
&= \frac{1}{n} \sum_{r=2}^n \frac{1}{r^{2-2H}} |a_0|^{r-1} \sum_{j=1-r}^{-1} |a_0|^j \frac{1}{j^{2H}} + \frac{1}{n} \sum_{r=1}^n \frac{1}{r^{2-2H}} |a_0|^{r-1} \sum_{j=1}^{n-r} |a_0|^j \frac{1}{j^{2H}} \\
&\leq \frac{1}{n} \sum_{r=1}^n \frac{1}{r^{1-2H}} |a_0|^{r-1} \max \left( \frac{1}{|a_0|}, \frac{1}{|a_0|^r r^{2H}} \right) + \frac{1}{n} \sum_{r=1}^n \frac{1}{r^{2-2H}} |a_0|^r \left( \frac{1 - |a_0|^{n-r}}{1 - |a_0|} \right) \\
&\xrightarrow{n \rightarrow \infty} 0.
\end{aligned}$$

Therefore, we deduce that

$$\Delta_{2,n,H} \xrightarrow{n \rightarrow \infty} 0.$$

It is obvious that by exchanging the roles of  $j_1$  and  $j_2$  and of  $r_2$  and  $r_4$  in the expression of  $\Delta_{4,n,H}$ , we obtain that  $\Delta_{4,n,H} = \Delta_{2,n,H}$ . Then, the sequence  $(\Delta_{4,n,H})_{n \geq 1}$  also converges to 0.

It remains to show that  $(\Delta_{3,n,H})_{n \geq 1}$  converges to 0 to obtain the convergence in mean square of  $(Z_{n,H}^T U_{n,H}(a_0)/n)_{n \geq 1}$  to 0 and thus the convergence in probability of the generalized least squares estimator  $\hat{a}_n$  defined in (9) to the true parameter  $a_0$ .

We use the stationarity of the process  $(e_t^H(a_0))_{t \in \mathbb{Z}}$  and we follow the same calculation procedures developed to obtain the explicit expression of  $\Delta_{2,n,H}$  given in (40). This leads to

$$\begin{aligned}
\Delta_{3,n,H} &= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{r_3, r_4} (\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2} (\Omega_{n+1,H})_{r_2+1, r_4+1} \\
&= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{r_3, r_4} (\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2} (\Omega_{n,H})_{r_2, r_4} \\
&= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n (\Omega_{n,H}^{-1})_{r_3, r_4} (\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2} (I_n)_{r_1, r_4} \\
&= \frac{1}{n^2} \sum_{j_1, j_2=1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=j_2}^n (\Omega_{n,H}^{-1})_{r_3, r_1} (\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2} \\
&= \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=j_2}^n (\Omega_{n,H}^{-1})_{r_3, r_1} (\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2} \\
&\quad + \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=j_1+1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=j_2}^n (\Omega_{n,H}^{-1})_{r_3, r_1} (\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2}.
\end{aligned}$$

The stationarity of  $(\epsilon_t^H)_{t \in \mathbb{Z}}$  implies that

- for  $j_2 \leq j_1$ :

$$(\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2} = (\Omega_{n,H})_{r_1+j_2-j_1, r_3},$$

- for  $j_2 > j_1$ :

$$(\Omega_{n,H})_{r_1+1-j_1, r_3+1-j_2} = (\Omega_{n,H})_{r_1, r_3+j_1-j_2}.$$

It follows that

$$\begin{aligned}
\Delta_{3,n,H} &= \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=j_2}^n \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1+j_2-j_1,r_3} \\
&\quad + \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=j_1+1}^n a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=j_2}^n \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1,r_3+j_1-j_2} \\
&= \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1+j_2-j_1,r_3} \\
&\quad - \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=1}^{j_2-1} \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1+j_2-j_1,r_3} \\
&\quad + \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=j_1+1}^n a_0^{j_1+j_2-2} \sum_{r_1=1}^n \sum_{r_3=j_2}^n \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1,r_3+j_1-j_2} \\
&\quad - \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=j_1+1}^n a_0^{j_1+j_2-2} \sum_{r_1=1}^{j_1-1} \sum_{r_3=j_2}^n \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1,r_3+j_1-j_2}.
\end{aligned} \tag{44}$$

Thus we have

$$\begin{aligned}
\Delta_{3,n,H} &= \frac{1}{n^2} \sum_{j_1=1}^n (n - j_1 + 1) a_0^{2j_1-2} - \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=1}^{j_1} a_0^{j_1+j_2-2} \sum_{r_1=j_1}^n \sum_{r_3=1}^{j_2-1} \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1+j_2-j_1,r_3} \\
&\quad - \frac{1}{n^2} \sum_{j_1=1}^n \sum_{j_2=j_1+1}^n a_0^{j_1+j_2-2} \sum_{r_1=1}^{j_1-1} \sum_{r_3=j_2}^n \left( \Omega_{n,H}^{-1} \right)_{r_3,r_1} (\Omega_{n,H})_{r_1,r_3+j_1-j_2} \\
&= \frac{1}{n} \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right].
\end{aligned} \tag{45}$$

Now, using Lemma 1 we deduce that

$$\Delta_{3,n,H} \xrightarrow[n \rightarrow \infty]{} 0.$$

The lemma is then proved.  $\square$

The convergences established in Lemmas 3 and 4 imply that the joint vector  $(Z_{n,H}^T U_{n,H}(a_0), \|Z_{n,H}\|_{\mathbb{R}^n}^2/n)$  converges in probability to  $(0, 1/(1-a_0^2))$ . Let now  $g$  be the continuous function defined on  $\mathbb{R} \times \mathbb{R}^*$  by  $g(x, y) = xy^{-1}$ , it is obvious that the vector  $(0, 1/(1-a_0^2))$  belongs to the set of continuity points of the function  $g$ . The continuous mapping theorem yields

$$\hat{a}_n - a_0 = g \left( Z_{n,H}^T U_{n,H}(a_0)/n, \|Z_{n,H}\|_{\mathbb{R}^n}^2/n \right) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} g \left( 0, 1/(1-a_0^2) \right) = 0.$$

The proof of Theorem 1 is then complete.

**Remark 4.** The following equivalent formulation of Equality (41)

$$\mathbb{E} \left[ \frac{1}{\sqrt{n}} \left\{ Z_{n,H}^T U_{n,H}(a_0) \right\} \right] = \frac{1}{\sqrt{n}} \sum_{j=1}^n a_0^{j-1} \sum_{r_1=j}^n \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1,r_2} \mathbb{E} \left[ e_{r_1+1-j}^H(a_0) e_{r_2+1}^H(a_0) \right] \tag{46}$$

will be helpful in the proof of Theorem 2.

### 5.3. Proof of Theorem 2

We establish in this section the proof of the asymptotic normality of the generalized least squares estimator  $\hat{a}_n$  given in (9) under the causality of  $(X_t)_{t \in \mathbb{Z}}$  (i.e. when the true parameter  $a_0$  is assumed to belong to  $(-1, 1)$ ) and in the case where  $H$  is known.

Recall that the normal distribution admits moments of all orders and is completely determined by its moments. The expected convergence in distribution stated in Theorem 2 can be obtained using the method of moments (see Theorem 30.2. of Billingsley [1986]).

In our case, Theorem 2 will be proved as soon as for any  $r \geq 1$

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left\{ \sqrt{n}(\hat{a}_n - a_0) \right\}^r \right] = \mathbb{E} [G_1^r], \quad (47)$$

where  $G_1$  is normally distributed with mean 0 and variance  $1 - a_0^2$ .

In view of the explicit expression of  $\hat{a}_n - a_0$  (see (10)), the random variable  $\sqrt{n}(\hat{a}_n - a_0)$  can be rewritten in the form:

$$\sqrt{n}(\hat{a}_n - a_0) = \frac{\frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0)}{\frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2}. \quad (48)$$

From Lemma 3, the random variable  $\|Z_{n,H}\|_{\mathbb{R}^n}^2/n$  converges in mean square to  $1/(1 - a_0^2)$ . This implies that the sequence  $(\|Z_{n,H}\|_{\mathbb{R}^n}^2/n)_{n \geq 1}$  converges in probability to  $1/(1 - a_0^2)$ .

The idea of the proof is to show using the method of moments that  $Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}$  has a limiting centered normal distribution with variance  $1/(1 - a_0^2)$ . The proof will be concluded by applying Slutsky's theorem to  $\|Z_{n,H}\|_{\mathbb{R}^n}^2/n$  and  $Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}$ .

The asymptotic distribution of the sequence  $(Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n})_{n \geq 1}$  is given in the following lemma.

**Lemma 5.** *Under Assumption (A0), the sequence of random variables  $(Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n})_{n \geq 1}$  has a limiting centered normal distribution with variance  $1/(1 - a_0^2)$ .*

*Proof.* Using the existence of moments of all orders of  $Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}$  and in view of Theorem 30.2. of Billingsley [1986], the proof of Lemma 5 is obtained once

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^r \right] = \mathbb{E} [G_2^r], \quad \text{for all } r \geq 1, \quad (49)$$

where  $G_2$  is normally distributed with mean 0 and variance  $1/(1 - a_0^2)$ .

