Parameter estimation for fractional autoregressive process with seasonal structure

Chunhao Cai & Yiwu Shang

Published online: 05 Aug 2025.

Submit your article to this journal ↗

Article views: 96

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Parameter estimation for fractional autoregressive process with seasonal structure

Chunhao Cai[a] and Yiwu Shang[b]

[a]School of Mathematics (Zhuhai), Sun Yat-sen University, Zhuhai, People's Republic of China; [b]School of Mathematical Sciences, Nankai University, Tianjin, People's Republic of China

ABSTRACT

This paper introduces a new kind of seasonal fractional autoregressive process (SFAR) driven by fractional Gaussian noise (fGn). The new model includes a standard seasonal AR model and fGn. The estimation of the parameters of this new model has to solve two problems: non-stationarity from the seasonal structure and long memory from fGn. We innovatively solve these by getting a stationary subsequence, making a stationary additive sequence and then obtaining their spectral density. Then we use one-step procedure for Generalized Least Squares Estimator (GLSE) and the Geweke Porter–Hudak (GPH) method to get better results. We prove that both the initial and one-step estimators are consistent and asymptotically normal. Finally, we use Monte Carlo simulations with finite-sized samples to demonstrate the performance of these estimators. Moreover, through empirical analysis, it is shown that the SFAR model can simulate some real-world phenomena better than general models.

1. Introduction

The long memory phenomenon and seasonal phenomenon play important roles in economics, geography and other fields. One classic type of seasonal model with long memory is the ARFISMA$(p, d, q) \times (P, D, Q)_s$ process, which has been extensively researched by Hosking (1984), Franco and Reisen (2007) and Chan and Terrin (1995). As demonstrated below, these models are equivalent to the ARUMA model, which can be expressed as

$$G(\mathcal{B})X_n = \epsilon_n, \tag{1}$$

where $\mathcal{B}$ is the lag operator, ϵ_n is a short memory process and $G(\cdot)$ satisfies the recurrence relation

$$G(z) = (1-z)^{d_0} \left\{ \prod_{j=1}^{r-1} (1 - 2z\cos\lambda_j + z^2)^{d_j} \right\} (1+z)^{d_r}, \tag{2}$$

where $|d_j| \le \frac{1}{2}$, and for $j = 0, 1, 2, \ldots, r$, denoting that r is a positive integer, the frequencies satisfy $0 \le \lambda_j \le \pi$. Additionally, $\lambda_0 = 0$ and $\lambda_r = \pi$.

In the present paper, we conduct a comprehensive study on a p-order seasonal fractional

CONTACT Yiwu Shang ✉ shangyiwu@mail.nankai.edu.cn 🖅 School of Mathematical Sciences, Nankai University, Weijin Road 94, Nankai District, Tianjin 300071, People's Republic of China

autoregressive process (SFAR), denoted as X_{nT+u}. Here, the nonnegative integer T signifies the number of seasons, suggesting that the time series data exhibits seasonal fluctuations with a period of T. For example, when $T = 4$, it represents a quarterly seasonal pattern, and when $T = 12$, it corresponds to a monthly seasonal pattern. For any $n \in \mathbb{N}$ (where n denotes the count of complete seasonal cycles), the model adheres to the following recursive relation:

$$X_{nT+u} = \sum_{i=1}^{p} \phi_i(nT + u)X_{(n-i)T+u} + \epsilon_{nT+u}^{H}, \quad u = 1, 2, \ldots, T, \tag{3}$$

where u denotes the specific time points within each seasonal cycle, thus taking values from 1 to T, and p represents the order of the autoregressive part of the model. The $\phi_i(nT + u)$ are autoregressive seasonal coefficients, which may change with time and satisfy $\phi_i(u) = \phi_i(u + T)$. ϵ_{nT+u}^{H} represents fractional Gaussian noise, which explains the nonseasonal fluctuations. Fractional Gaussian noise exhibits long memory when $\frac{1}{2} < H < 1$. The long memory phenomenon indicates strong autocorrelation or dependence in time series data. We typically say that X_t has long memory if its covariance satisfies

$$\gamma_j \sim Q j^{2H-2}, \quad j \to \infty, \tag{4}$$

and the spectral density is defined by the scheme

$$f(\lambda) \sim V \lambda^{1-2H}, \quad \lambda \to 0^+, \tag{5}$$

where $\frac{1}{2} < H < 1$, Q and V are constants greater than 0. Robinson (2010), Bisognin and Lopes (2009) and Beran et al. (2013) did a great deal of detailed and excellent work in fractional Gaussian noise (fGn), especially in the estimation of H.

The Seasonal Fractional Autoregressive (SFAR) model represents a natural expansion of the fractional autoregressive process (FAR). The FAR recognized as a long memory model is formulated as

$$X_n = \sum_{i=1}^{p} a_i X_{n-i} + \epsilon_n^{H}, \quad n \in \mathbb{N}, \tag{6}$$

where $a_i \in \mathbb{R}$. It is composed of fractional Gaussian noise, and its long-range dependence characteristics are determined by the value of H.

In this paper, we focus on the estimation and asymptotic properties of parameters in the SFAR model. Geweke and Porter-Hudak (1983) and Carlin and Dempster (1989) conducted some research on such models in the early stage. We aim to extend related research and will use a one-step procedure to optimize our approach.

For the parameter estimation of SFAR model, two key problems need to be addressed: the nonstationarity resulting from the seasonal structure and the dependence within the fractional Gaussian noise.

Seasonality is a distinctive feature of time series data where patterns repeat at regular intervals, typically defined by a specific period T. Seasonal time series models are often nonstationary, which present certain challenges for our research. A common solution is to perform seasonal differencing on the time series. Seasonal autoregressive process is a classical model proposed by Harrison (1965) and Chatfield and Prothero (1973). Tsay (2013) has explored SAR model with white noise in detail. However, there still remain many interesting

variations worthy of research. For instance, the study in Kong and Lund (2023) employed particle filtering likelihood methods to estimate seasonal count time series. The other category is the research on the SAR model driven by fGn.

Previous studies by Brouste et al. (2014) and Soltane (2024) have laid a foundation for the estimation of the parameters $\phi_i(u)$ in FAR models. In this paper, we use the modified Generalized Least Squares Estimation (GLSE) proposed by Esstafa (2019) and Hariz et al. (2024) to obtain a consistent estimator of $\phi_i(u)$. Additionally, we will prove that this estimator is asymptotically normal.

Time series models with long memory show long-range dependencies between distant observations, posing challenges to traditional statistical analysis and forecasting. In the SFAR model, long memory comes from fractional Gaussian noise, where the parameter H determines this characteristic. Thus estimating H is crucial. The first method for estimating H was the rescaled range analysis by Hurst (1951), but its lack of a limiting distribution complicates statistical inference. Now, popular estimation techniques are the GPH estimation by Geweke and Porter-Hudak (1983) and the local Whittle estimation by Robinson (1995).

For the estimation of the Hurst index H, we will adopt the Geweke Porter–Hudak (GPH) method, which exhibits a smaller bias, for an additive stationary time series derived from the samples. It is worth noting that it would be more straightforward to estimate $\widehat{H}_n(u)$ by $(X_{nT+u})_{n\in\mathbb{N}}$. However, this approach is not fundamentally different from the method in Hariz et al. (2024) and each $\widehat{H}_n(u)$ cannot contain information about all the data. Meanwhile, considering that sequence $(X_{nT+u})_{n\in\mathbb{N},u=1,\ldots,T}$ represents data of the same nature, we assume that the long memory parameter is the same for each season and is independent of the season u, and the differences between different seasons are only determined by the seasonal parameters. To obtain a unique $\widehat{H}_n$, we sum up the data in each cycle to obtain a new sequence $(Y_n)_{n\in\mathbb{N}}$, and then we prove the stationarity of $(Y_n)_{n\in\mathbb{N}}$, calculate its spectral density and finally use the GPH method to get $\widehat{H}_n(u)$. This improvement enables us to address the issue of parameter estimation for H in nonstationary time series with seasonality.