To prove (49), in a first step we consider an odd order  $r$  and in a second step  $r$  will be assumed to be even.

Recall that for all  $r \geq 1$ , we have

$$\mathbb{E} [G_2^r] = \begin{cases} 0 & \text{if } r \text{ is odd} \\ \{1/(1 - a_0^2)\}^{r/2} (r-1)!! & \text{if } r \text{ is even,} \end{cases} \quad (50)$$

where  $r!! = r(r-2)(r-4) \cdots 1$ .

We start from (19), we use the fact that the random vector  $(e_1^H(a_0), \dots, e_{n+1}^H(a_0))^T$  is a centered Gaussian random vector and we exploit the expression of moments in terms of cumulants to obtain

that

$$\begin{aligned}
& \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^\top U_{n,H}(a_0) \right\}^r \right] \\
&= \frac{1}{n^{r/2}} \sum_{j_1, j_2, \dots, j_r=1}^n a_0^{\sum_{i=1}^r (j_i-1)} \sum_{k_1=j_1}^n \sum_{k_2=1}^n \sum_{k_3=j_2}^n \sum_{k_4=1}^n \dots \sum_{k_{2r-1}=j_r}^n \sum_{k_{2r}=1}^n \left\{ \prod_{i=1}^r (\Omega_{n,H}^{-1})_{k_{2i-1}, k_{2i}} \right. \\
&\quad \left. \times \mathbb{E} \left[ \prod_{i=1}^r e_{k_{2i-1}+1-j_i}^H(a_0) e_{k_{2i}+1}^H(a_0) \right] \right\} \\
&= \frac{1}{n^{r/2}} \sum_{j_1, j_2, \dots, j_r=1}^n a_0^{\sum_{i=1}^r (j_i-1)} \sum_{k_1=j_1}^n \sum_{k_2=1}^n \sum_{k_3=j_2}^n \sum_{k_4=1}^n \dots \sum_{k_{2r-1}=j_r}^n \sum_{k_{2r}=1}^n \left\{ \prod_{i=1}^r (\Omega_{n,H}^{-1})_{k_{2i-1}, k_{2i}} \right. \\
&\quad \left. \times \sum_{\pi} \prod_{B \in \pi} \text{cum} \left( e_i^H(a_0) : i \in B \right) \right\}, \tag{51}
\end{aligned}$$

where  $\pi$  runs through the list of all unordered pair partitions of  $\{k_1+1-j_1, k_2+1, k_3+1-j_2, k_4+1, \dots, k_{2r-1}+1-j_r, k_{2r}+1\}$  and  $B$  runs through the list of all blocks of the partition  $\pi$  (see Brillinger [1981]).

**Remark 5.** In the above expression, we only kept unordered pair partitions of  $\{k_1+1-j_1, k_2+1, k_3+1-j_2, k_4+1, \dots, k_{2r-1}+1-j_r, k_{2r}+1\}$ . This is due to the fact that the random vector  $(e_1^H(a_0), \dots, e_{n+1}^H(a_0))^\top$  is a centered Gaussian vector.

Any joint cumulant involving three (or more) elements of  $(e_1^H(a_0), \dots, e_{n+1}^H(a_0))^\top$  is zero. Moreover, since the joint cumulant of just one random variable is its expectation and the process  $(e_t^H(a_0))_{t \in \mathbb{Z}}$  is centered, the first cumulant of any element of  $(e_1^H(a_0), \dots, e_{n+1}^H(a_0))^\top$  vanishes.

Thanks to the above remark, we may write (51) as

$$\begin{aligned}
& \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^\top U_{n,H}(a_0) \right\}^r \right] \\
&= \frac{1}{n^{r/2}} \sum_{j_1, j_2, \dots, j_r=1}^n a_0^{\sum_{i=1}^r (j_i-1)} \sum_{k_1=j_1}^n \sum_{k_2=1}^n \sum_{k_3=j_2}^n \sum_{k_4=1}^n \dots \sum_{k_{2r-1}=j_r}^n \sum_{k_{2r}=1}^n \left\{ \prod_{i=1}^r (\Omega_{n,H}^{-1})_{k_{2i-1}, k_{2i}} \right. \\
&\quad \left. \times \sum_{\pi} \prod_{B \in \pi} \mathbb{E} \left[ e_i^H(a_0) : i \in B \right] \right\}. \tag{52}
\end{aligned}$$

We strength the fact that when  $r = 1$ , (52) is (46) and for  $r = 2$  we retrieve an analogous formula to (39).

The following remark will help us to develop (51) more explicitly.

**Remark 6.** • In Equation (51), the exchange of the roles of indices with even sub-indices does not modify the term in the right hand side. It is the same when we exchange the roles of



indices with odd sub-indices. For example, for the indices  $k_1$  and  $k_3$  it holds that

$$\begin{aligned} & \frac{1}{n^{r/2}} \sum_{j_1, j_2, \dots, j_r=1}^n a_0^{\sum_{i=1}^r (j_i-1)} \sum_{k_1=j_1}^n \sum_{k_2=1}^n \sum_{k_3=j_2}^n \sum_{k_4=1}^n \dots \sum_{k_{2r-1}=j_r}^n \sum_{k_{2r}=1}^n \left\{ \prod_{i=1}^r (\Omega_{n,H}^{-1})_{k_{2i-1}, k_{2i}} \right. \\ & \quad \left. \times \sum_{\pi} \prod_{B \in \pi} \text{cum} \left( e_i^H(a_0) : i \in B \right) \right\} \\ &= \frac{1}{n^{r/2}} \sum_{j_1, j_2, \dots, j_r=1}^n a_0^{\sum_{i=1}^r (j_i-1)} \sum_{k_1=j_2}^n \sum_{k_2=1}^n \sum_{k_3=j_1}^n \sum_{k_4=1}^n \dots \sum_{k_{2r-1}=j_r}^n \sum_{k_{2r}=1}^n \left\{ \prod_{i=1}^r (\Omega_{n,H}^{-1})_{k_{2i-1}, k_{2i}} \right. \\ & \quad \left. \times \sum_{\pi} \prod_{B \in \pi} \text{cum} \left( e_i^H(a_0) : i \in B \right) \right\}, \end{aligned}$$

- Again in Equation (51) when  $r$  is odd, every pair partition  $\pi$  of the set of indices  $\{k_1 + 1 - j_1, k_2 + 1, k_3 + 1 - j_2, k_4 + 1, \dots, k_{2r-1} + 1 - j_r, k_{2r} + 1\}$  contains at least one block of the type  $\{k_{2i-1} + 1 - j_i, k_{i'} + 1\}$ . To see that, let us consider the case when  $r = 3$ . The third moment of  $Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}$  is given by:

$$\begin{aligned} & \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^3 \right] \\ &= \frac{1}{n^{3/2}} \sum_{j_1, j_2, j_3=1}^n a_0^{j_1+j_2+j_3-3} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n \sum_{r_5=j_3}^n \sum_{r_6=1}^n \left\{ (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{r_3, r_4} (\Omega_{n,H}^{-1})_{r_5, r_6} \right. \\ & \quad \left. \times \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1}^H(a_0) e_{r_3+1-j_2}^H(a_0) e_{r_4+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) e_{r_6+1}^H(a_0) \right] \right\} \\ &= \frac{1}{n^{3/2}} \sum_{j_1, j_2, j_3=1}^n a_0^{j_1+j_2+j_3-3} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \sum_{r_3=j_2}^n \sum_{r_4=1}^n \sum_{r_5=j_3}^n \sum_{r_6=1}^n \left\{ (\Omega_{n,H}^{-1})_{r_1, r_2} (\Omega_{n,H}^{-1})_{r_3, r_4} (\Omega_{n,H}^{-1})_{r_5, r_6} \right. \\ & \quad \times \left[ \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_5+1-j_3}^H(a_0) e_{r_6+1}^H(a_0) \right] \right. \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_5+1-j_3}^H(a_0) e_{r_6+1}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right] \mathbb{E} \left[ e_{r_5+1-j_3}^H(a_0) e_{r_6+1}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_6+1}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \\ & \quad + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \end{aligned}$$

$$\begin{aligned}
& + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_6+1}^H(a_0) \right] \\
& + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_4+1}^H(a_0) \right] \\
& + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_3+1-j_2}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \\
& + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \\
& + \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_2+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_4+1}^H(a_0) \right] \Bigg\}.
\end{aligned}$$

We can use this last remark and Equations (46) and (39) to show that

$$\begin{aligned}
\mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^3 \right] & = \left( \frac{1}{\sqrt{n}} \sum_{j_1=1}^n a_0^{j_1-1} \sum_{r_1=j_1}^n \sum_{r_2=1}^n \left( \Omega_{n,H}^{-1} \right)_{r_1, r_2} \mathbb{E} \left[ e_{r_1+1-j_1}^H(a_0) e_{r_2+1}^H(a_0) \right] \right) \\
& \times \left( \frac{1}{n} \sum_{j_2, j_3=1}^n a_0^{j_2+j_3-2} \sum_{r_3=j_2}^n \sum_{r_4=1}^n \sum_{r_5=j_3}^n \sum_{r_6=1}^n \left\{ \left( \Omega_{n,H}^{-1} \right)_{r_3, r_4} \left( \Omega_{n,H}^{-1} \right)_{r_5, r_6} \right. \right. \\
& \times \left[ \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_4+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_5+1-j_3}^H(a_0) e_{r_6+1}^H(a_0) \right] \right. \\
& + \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_6+1}^H(a_0) \right] \\
& \left. \left. + \mathbb{E} \left[ e_{r_3+1-j_2}^H(a_0) e_{r_6+1}^H(a_0) \right] \mathbb{E} \left[ e_{r_4+1}^H(a_0) e_{r_5+1-j_3}^H(a_0) \right] \right\} \right) \\
& = \mathbb{E} \left[ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right] \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^2 \right].
\end{aligned}$$

In view of (39), (42) and Remarks 5 and 6, one can rewrite the  $r$ -th moment of the random variable  $Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}$  when  $r$  is odd in the form:

$$\mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^r \right] = \mathbb{E} \left[ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right] \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^{r-1} \right]. \quad (53)$$

The calculation given in the proof of Lemma 4 to show the convergence to 0 of the sequence  $(\mathbb{E}[Z_{n,H}^T U_{n,H}(a_0)/n])_{n \geq 1}$  remains valid when we multiply the general term of the last sequence by  $\sqrt{n}$  and when we use the fractional version of Cesàro's Lemma<sup>1</sup> to conclude.

In fact, from (43) we have

$$\left| \mathbb{E} \left[ \frac{1}{\sqrt{n}} \left\{ Z_{n,H}^T U_{n,H}(a_0) \right\} \right] \right| \leq K \left( \frac{1}{\sqrt{n}} \sum_{j=1}^n |a_0|^{j-1} + \frac{1}{\sqrt{n}} \sum_{j=1}^n |a_0|^{j-1} \sum_{\substack{1 \leq r \leq n \\ r \neq j}} \frac{1}{(r-j)^{2H}} \frac{1}{r^{2-2H}} \right). \quad (54)$$

<sup>1</sup>Recall that the fractional version of Cesàro's Lemma states that for  $(h_t)_{t \geq 0}$  a sequence of positive real numbers,  $\kappa > 0$  and  $c \geq 0$  we have

$$\lim_{t \rightarrow \infty} h_t t^{1-\kappa} = |\kappa| c \Rightarrow \lim_{n \rightarrow \infty} \frac{1}{n^\kappa} \sum_{t=0}^n h_t = c.$$

Since  $\sqrt{j} |a_0|^{j-1} \rightarrow 0$  when  $j \rightarrow \infty$ , fractional version of Cesàro's Lemma implies that the first term in the right hand side of (54) tends to 0 when  $n$  goes to infinity. Based on the calculation given in the proof of Lemma 4 and using the fractional version of Cesàro's Lemma, the second term in the right hand side of (54) verifies

$$\begin{aligned} & \frac{1}{\sqrt{n}} \sum_{j=1}^n |a_0|^{j-1} \sum_{\substack{1 \leq r \leq n \\ r \neq j}} \frac{1}{(r-j)^{2H}} \frac{1}{r^{2-2H}} \\ & \leq \frac{1}{\sqrt{n}} \sum_{r=1}^n \frac{1}{r^{1-2H}} |a_0|^{r-1} \max \left( \frac{1}{|a_0|}, \frac{1}{|a_0|^r r^{2H}} \right) + \frac{1}{\sqrt{n}} \sum_{r=1}^n \frac{1}{r^{2-2H}} |a_0|^r \left( \frac{1 - |a_0|^{n-r}}{1 - |a_0|} \right) \\ & \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

Then the sequence  $(\mathbb{E}[Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}])_{n \geq 1}$  converges to 0.

Furthermore, Equation (45) and Lemma 1 imply that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^2 \right] = \lim_{n \rightarrow \infty} \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] = \frac{1}{1 - a_0^2}.$$

Using (53) we deduce that

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^3 \right] = 0. \quad (55)$$

When  $r = 4$ , the set  $\{k_1 + 1 - j_1, k_2 + 1, k_3 + 1 - j_2, k_4 + 1, \dots, k_{2r-1} + 1 - j_r, k_{2r} + 1\}$  can be partitioned by pairs in two different ways. The first is to consider partitions that contain at least one block of type  $\{k_{2i-1} + 1 - j_i, k_i + 1\}$  and the second is to take partitions that do not contain any such block. In the first case, we use the same remarks introduced to obtain (53) and we factorize by the expectation of  $Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}$ . In the second case, we use the result stated in Lemma 1. More precisely, we have

$$\mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^4 \right] = 3 \left( \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] \right)^2 + \mathbb{E} \left[ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right] \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^3 \right].$$

Therefore, in view of Lemma 1 and (55) we have

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^4 \right] = \frac{3}{(1 - a_0^2)^2}.$$

We can generalize this result by an immediate recurrence on  $r$  to obtain

$$\begin{aligned} & \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^r \right] \\ & = \begin{cases} \mathbb{E} \left[ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right] \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^{r-1} \right] & \text{if } r \text{ is odd} \\ (r-1)!! \left( \mathbb{E} \left[ \frac{1}{n} \|Z_{n,H}\|_{\mathbb{R}^n}^2 \right] \right)^{r/2} + \mathbb{E} \left[ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right] \mathbb{E} \left[ \left\{ \frac{1}{\sqrt{n}} Z_{n,H}^T U_{n,H}(a_0) \right\}^{r-1} \right] & \text{if } r \text{ is even.} \end{cases} \end{aligned}$$

This implies that the  $r$ -th moment of the random variable  $Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}$  converges to

$$\begin{cases} 0 & \text{if } r \text{ is odd} \\ (r-1)!! \left(\frac{1}{1-a_0^2}\right)^{r/2} & \text{if } r \text{ is even,} \end{cases} \quad (56)$$

and this is exactly the result recalled in (50). The lemma is then proved. □

The convergences established in Lemmas 3 and 5 imply (by Slutsky's theorem) that the vector  $(Z_{n,H}^T U_{n,H}(a_0)/\sqrt{n}, \|Z_{n,H}\|_{\mathbb{R}^n}^2/n)$  converges in distribution to  $(G_2, 1/(1-a_0^2))$ , where  $G_2$  is the random variable defined in the proof of Lemma 5. Continuous mapping theorem allows to complete the proof of Theorem 2.

6. Figures

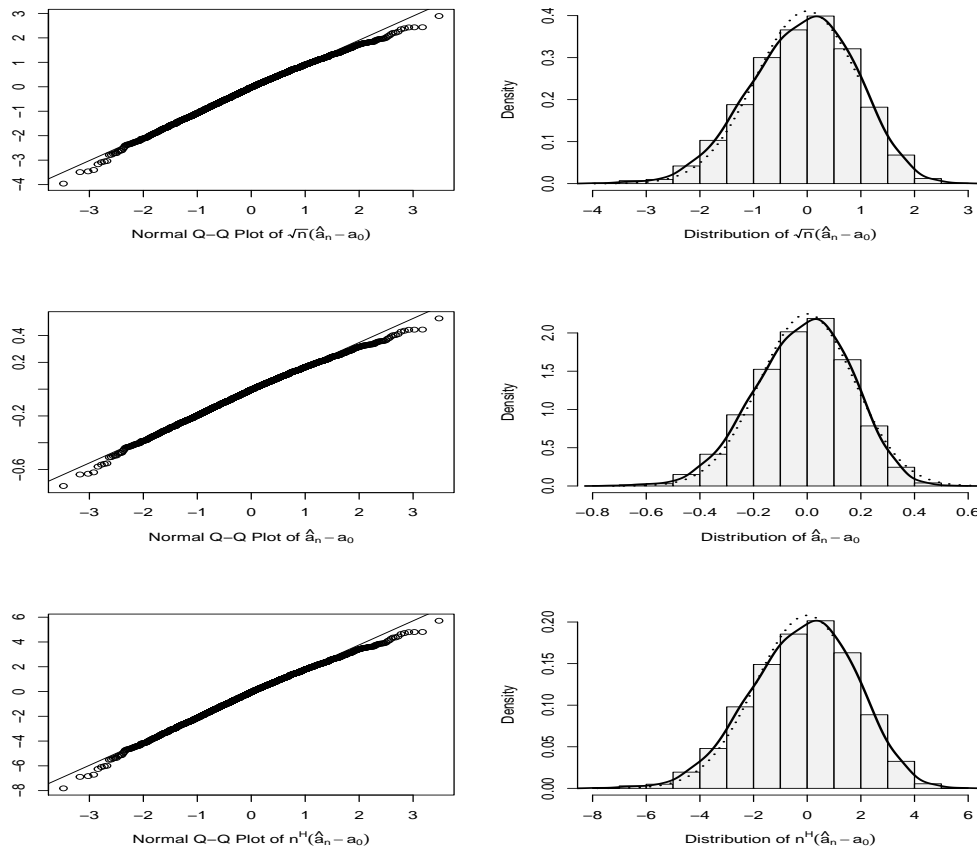


FIG 1. Distribution of the generalized least squares estimator of  $N = 2,000$  independent simulations of the fractional autoregressive model (14) with size  $n = 30$  and unknown parameter  $a = 0.3$ . The Hurst index is assumed to be equal to 0.7.