After obtaining the initial estimators of $\phi_i(u)$ and H, we modify our approach using a faster and asymptotically efficient method known as the one-step estimator. This method, first proposed by Le Cam (1956), has been widely applied in ergodic Markov chains (Kutoyants & Motrunich, 2016), diffusion processes (Gloter & Yoshida, 2021) and fractional autoregressive processes (Hariz et al., 2024). The primary challenge lies in calculating the Fisher information matrix, as discussed in Cohen et al. (2013). To tackle this issue, we extract the data from each season to form a new series, proving the stationarity of this new series and deriving its spectral density. Subsequently, we can utilize the results from Cohen et al. (2013) and Hariz et al. (2024) to obtain related findings.

This paper is organized as follows. Sections 2 and 3 present the main results. Section 2 introduces the initial estimator of the Hurst index $\phi_i(u)$ and discusses its asymptotic properties. Section 3 derives the one-step estimator and its asymptotic properties. Section 4 provides numerical illustrations to demonstrate the performance of both the initial and one-step estimators. Section 5 concludes this paper and considers the prospects and significance of our research. Section 6 illustrates that the SFAR model is superior to the traditional seasonal autoregressive model through a practical application. All technical proofs are gathered in Section 7, while Section 8 presents auxiliary results.

2. Initial estimator of SFAR(1) models

2.1. Problem statements and assumptions

Without loss of generality, based on the representation of the SFAR model in (3), we can consider the first-order model in this paper and denote $\phi_1(u) = \phi(u)$.

X_{nT+u} is said to be an SFAR(1) model if it admits the representation

$$X_{nT+u} = \phi(nT + u)X_{(n-1)T+u} + \epsilon_{nT+u}^H, \quad u = 1, 2, 3, \ldots, T, \ n \in \mathbb{N}, \tag{7}$$

where $\phi(u) = \phi(u + nT)$, T represents the season length and u denotes the uth season of the nth cycle. The term ϵ_{nT+u}^H represents a stationary fractional Gaussian noise with a Hurst index H. It is defined as the increment of the fractional Brownian motion, specifically $\epsilon_{nT+u}^H = B_{nT+u+1}^H - B_{nT+u}^H$, where B_{nT+u}^H is the fractional Brownian motion. The autocovariance of sequence $(\epsilon_n^H)_{n \in \mathbb{N}}$ takes the form of

$$\rho(k) = \frac{1}{2}\left(|k+1|^{2H} - 2|k|^{2H} + |k-1|^{2H}\right), \tag{8}$$

the spectral density of $(\epsilon_n^H)_{n \in \mathbb{N}}$ defined by

$$f_{\epsilon_n^H}(\lambda) = C_H(1 - \cos(\lambda)) \sum_{j \in \mathbf{Z}} \frac{1}{|\lambda + 2j\pi|^{2H+1}}, \tag{9}$$

where $C_H = \frac{1}{2\pi}\Gamma(2H+1)\sin(\pi H)$ and $\lambda \in [-\pi, \pi]$, $\Gamma(\cdot)$ is Gamma function.

Here are some assumptions and notations below.

A_0: Denote Θ_u^{l*} as a compact set with the following expression:

$$\Theta_u^{l*} = \{\phi(u) \in \mathbb{R} : \text{ the roots of } 1 - \phi(u)z = 0 \text{ have modulus } \geq 1 + l\}.$$

We define the set Θ_u^l as the Cartesian product $\Theta_u^{l*} \times [d_1, d_2]$, where l is a positive constant and $[d_1, d_2] \in (0, 1)$.

A_1: $\phi(u) \in (-1, 1)$ and $H \in (0, 1)$.

Notation: By $\xrightarrow{\mathcal{L}}$ and $\xrightarrow{\mathbb{P}}$, respectively, we denote convergence in law and convergence in probability. Let $\underline{\phi} = (\phi(1), \phi(2), \ldots, \phi(T))$. Denote the parameters $\theta(u) = (\phi(u), H)$, where $\theta(u) \in \Theta_u^l$, and $\overset{\circ}{\Theta}_u^l$ represents the interior of Θ_u^l.

Define the parameter space $\Theta^l = \Theta_1^{l*} \times \Theta_2^{l*} \times \cdots \times \Theta_T^{l*} \times [d_1, d_2]$, which encompasses all the required parameters. Given samples of size n, we obtain the estimators $\hat{\theta}_n = (\hat{\phi}_n(1), \hat{\phi}_n(2), \ldots, \hat{\phi}_n(T), \widehat{H}_n)$ and $\hat{\theta}_n(u) = (\hat{\phi}_n(u), \widehat{H}_n)$.

In this paper, we will present both the initial estimator and the one-step estimator for the parameters of the SFAR(1) model. The following sections will delve into the asymptotic properties and characteristics of these estimators in detail.

2.2. The GPH estimator for the hurst index

Due to the nonstationarity of X_n, obtaining an estimator for H using standard semiparametric methods is not feasible. To address this, we can extract stationarity from the data by splitting

the time series $(X_n)_{n \in \mathbb{N}}$ into seasonal components, resulting in T stationary subsequences $\underline{X}(u) = (X_u, X_{T+u}, \dots, X_{nT+u})$ and we construct a stationary additive series defined as $Y_n = \sum_{u=1}^{T} X_{nT+u}$.

In this section, we will estimate H using the log-periodogram method, specifically the GPH estimator, applied to the additive series$(Y_n)_{n \in \mathbb{N}}$. The spectral density and stationarity properties of $(Y_n)_{n \in \mathbb{N}}$ and $(X_{nT+u})_{u \in \mathbb{Z}}$ are outlined in the following three propositions.

Proposition 2.1: *For each $u = 1, 2, \dots, T$ and any $n \in \mathbb{N}$, under conditions A_0 and A_1, the processes*

$$Y_n = \sum_{u=1}^{T} X_{nT+u}, \quad a.s.,$$

$$X_{nT+u} = \sum_{j=0}^{\infty} \phi^j(u) \epsilon_{(n-j)T+u}^{H}, \quad a.s., \tag{10}$$

are stationary.

According to the above formula and Theorem 4.4.1 in Brockwell and Davis (1991), we deduce the spectral density of Y_n from the spectral density of ϵ_t^H. The proof will be presented in detail in Section 6.

Remark 2.1: The stationary process (Y_n) encompasses all the information of ϵ_n^H. Therefore, we will utilize (Y_n) to obtain the estimation of H and the one-step estimator.

Proposition 2.2: *Let $f_{H,\phi(u)}(\lambda)$ be the spectral density of $(X_{nT+u})_{n \in \mathbb{N}}$. Then it can be rewritten as*

$$f_{H,\phi(u)}(\lambda) = \left(1 - 2\phi(u)\cos\lambda T + \phi^2(u)\right)^{-1} f_{\epsilon_n^H}(\lambda).$$

Proposition 2.3: *Let $g_{H,\underline{\phi}}(\lambda)$ be the spectral density of $(Y_n)_{n \in \mathbb{N}}$. Then it can be rewritten as*

$$g_{H,\underline{\phi}}(\lambda) = \left| \sum_{p=0}^{T-1} \Phi_{\phi(T-p)}(\lambda) \right|^2 f_{\epsilon_n^H}(\lambda), \tag{11}$$

where $\Phi_{\phi(T-p)}(\lambda) = \frac{e^{-ip\lambda}}{1 - \phi(T-p)e^{-i\lambda T}}, p = 0, 1, \dots, T-1$.

Because the GPH estimator is a type of semi-parametric estimation as discussed in Geweke and Porter-Hudak (1983), the explicit expression of $|\sum_{p=0}^{T-1} \Phi_{\phi(T-p)}(\lambda)|^2$ does not affect the estimation of H. Thus the equation $\widehat{H}_n = \hat{d}_n + \frac{1}{2}$ remains valid. We can then apply the GPH method directly to the stationary process Y_n.

Let new series $(Y_n)_{n \in \mathbb{N}}$ be an observation sample generated via Equation (10) and choose a suitable integer m which can decrease the mean square error of estimation, where $m < n$.

We get the periodogram of Y_n given by

$$I(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=0}^{n} Y_t \exp(\mathrm{i}t\lambda) \right|^2, \quad \lambda_j = \frac{2\pi j}{n}, \quad j \in \{1, 2, \ldots, m\},$$

$$a_j = \log\left(2\sin\frac{\lambda_j}{2}\right), \quad \overline{a}_m = \frac{1}{m}\sum_{j=1}^{m} a_j, \quad S_m = \sum_{j=1}^{m}(a_j - \overline{a}_m)^2.$$

We estimate d by regressing $\log I(\lambda_j)$ with respect to a_j, such that

$$\hat{d}_n = -\frac{1}{2S_m}\sum_{j=1}^{m}(a_j - \overline{a}_m)\log I(\lambda_j).$$

The estimator $\widehat{H}_n$ is defined by

$$\widehat{H}_n = \hat{d}_n + \frac{1}{2}. \tag{12}$$

Remark 2.2: There are several semi-parametric methods for estimating the long memory parameters d and H, such as whittle estimation and R/S estimation method proposed by Robinson (1995) and Marinucci and Robinson (2001). These models rely on the log-periodogram approach. However, these methods tend to exhibit greater bias compared to the GPH estimator.

2.3. *Generalized least squares estimation of SFAR(1) models*

We now focus on estimating $\phi(u)$ given that the parameter H has been estimated. When the noise in the seasonal autoregressive model is white noise, we can easily obtain the estimator of the parameters of these models using Least Squares Estimation (LSE). However, when the noise is fractional Gaussian noise (fGn), the covariance matrix of fGn is no longer diagonal, making LSE inappropriate. Therefore, we consider using Generalized Least Squares Estimation (GLSE).

To address the effect of seasonal structure on parameter estimation, we apply GLSE to the subsequences $(X_u, X_{T+u}, \ldots, X_{nT+u})$, where $u = 1, 2, \ldots, T$. This allows us to estimate the parameters $\phi(1), \phi(2), \ldots, \phi(T)$ sequentially, assuming the Hurst index is known.

We deduce the time series can be written in the form

$$\Phi_i^j(u) = (X_{u+iT}, X_{u+(i+1)T}, \ldots, X_{u+jT})^*, \quad i \leq j,$$

and the autocovariance matrix is given by

$$\Gamma_{nT,u}(H) = \rho(|i-j|T)_{1 \leq i,j \leq n} = \begin{pmatrix} \gamma_{0,u} & \gamma_{T,u} & \gamma_{2T,u} & \cdots & \gamma_{(n-1)T,u} \\ \gamma_{T,u} & \gamma_{0,u} & \gamma_{3T,u} & \cdots & \gamma_{(n-2)T,u} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_{(n-1)T,u} & \gamma_{(n-2)T,u} & \gamma_{(n-3)T,u} & \cdots & \gamma_{0,u} \end{pmatrix},$$

where $\gamma_{nT,u} = \mathrm{Cov}(X_{nT+u}, X_u), n \in \mathbb{N}$. We can easily show that $\Gamma_{nT,u}(H)$ depends only on n and not on u. Hence, we will denote $\Gamma_{nT,u}$ simply as Γ_{nT} from now on, without distinguishing between them.

The estimators $\{\hat{\phi}_n(u)\}_{u \geq 0}$ are defined by

$$\hat{\phi}_n(u) = \frac{\Phi_2^n(u)^* \Gamma_{(n-1)T}^{-1}(\widehat{H}_n) \Phi_1^{n-1}(u)}{\Phi_1^{n-1}(u)^* \Gamma_{(n-1)T}^{-1}(\widehat{H}_n) \Phi_1^{n-1}(u)}. \tag{13}$$

Now, due to the seasonal structure, we need to examine whether the elements of $\Gamma_{nT}^{-1}(\widehat{H}_n)$ are finite to assess the feasibility of this method.

Thanks to Fox and Taqqu (1986) and Esstafa (2019), we know that the elements of $\Gamma_{nT}^{-1}(\widehat{H}_n)$ can be expressed as a function of the spectral density of fGn. The spectral representation of $(\Gamma_{nT}^{-1})_{j,k}$ implies that

$$(\Gamma_{nT}^{-1})_{j,k} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{1}{f_{\epsilon_n^H}(\lambda)} e^{i(k-j)T\lambda} \, d\lambda.$$

As $\lambda \to 0$, according to the definition of fractional Gaussian noise, we have

$$f_{\epsilon_n^H}(\lambda) \sim \frac{C_H}{2} |\lambda|^{1-2H},$$

where C_H is a constant. We can categorize the elements of the matrix into two types: diagonal elements and off-diagonal elements.

When $j = k$, we have

$$(\Gamma_{nT}^{-1})_{j,j} = \frac{1}{(2\pi)^2} \int_{-\pi}^{\pi} \frac{1}{f_{\epsilon_n^H}(\lambda)} \, d\lambda = \frac{1}{2\pi^2} \int_{0}^{\pi} \frac{1}{f_{\epsilon_n^H}(\lambda)} \, d\lambda. \tag{14}$$

One has when $\lambda \to 0$ that

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

This implies that for $l > 0$ there exists $\delta_l > 0$ such that for any $\lambda \in (-\delta_l, \delta_l)$, we have

$$(1-l)\frac{2}{C_H}|\lambda|^{2H-1} \leq \frac{1}{f_{\epsilon_n^H}(\lambda)} \leq (1+l)\frac{2}{C_H}|\lambda|^{2H-1}. \tag{15}$$

Thus Equation (14) has an upper bound when $\lambda \in (-\delta_l, \delta_l)$:

$$|(\Gamma_{nT}^{-1})_{j,j}| \leq \frac{1+l}{C_H \pi^2} \int_0^{\delta_l} \lambda^{2H-1} \, d\lambda + \frac{1}{2\pi^2} \int_{\delta_l}^{\pi} \frac{1}{f_{\epsilon_n^H}(\lambda)} \, d\lambda$$

$$\leq \frac{\delta_l^{2H}(1+l)}{2HC_H\pi^2} + \frac{\pi - \delta_l}{2\pi^2} \sup_{\lambda \in (\delta_l, \pi]} \frac{1}{f_{\epsilon_n^H}(\lambda)}$$

$$\leq K_1,$$

where K_1 is a constant.

When $j \neq k$, according to Esstafa (2019), there exists a positive constant K_2 such that for any $j, k = 1, 2, \ldots,$

$$|(\Gamma_{nT}^{-1})_{j,k}| \leq K_2 \left| \frac{1}{(k-j)T} \right|^{2H}.$$

Therefore, we have shown that the elements of $\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)$ are finite, which implies that $\hat{\phi}_n(u)$ is bounded. From these points, we use the notations $H, \phi(u)$ and estimator $\widehat{H}_n(u), \hat{\phi}_n(u)$ to present our results concerning the asymptotic properties of the initial estimator.

Theorem 2.4: *Letting $m = [n^\delta]$ for some $0 < \delta < 1$, $(X_{nT+u})_{n \geq 0}$ satisfies Equation (7). Under conditions A_0 and A_1, we have*

$$\begin{pmatrix} \widehat{H}_n \\ \hat{\phi}_n(u) \end{pmatrix} \xrightarrow[n \to \infty]{\mathbb{P}} \begin{pmatrix} H \\ \phi(u) \end{pmatrix},$$

for every $u = 1, 2, \ldots, T$, as $n \to \infty$, where $\widehat{H}_n$ and $\hat{\phi}_n(u)$ are initial estimators defined in Equations (12) and (13). $[\cdot]$ denotes the integer part function.