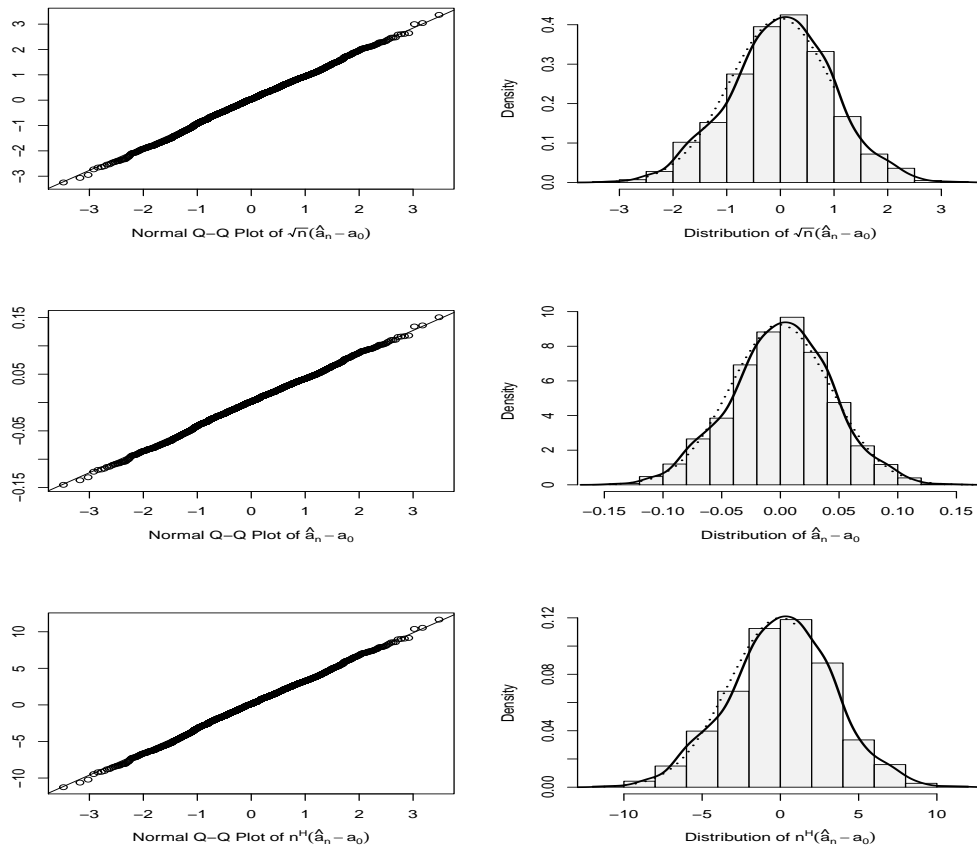


FIG 2. Distribution of the generalized least squares estimator of  $N = 2,000$  independent simulations of the fractional autoregressive model (14) with size  $n = 500$  and unknown parameter  $a = 0.3$ . The Hurst index is assumed to be equal to 0.7.

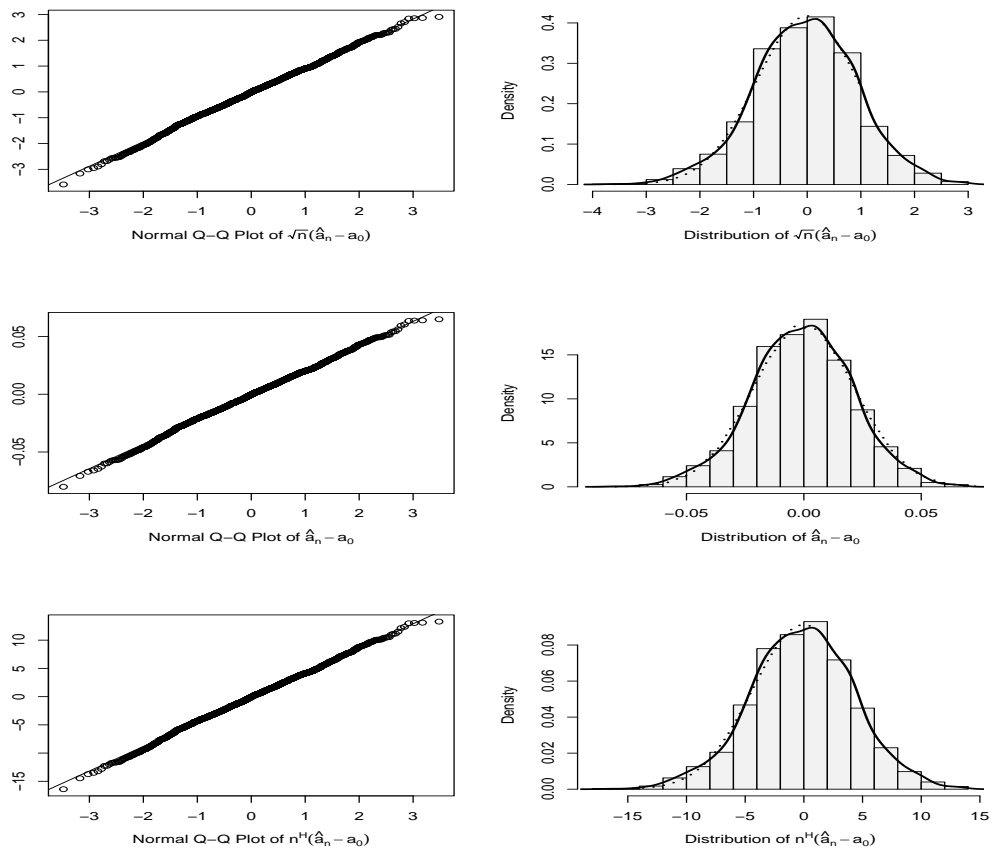


FIG 3. Distribution of the generalized least squares estimator of  $N = 2,000$  independent simulations of the fractional autoregressive model (14) with size  $n = 2,000$  and unknown parameter  $a = 0.3$ . The Hurst index is assumed to be equal to 0.7.

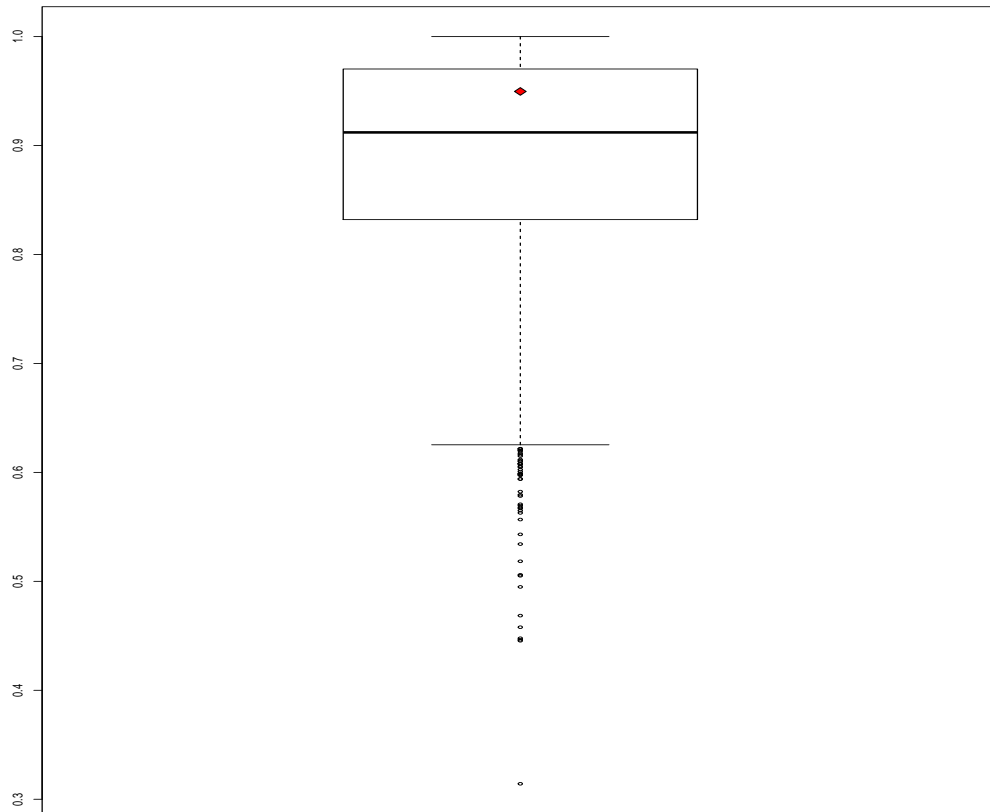


FIG 4. Box plot describing the distribution of the variance estimates of the generalized least squares estimator over  $N = 2,000$  independent simulations of the fractional autoregressive model (14) with size  $n = 30$  and unknown parameter  $a = 0.3$ . The Hurst index is assumed to be equal to 0.7. The diamond symbol represent the mean of the standardized errors  $n(\hat{a}_n - 0.3)^2$ .



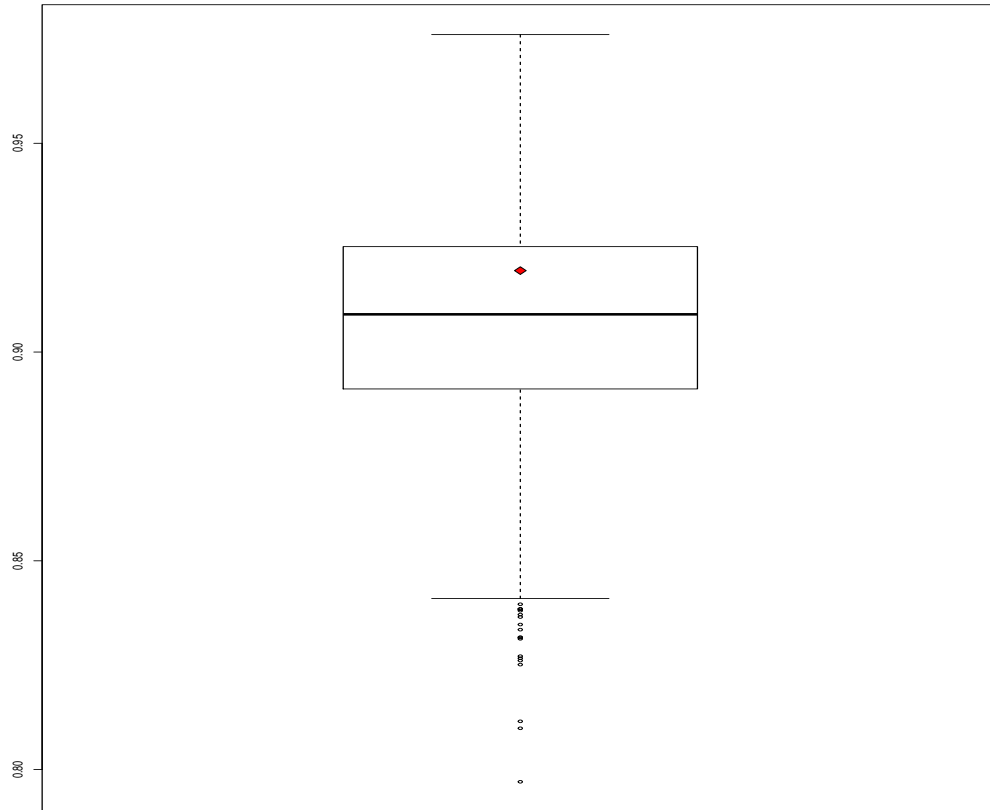


FIG 5. Box plot describing the distribution of the variance estimates of the generalized least squares estimator over  $N = 2,000$  independent simulations of the fractional autoregressive model (14) with size  $n = 500$  and unknown parameter  $a = 0.3$ . The Hurst index is assumed to be equal to 0.7. The diamond symbol represent the mean of the standardized errors  $n(\hat{a}_n - 0.3)^2$ .

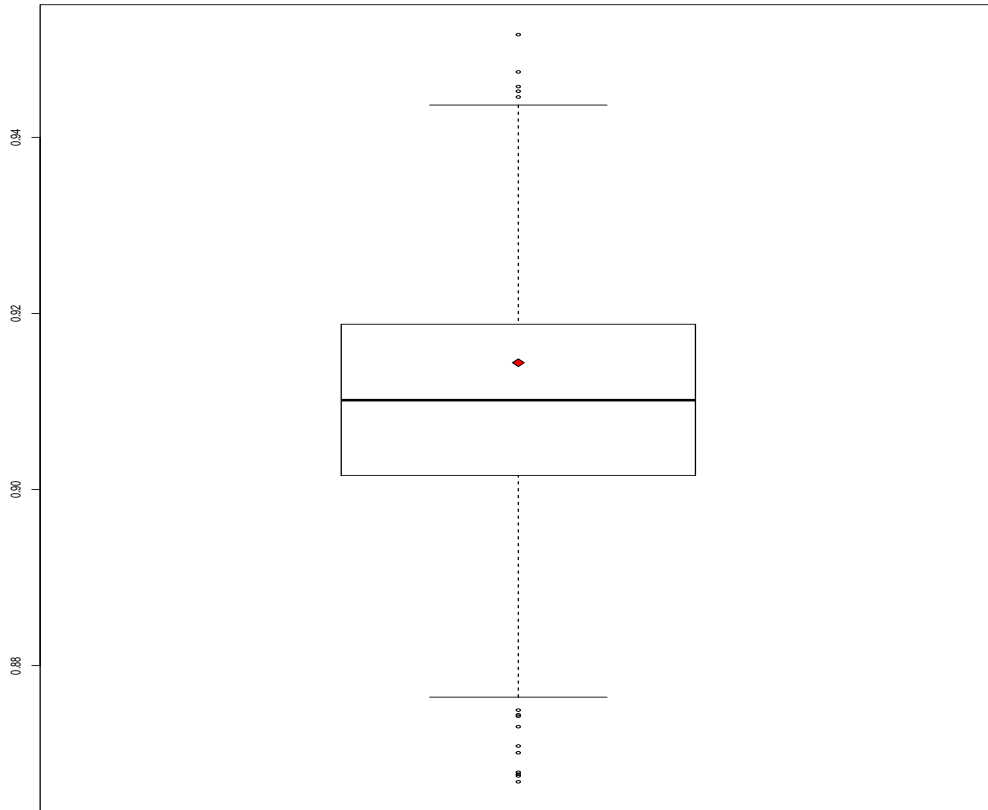


FIG 6. Box plot describing the distribution of the variance estimates of the generalized least squares estimator over  $N = 2,000$  independent simulations of the fractional autoregressive model (14) with size  $n = 2,000$  and unknown parameter  $a = 0.3$ . The Hurst index is assumed to be equal to 0.7. The diamond symbol represent the mean of the standardized errors  $n(\hat{a}_n - 0.3)^2$ .

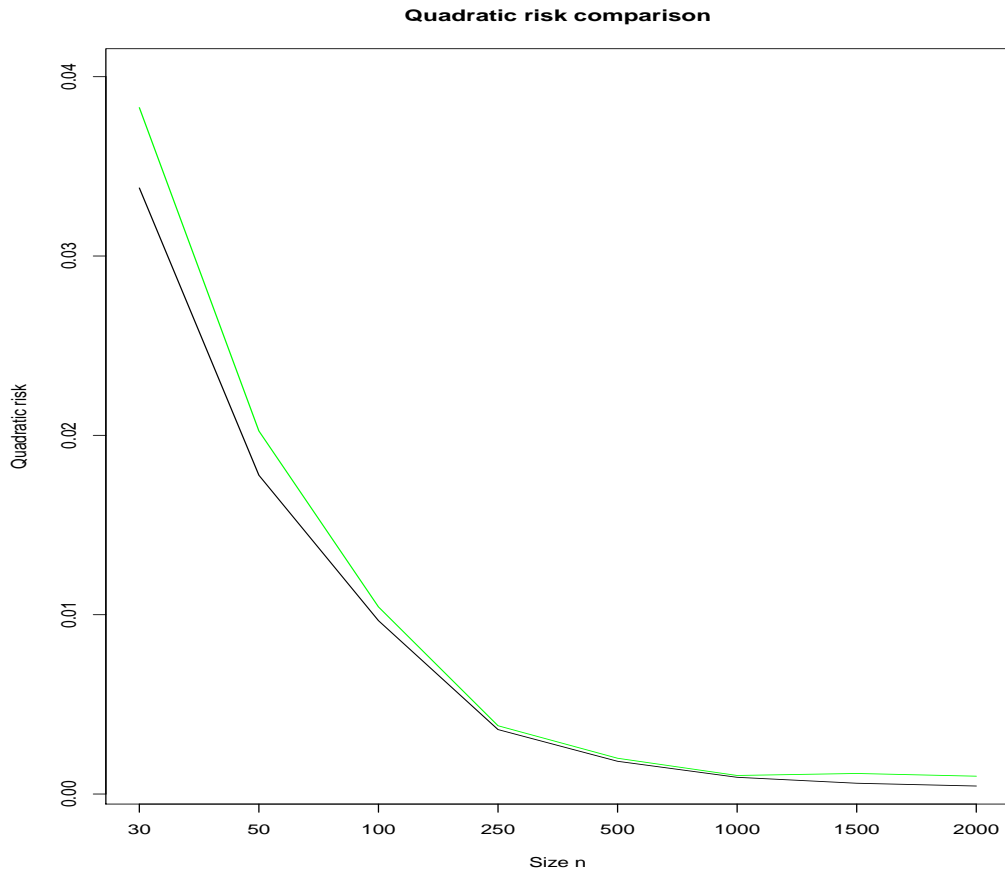


FIG 7. Comparison of quadratic risk of the generalized least squares estimator (in black) and Whittle estimator (in green) of  $N = 2,000$  independent simulations of the fractional autoregressive model (14) with  $a = 0.3$  and with  $H = 0.7$ . The sample size  $n$  varies and takes its values from the set  $\{30, 50, 100, 250, 500, 1000, 1500, 2000\}$ .