Remark 2.3: In this proof, we demonstrate that the estimators for each pair of parameters are individually consistent. Consequently, it follows that the estimators for all parameters together are also consistent.

Theorem 2.5: *Let $m = [n^\delta]$ for some $\frac{1}{2} < \delta < \frac{2}{3}$. Under conditions A_0 and A_1, $\hat{\theta}_n$ has a (T+1) dimension limiting normal distribution given by*

$$\sqrt{m} \begin{pmatrix} \widehat{H}_n - H \\ \hat{\phi}_n(1) - \phi(1) \\ \vdots \\ \hat{\phi}_n(T) - \phi(T) \end{pmatrix} \xrightarrow[n \to \infty]{\mathcal{L}} \mathcal{N}(0, \Sigma_\theta),$$

where $\hat{\theta}_n = (\widehat{H}_n, \hat{\phi}_n(1), \hat{\phi}_n(2), \ldots, \hat{\phi}_n(T))$. The covariance matrix Σ_θ is of the form $\Sigma_\theta = V_H \widetilde{\Sigma}_\theta$, V_H is the asymptotic variance of $\sqrt{m}(\widehat{H}_n - H)$, and $\widetilde{\Sigma}_\theta$ is a built-in singular matrix.

Remark 2.4: Hariz et al. (2024) presented $\frac{1}{2} < \delta < \frac{2}{3}$, and Hurvich et al. (1998) stated that if $m = n^\delta$, where $0 < \delta < 1$, it can ensure the asymptotic normality of $\widehat{H}_n$. The condition $\frac{1}{2} < \delta$ is to ensure that $\hat{\phi}_n(1)$ and $\hat{\phi}_n(2)$ are asymptotically normal. But according to Kutoyants and Motrunich (2016), if $\frac{2}{3} < \delta$, a multi-step estimator may be required, which contradicts our consideration of a one-step estimator. Thus we consider restricting δ to the interval $\left(\frac{1}{2}, \frac{2}{3} \right)$.

Remark 2.5: These results can be extended to the SFAR(p), provided that $\underline{X}(u)$ is stationary.

Remark 2.6: Additionally, the estimation of $\underline{\phi}$ can be also approached using methods from Brouste et al. (2014) and Soltane (2024).

3. One-step estimator of SFAR(1) models

In this section, we explore modifications to the initial estimator $\hat{\theta}_n$ to develop a one-step estimator $\tilde{\theta}_n$.

We assume that Y_n is stationary with a spectral density $g_{H,\underline{\phi}}(\lambda)$, as obtained in Equation (11). And $g_{H,\underline{\phi}}(\lambda)$ satisfies the necessary regularity conditions as follows.

Condition 1: For any $\theta = (\theta_{j_1}, \theta_{j_2}, \ldots, \theta_{j_{T+1}}) \in \Theta^l$, where Θ^l is an open subset of R^{T+1}, $g_{H,\underline{\phi}}(\lambda)$ is three times continuously differentiable on Θ^l. In addition, for any $0 \le k \le 3$ and $j_1, \ldots, j_{T+1}$, the partial derivative

$$\frac{\partial^k}{\partial \theta_{j_1} \ldots \partial \theta_{j_k}} g_{H,\underline{\phi}}(\lambda)$$

is a continuous equation on $\Theta^l \times [-\pi, \pi] \backslash \{0\}$, is continuously differentiable with respect to λ and its partial derivative

$$\frac{\partial^{k+1}}{\partial \lambda \partial \theta_{j_1} \ldots \partial \theta_{j_k}} g_{H,\underline{\phi}}(\lambda),$$

and is continuous on $\Theta^l \times [-\pi, \pi] \backslash \{0\}$.

Condition 2: There also exists a continuous function $\alpha \colon \Theta^l \longrightarrow (-1, 1)$, such that for any compact set $\Theta^{l^\star} \subset \Theta^l$ and $\delta > 0$, the following conditions hold for every $(\theta, \lambda) \in \Theta^\star \times [-\pi, \pi] \backslash \{0\}$:

$$c_{1,\delta,\Theta^{l^\star}} |\lambda|^{-\alpha(\theta)+\delta} \le g_{H,\underline{\phi}}(\lambda) \le c_{2,\delta,\Theta^{l^\star}} |\lambda|^{-\alpha(\theta)-\delta}$$

and

$$\left| \frac{\partial}{\partial \lambda} g_{H,\underline{\phi}}(\lambda) \right| \le c_{2,\delta,\Theta^{l^\star}} |\lambda|^{-\alpha(\theta)-1-\delta},$$

for any $k \in \{1, 2, 3\}$ and any $j \in (1, \ldots, T+1)^k$, where

$$\left| \frac{\partial^k}{\partial \theta_{j_1} \ldots \partial \theta_{j_k}} g_{H,\underline{\phi}}(\lambda) \right| \le c_{2,\delta,\Theta^{l^\star}} |\lambda|^{-\alpha(\theta)-\delta}.$$

Here, $c_{1,\delta,\Theta^{l^\star}}$ and $c_{2,\delta,\Theta^{l^\star}}$ are some positive finite constants which only depend upon δ and $\Theta^{l^\star}$. We will prove the spectral density of Y_n satisfies regular condition in auxiliary results.

Proposition 3.1: *Let l_n be the log-likelihood function of a stationary process $(Y_n)_{n \in \mathbb{N}}$. Assume that $g_{H,\underline{\phi}}(\lambda)$ satisfies the regularity conditions and let $B(\theta, R)$ (open ball of centre θ and radius R) for some $R > 0$. Then, for any $t \in B(\theta, R)$, $u \in \mathbb{N}$,*

$$l_n \left(\theta + \frac{t}{\sqrt{n}} \right) - l_n(\theta) = t \frac{\nabla l_n(\theta)}{\sqrt{n}} - \frac{t \mathcal{I}(\theta) t^*}{2} + r_{n,\theta}(t).$$

When $n \to \infty$, the score function $\nabla(\cdot)$ satisfies

$$\frac{\nabla l_n(\theta)}{\sqrt{n}} \xrightarrow[n \to \infty]{\mathbb{P}} \mathcal{N}(0, \mathcal{I}(\theta))$$

and

$$r_{n,\theta}(t) \xrightarrow[n \to \infty]{a.s.} 0,$$

uniformly on each compact set. The Fisher information matrix is given in our case by

$$\mathcal{I}(\theta) = \frac{1}{4\pi} \left(\int_{-\pi}^{\pi} \frac{\partial \log g_{H,\underline{\phi}}(\lambda)}{\partial \theta_k} \frac{\partial \log g_{H,\underline{\phi}}(\lambda)}{\partial \theta_j} \, d\lambda \right)_{1 \le k,j \le T+1}. \tag{16}$$

This result is a direct consequence of Theorem from Cohen et al. (2013).

Since $g_{H,\underline{\phi}}(\lambda)$ satisfies the regularity conditions, the elements of the Fisher information matrix $\mathcal{I}(\theta)$ are finite. After obtaining the Fisher information matrix $\mathcal{I}(\theta)$ and the log-likelihood function of $(Y_n)_{n \in \mathbb{N}}$, we can compute the one-step estimator as follows:

$$\tilde{\theta}_n = \hat{\theta}_n + \mathcal{I}(\hat{\theta}_n)^{-1} \frac{1}{n} \nabla l_n(\hat{\theta}_n). \tag{17}$$

We can now analyse the asymptotic properties of the one-step estimator.

Theorem 3.2: *Let $\hat{\theta}_n$ be the initial estimator of θ and $\tilde{\theta}_n$ be the one-step estimator of θ. When $g_{H,\underline{\phi}}(\lambda)$ satisfies the regular conditions, we have a asymptotic normal distribution of $\tilde{\theta}_n$ that*

$$\sqrt{n}(\tilde{\theta}_n - \theta) \xrightarrow[n \to \infty]{\mathbb{P}} \mathcal{N}(0, \mathcal{I}(\theta)^{-1}).$$

Remark 3.1: The parameter θ should not lie on the boundary of the parameter space Θ^l.

Remark 3.2: The one-step estimator can be applied more generally even if the initial estimator $\hat{\theta}_n$ does not satisfy asymptotic normality. According to Proposition 2.3 in Hariz (2025), if the initial estimator with convergence speed is lower than $\sqrt{n}$ and the spectral density of time series meets the regular condition, then the one-step estimator $\tilde{\theta}_n$ can still achieve asymptotic normality.

Remark 3.3: One-step estimator can achieve Hájek's lower bound, and thus it is asymptotically efficient in the local minimax sense. We can find related conclusions in Brouste et al. (2020) and Cohen et al. (2013).
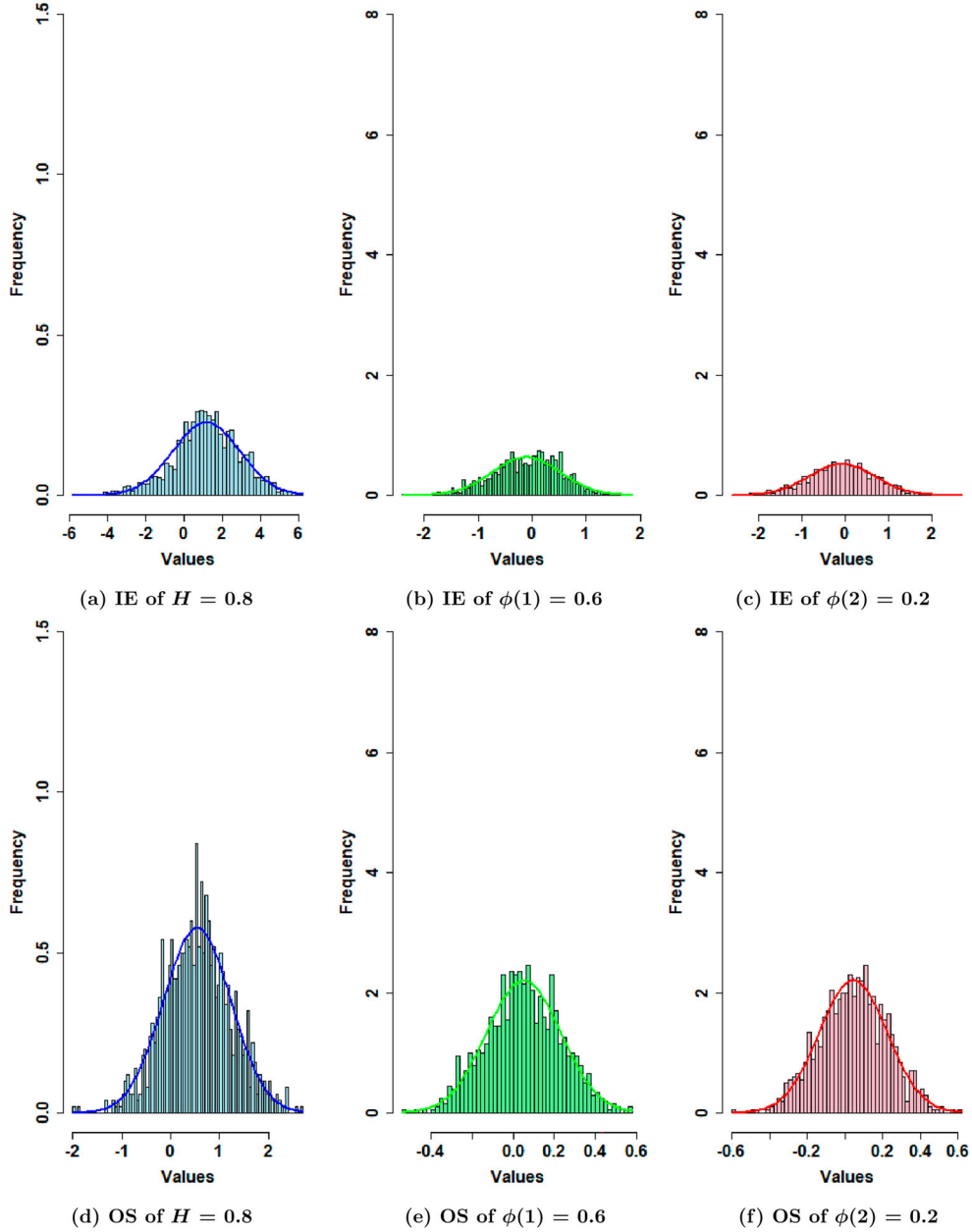
4. Simulation study

According to Equation (10), the likelihood function based on the sample $\underline{Y}^{(n)} = (Y_0, Y_1, \ldots, Y_{n-1})$ is given by

$$l_n(\theta) = -\frac{1}{2} \log \det \left(\Gamma_n^Y(\theta) \right) - \frac{1}{2} \underline{Y}^{(n)*} \Gamma_n^Y(\theta) \underline{Y}^{(n)},$$

where $\Gamma_n^Y(\theta)$ is the covariance matrix of $\underline{Y}^{(n)}$. For any $K \in \mathbb{N}$,

$$\text{Cov}(Y_0, Y_k) = \int_{-\pi}^{\pi} \exp(ik\lambda) \, g_{H,\underline{\phi}}(\lambda) \, d\lambda,$$
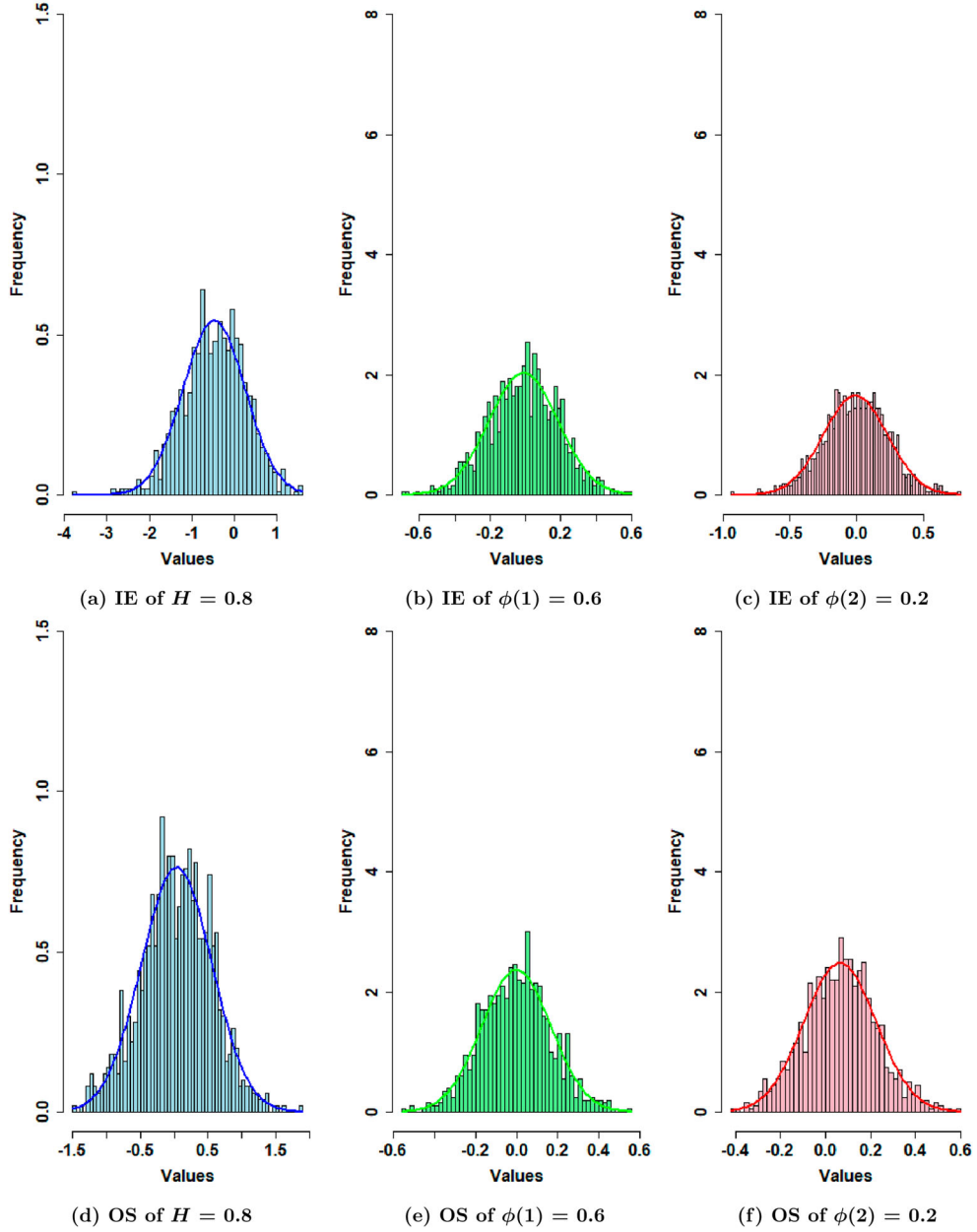
where $\text{Cov}(\cdot, \cdot)$ denotes the covariance. The score function with respect to θ is given by