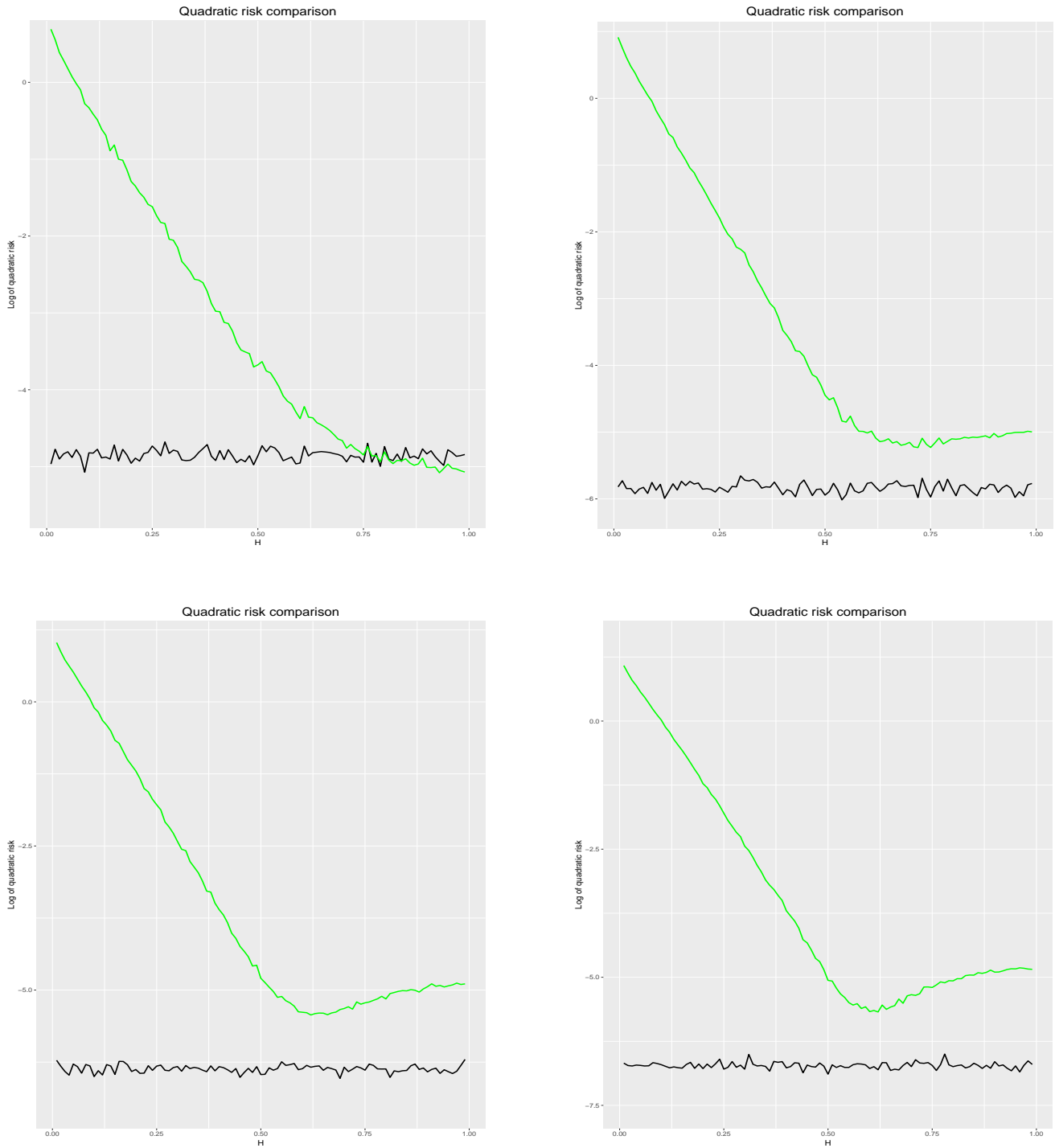


FIG 8. Comparison of the logarithm of quadratic risk of the generalized least squares estimator (in black) and Whittle estimator (in green) of  $N = 1,000$  independent simulations of the fractional autoregressive model (14) with  $a = -0.9$  and with different values of  $H$ . The sample size  $n$  is (from left to right and from top to bottom)  $n = 50, 100, 150, 200$ .

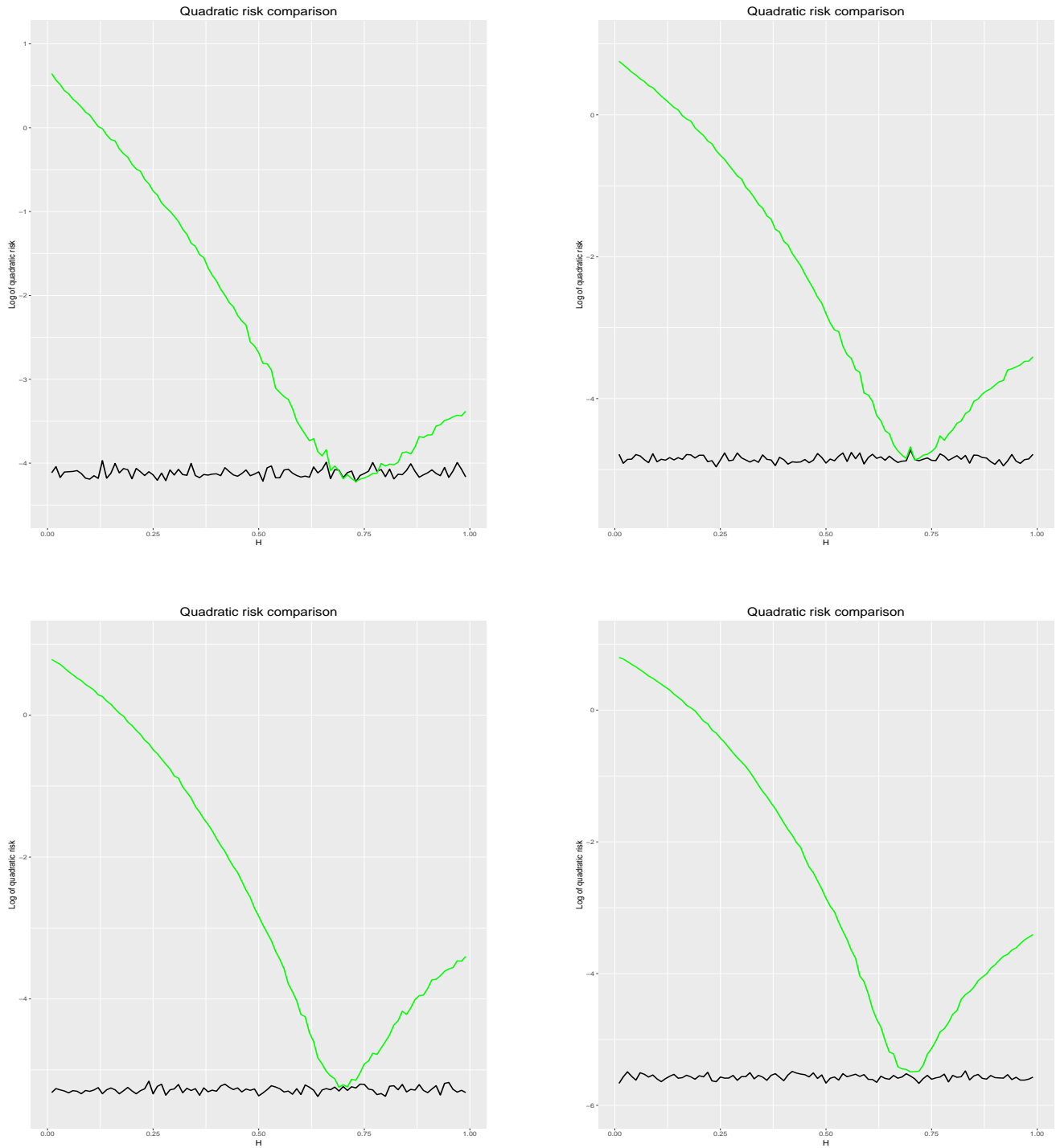


FIG 9. Comparison of the logarithm of quadratic risk of the generalized least squares estimator (in black) and Whittle estimator (in green) of  $N = 1,000$  independent simulations of the fractional autoregressive model (14) with  $a = -0.5$  and with different values of  $H$ . The sample size  $n$  is (from left to right and from top to bottom)  $n = 50, 100, 150, 200$ .