Figure 1. The simulation of initial estimator and one-step estimator where $\theta = (0.6, 0.2, 0.8)$ for $m = [n^{3/5}], n = 100$.

$$\frac{\partial l_n(\theta)}{\partial \theta_i} = -\frac{1}{2}\mathrm{tr}\left(\left(\Gamma_n^Y(\theta)\right)^{-1}\frac{\partial \Gamma_n^Y(\theta)}{\partial \theta_i}\right) + \frac{1}{2}\underline{Y}^{(n)^*}\left(\Gamma_n^Y(\theta)\right)^{-1}\frac{\partial \Gamma_n^Y(\theta)}{\partial \theta_i}\left(\Gamma_n^Y(\theta)\right)^{-1}\underline{Y}^{(n)},$$
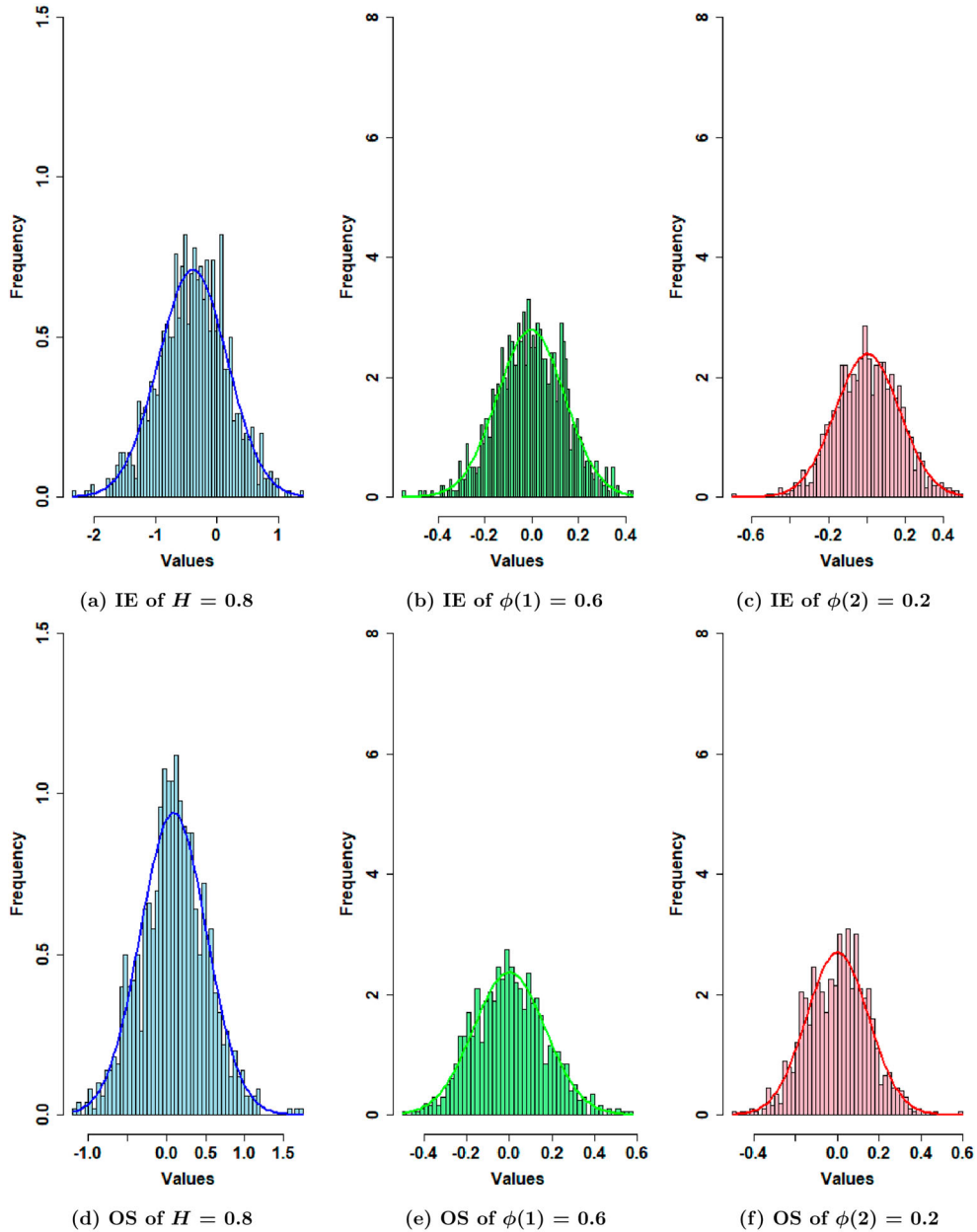
where $\mathrm{tr}(\cdot)$ denotes the trace of a matrix. The Fisher information matrix (FIM) can be deduced from Equation (16). We simulate the spectral density and its derivatives using the method described in Hariz et al. (2024), and then plug the FIM and score functions into Equation (17) to compute the one-step estimator numerically.

(a) IE of $H = 0.8$

(b) IE of $\phi(1) = 0.6$

(c) IE of $\phi(2) = 0.2$

(d) OS of $H = 0.8$

(e) OS of $\phi(1) = 0.6$

(f) OS of $\phi(2) = 0.2$

Figure 2. The simulation of initial estimator and one-step estimator where $\theta = (0.6, 0.2, 0.8)$ for $m = [n^{3/5}], n = 1000$.