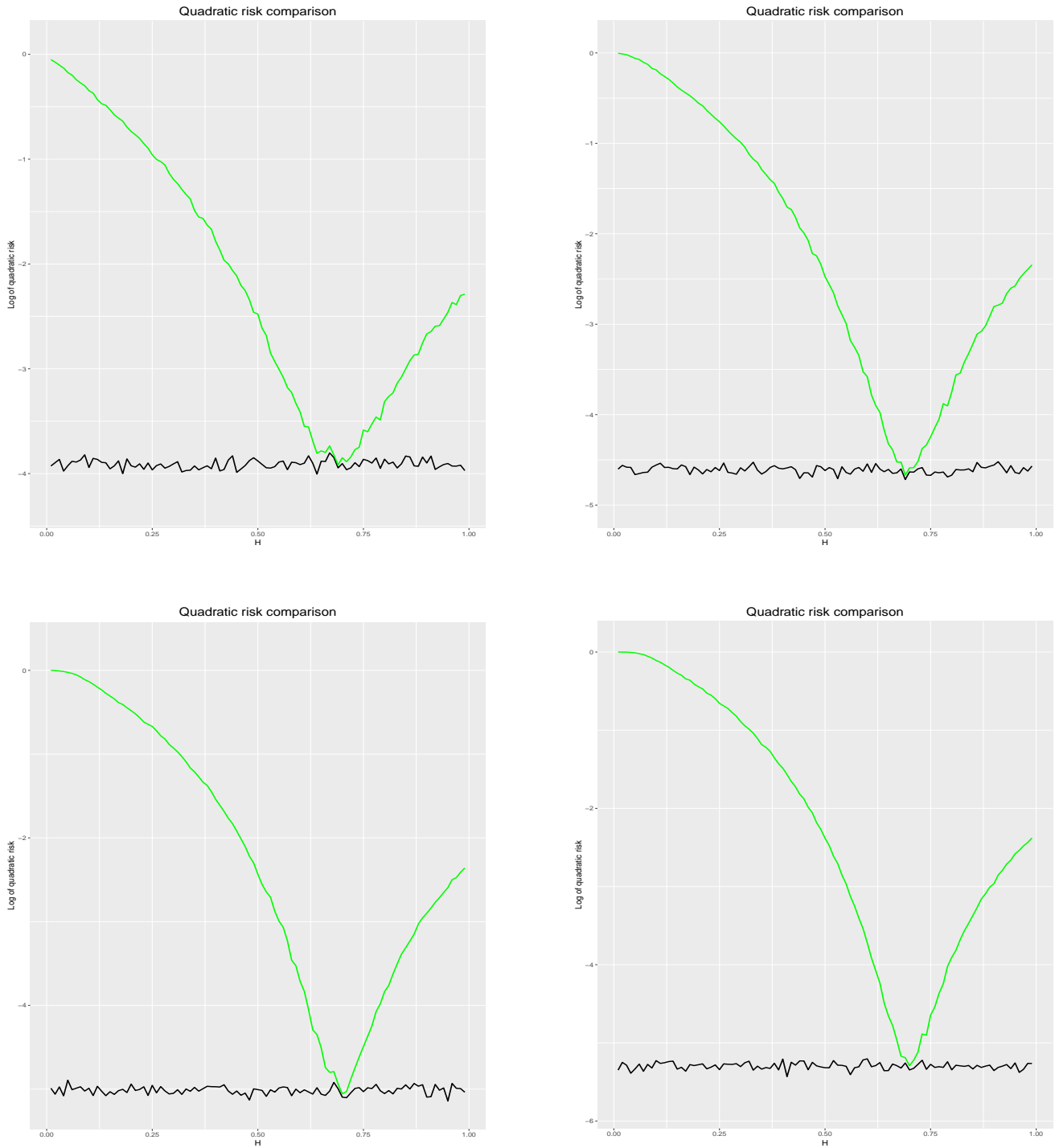


FIG 10. Comparison of the logarithm of quadratic risk of the generalized least squares estimator (in black) and Whittle estimator (in green) of  $N = 1,000$  independent simulations of the fractional autoregressive model (14) with  $a = 0$  and with different values of  $H$ . The sample size  $n$  is (from left to right and from top to bottom)  $n = 50, 100, 150, 200$ .

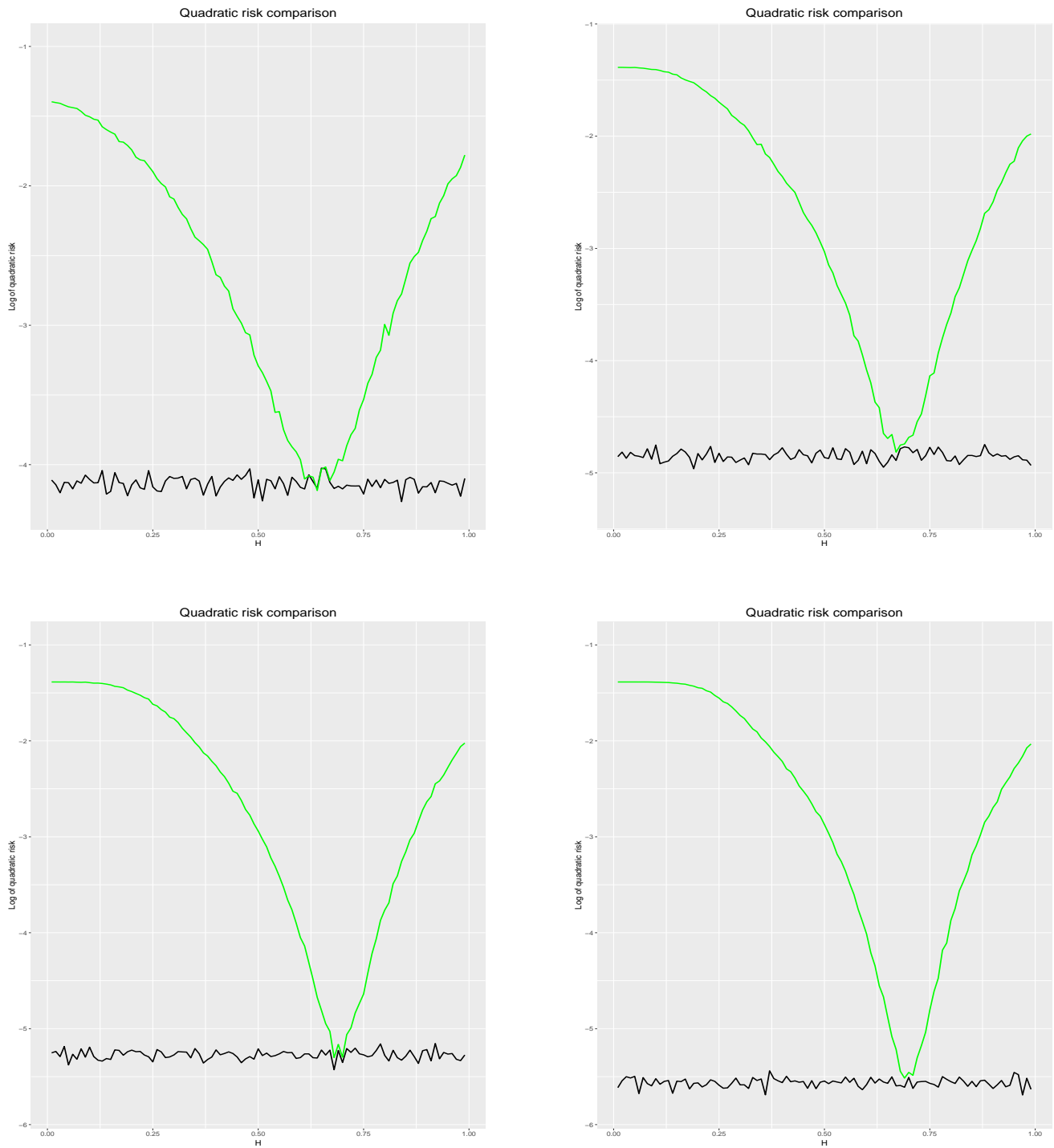


FIG 11. Comparison of the logarithm of quadratic risk of the generalized least squares estimator (in black) and Whittle estimator (in green) of  $N = 1,000$  independent simulations of the fractional autoregressive model (14) with  $a = 0.5$  and with different values of  $H$ . The sample size  $n$  is (from left to right and from top to bottom)  $n = 50, 100, 150, 200$ .