For each set of parameters, specifically $(\phi(1), \phi(2), H) = (0.6, 0.2, 0.8)$ and $(\phi(1), \phi(2), H) = (0.2, 0.8, 0.6)$, we conduct $M = 1000$ Monte Carlo simulations. The sample sizes considered are $n = 100$, $n = 1000$ and $n = 2000$. The number of Fourier frequencies for the initial estimations is set as $m = [n^{0.6}]$ and remains fixed throughout the simulations. Without loss of generality, we assume $T = 2$, and the spectral density of Y_n in this case is given by

$$g_{H,\phi(1),\phi(2)}(\lambda) = \frac{2 + \phi^2(1) + \phi^2(2) + A\cos\lambda - B\cos 2\lambda - C\cos 3\lambda}{\left(1 - 2\phi(2)\cos 2\lambda + \phi^2(2)\right)\left(1 - 2\phi(1)\cos 2\lambda + \phi^2(1)\right)} f_{\epsilon_n^H}(\lambda), \quad (18)$$
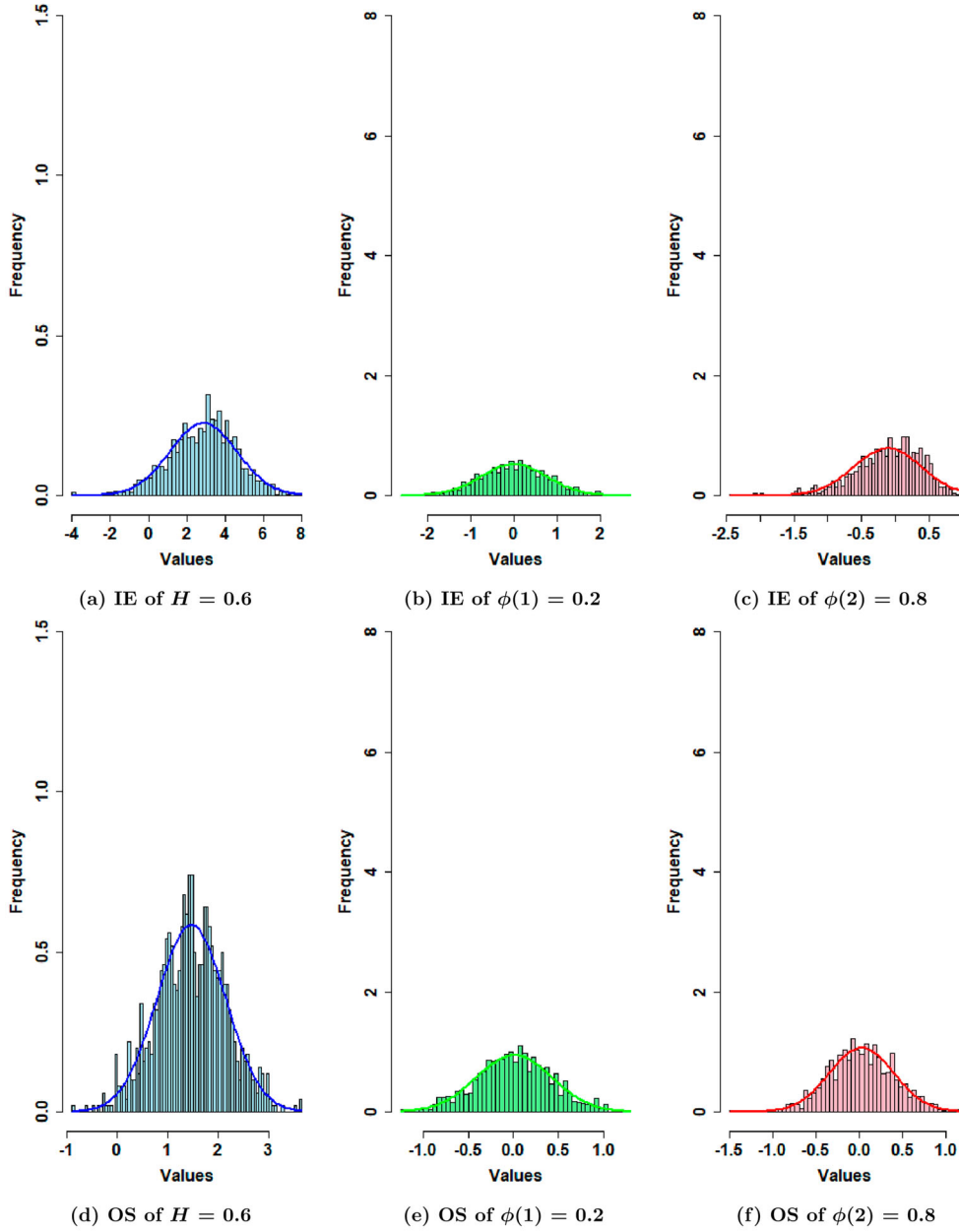
Figure 3. The simulation of initial estimator and one-step estimator where $\theta = (0.6, 0.2, 0.8)$ for $m = [n^{3/5}], n = 2000$.

where $A = 2 + 2\phi(1)\phi(2) - 2\phi(1)$, $B = 2(\phi(1) + \phi(2))$ and $C = 2\phi(2)$.

Figures 1, 2 and 3 depict the frequency distribution of statistical errors for the initial estimator and one-step estimator of the SFAR(1) model with parameters $\phi(1) = 0.6$, $\phi(2) = 0.2$ and $H = 0.8$. Figures 4, 5 and 6 depict the frequency distribution of statistical errors for the initial estimator and one-step estimator of the SFAR(1) model with parameters $\phi(1) = 0.2$, $\phi(2) = 0.8$ and $H = 0.6$.

In all the tables (Tables 1–6), B stands for Bias, IE represents initial estimator and OS denotes one-step estimator. From the above tables, it can be seen that the OS estimator shows

Figure 4. The simulation of initial estimator and one-step estimator where $\theta = (0.2, 0.8, 0.6)$ for $m = [n^{\frac{3}{5}}], n = 100$.

a significant improvement in the estimation of H. From these figures and the accompanying tables, it is evident that the one-step estimator outperforms the initial estimation, with a particularly notable improvement in estimating the parameter H. At the same time, we found that as the sample size increases, the estimation becomes more efficient. According to Hariz et al. (2024) and our simulations, the one-step estimation also has a faster running speed.
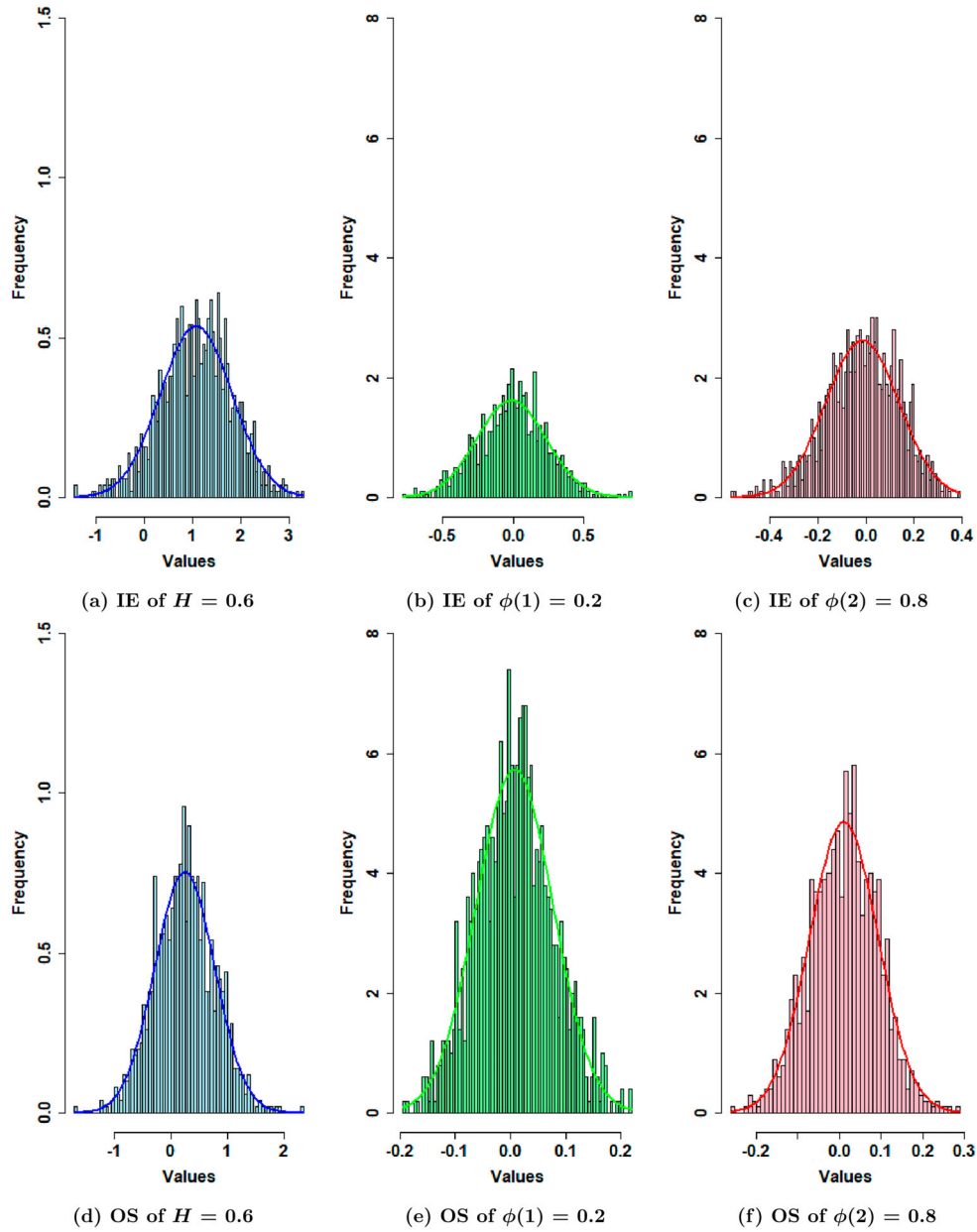
Figure 5. The simulation of initial estimator and one-step estimator where $\theta = (0.2, 0.8, 0.6)$ for $m = [n^{\frac{3}{5}}], n = 1000$.