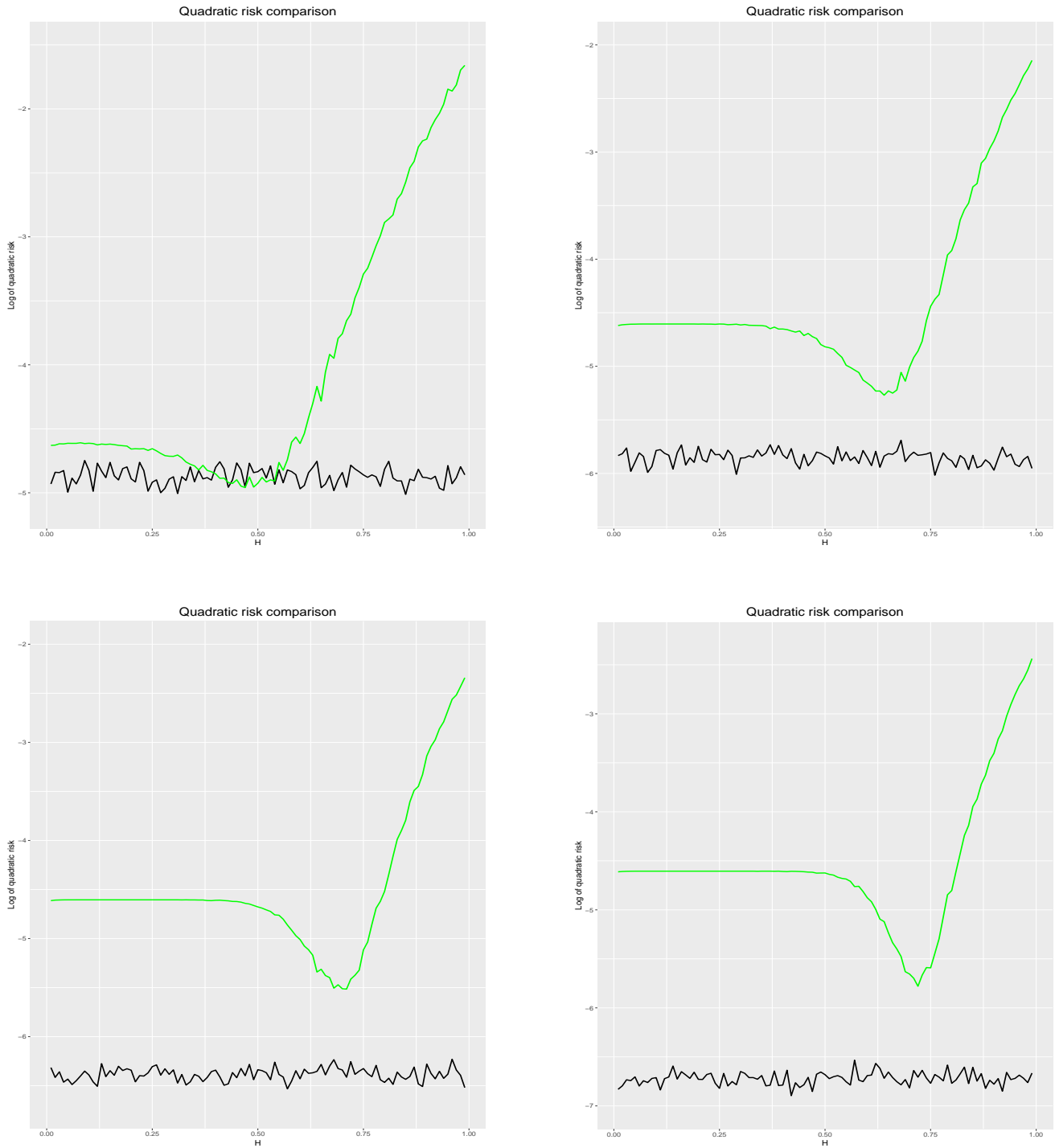


FIG 12. Comparison of the logarithm of quadratic risk of the generalized least squares estimator (in black) and Whittle estimator (in green) of  $N = 1,000$  independent simulations of the fractional autoregressive model (14) with  $a = 0.9$  and with different values of  $H$ . The sample size  $n$  is (from left to right and from top to bottom)  $n = 50, 100, 150, 200$ .



## Acknowledgments

The author acknowledges the support of the ANR project "Efficient inference for large and high-frequency data" (ANR-21-CE40-0021).

## References

- Ablowitz, M. J. and Fokas, A. S. (2003). *Complex Variables: Introduction and Applications*. Cambridge Texts in Applied Mathematics. Cambridge University Press, 2 edition.
- Anderson, T. W. and Taylor, J. B. (1976). Strong consistency of least squares estimates in normal linear regression. *Ann. Statist.*, 4(4):788–790.
- Anderson, T. W. and Taylor, J. B. (1979). Strong consistency of least squares estimates in dynamic models. *Ann. Statist.*, 7(3):484–489.
- Ben Hariz, S., Brouste, A., Cai, C., and Soltane, M. (2021). Fast and Asymptotically-efficient estimation in a Fractional autoregressive process. working paper or preprint.
- Beran, J., Feng, Y., Ghosh, S., and Kulik, R. (2013). *Long-memory processes*. Springer, Heidelberg. Probabilistic properties and statistical methods.
- Billingsley, P. (1986). *Probability and measure*. Wiley series in probability and mathematical statistics. Wiley.
- Brillinger, D. R. (1981). *Time series: data analysis and theory*, volume 36. Siam.
- Brouste, A., Cai, C., and Kleptsyna, M. (2014). Asymptotic properties of the MLE for the autoregressive process coefficients under stationary gaussian noise. *Mathematical Methods of Statistics*, 23(2):103–115.
- Dahlhaus, R. (1989). Efficient parameter estimation for self-similar processes. *Ann. Statist.*, 17(4):1749–1766.
- Fox, R. and Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *Ann. Statist.*, 14(2):517–532.
- Francq, C. and Zakoïan, J.-M. (1998). Estimating linear representations of nonlinear processes. *J. Statist. Plann. Inference*, 68(1):145–165.
- Geweke, J. and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4):221–238.
- Graf, H.-P. (1980). *Long-range correlations and estimation of the self-similarity parameter*. Ph.D. dissertation, Swiss Federal Institute of Technology.
- Granger, C. W. J. and Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *J. Time Ser. Anal.*, 1(1):15–29.
- Hosking, J. R. M. (1981). Fractional differencing. *Biometrika*, 68(1):165–176.
- Ito, K. (1944). On the ergodicity of a certain stationary process. *Proc. Imp. Acad.*, 20(2):54–55.
- Khinchin, A. I. (1949). *Mathematical foundations of statistical mechanics* dover publications. Inc. New York.
- Lai, T., Robbins, H., and Wei, C. (1979). Strong consistency of least squares estimates in multiple regression ii. *Journal of Multivariate Analysis*, 9(3):343 – 361.
- Lai, T. L. and Robbins, H. (1977). Strong consistency of least-squares estimates in regression models. *Proceedings of the National Academy of Sciences*, 74(7):2667–2669.
- Lai, T. L., Robbins, H., and Wei, C. Z. (1978). Strong consistency of least squares estimates in multiple regression. *Proceedings of the National Academy of Sciences*, 75(7):3034–3036.
- Mandelbrot, B. and Taqqu, M. (1979). *Robust R/S Analysis of Long Run Serial Correlation*. IBM Thomas J. Watson Research Division.

- Mandelbrot, B. B. (1975). Limit theorems on the self-normalized range for weakly and strongly dependent processes. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 31(4):271–285.
- Mandelbrot, B. B. and Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, 10:422–437.
- McLeod, A. I. and Hipel, K. W. (1978). Preservation of the rescaled adjusted range: 1. a reassessment of the hurst phenomenon. *Water Resources Research*, 14(3):491–508.
- Miller, P. D. (2006). *Applied asymptotic analysis*, volume 75. American Mathematical Soc.
- Mohr, D. (1981). *Modeling data as a fractional Gaussian noise*. Ph.D. dissertation, Princeton University.
- Palma, W. (2007). *Long-memory time series*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ. Theory and methods.
- Sinai, Y. (1976). Self-similar probability distributions. *Theory of Probability & Its Applications*, 21(1):64–80.
- Solo, V. (1981). Strong consistency of least squares estimators in regression with correlated disturbances. *Ann. Statist.*, 9(3):689–693.
- Sweeting, T. J. (1980). Uniform asymptotic normality of the maximum likelihood estimator. *The Annals of Statistics*, 8(6):1375–1381.
- Watson, G. N. (1918). The harmonic functions associated with the parabolic cylinder. *Proceedings of the London Mathematical Society*, s2-17(1):116–148.
- Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. *Ann. Statist.*, 16(2):791–807.

## Contents

1	Introduction	1
2	Generalized least squares estimation of FrAR models of order 1	3
2.1	Statement of the problem and notations	3
2.2	Asymptotic properties	5
3	Generalized least squares estimation of FrAR models of order $p$	5
4	Numerical illustrations	6
4.1	Simulation studies	7
5	Proofs	7
5.1	Preliminary results	8
5.1.1	The expression of $Z_{n,H}$ as a function of $U_{n,H}(a_0)$	8
5.1.2	Asymptotic properties of the components of the inverse covariance matrix of $(\epsilon_t^H)_{1 \leq t \leq n}$	9
5.2	Proof of Theorem 1	11
5.2.1	Proof of the convergence in mean square of $(\ Z_{n,H}\ _{\mathbb{R}^n}^2/n)_{n \geq 1}$ to $1/(1 - a_0^2)$	11
5.2.2	Proof of the convergence in mean square of $(Z_{n,H}^T U_{n,H}(a_0)/n)_{n \geq 1}$ to 0	16
5.3	Proof of Theorem 2	22
6	Figures	28
	Acknowledgments	40
	References	40