5. Conclusion and perspectives

In this paper, we propose a simple and effective method for estimating the parameters of the SFAR model individually, and we derive the asymptotic properties of this method. We address the difficulty of parameter estimation caused by the non-stationarity of the model by creating new subseries and obtaining an explicit form for the spectral density of the additive series.

(a) IE of $H = 0.6$ (b) IE of $\phi(1) = 0.2$ (c) IE of $\phi(2) = 0.8$

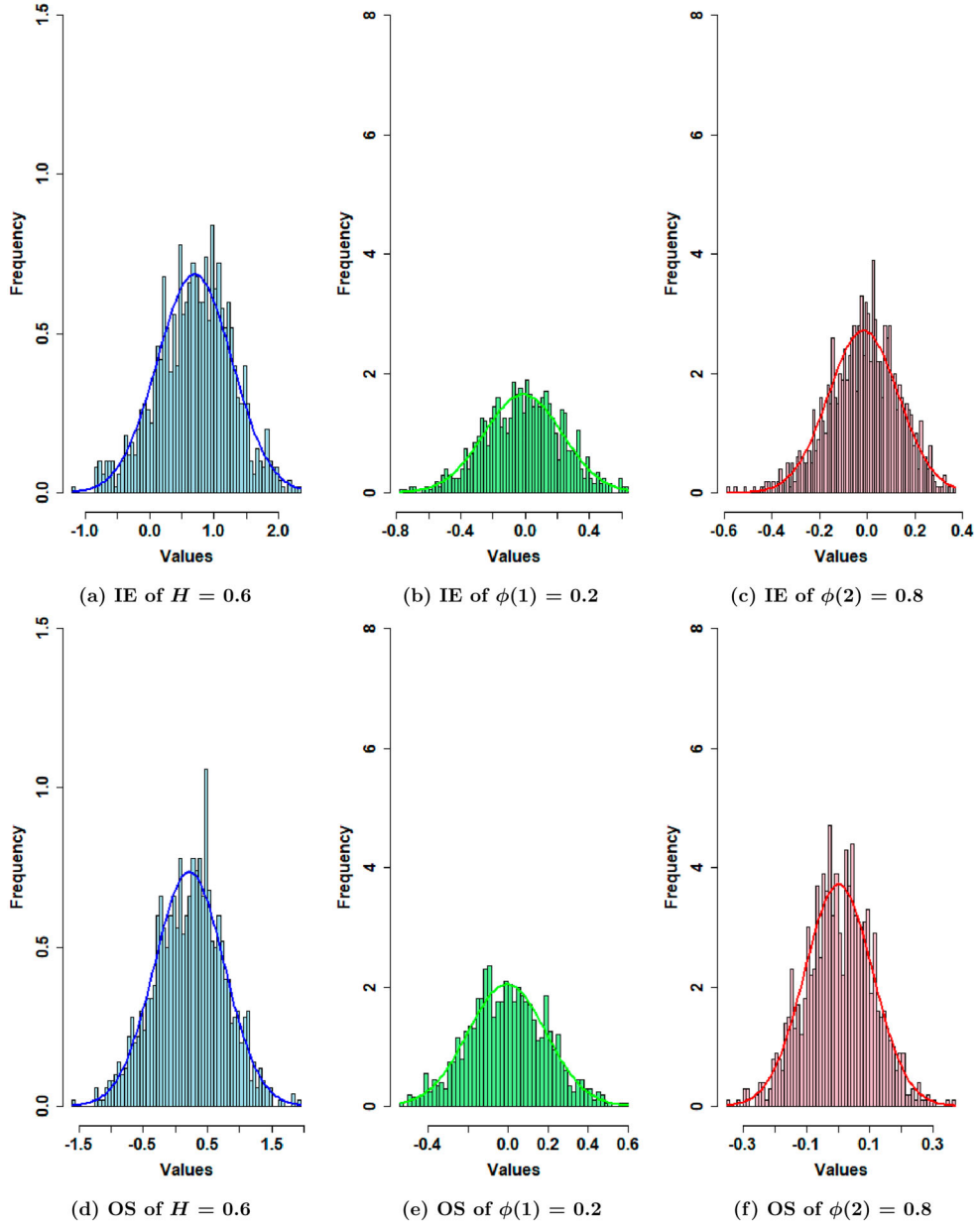(d) OS of $H = 0.6$ (e) OS of $\phi(1) = 0.2$ (f) OS of $\phi(2) = 0.8$

Figure 6. The simulation of initial estimator and one-step estimator where $\theta = (0.2, 0.8, 0.6)$ for $m = [n^{\frac{3}{5}}], n = 2000$.

The one-step procedure is essentially a gradient descent approach, achieving the $\sqrt{n}$ rate with optimal variance.

Our results can be extended to SARIMA models by adjusting the calculation of the covariance matrix of the noise and the spectral density of X_n. Additionally, more effective initial estimators can be utilized for the one-step procedure, similar to the approach taken by Hariz et al. (2024) in the estimation of FARIMA models.

Table 1. The bias and RMSE of initial estimator and one-step estimator for $\theta = (0.6, 0.2, 0.8)$ when $n = 100$.

$n = 100$	B IE	B OS	RMSE IE	RMSE OS
H	0.1462	0.0683	0.2656	0.1106
$\phi(1)$	−0.0143	0.0064	0.0807	0.0237
$\phi(2)$	−0.0082	−0.0055	0.0972	0.0234

Table 2. The bias and RMSE of initial estimator and one-step estimator for $\theta = (0.6, 0.2, 0.8)$ when $n = 1000$.

$n = 1000$	B IE	B OS	RMSE IE	RMSE OS
H	−0.0599	0.0471	0.1101	0.0660
$\phi(1)$	−0.0015	−0.0002	0.0248	0.0212
$\phi(2)$	−0.0010	−0.0081	0.0304	0.0218

Table 3. The bias and RMSE of initial estimator and one-step estimator for $\theta = (0.6, 0.2, 0.8)$ when $n = 2000$.

$n = 2000$	B IE	B OS	RMSE IE	RMSE OS
H	−0.0497	0.0112	0.0864	0.0545
$\phi(1)$	−0.0004	0.0001	0.0180	0.0212
$\phi(2)$	0.0005	0.0003	0.0210	0.0186

Table 4. The bias and RMSE of initial estimator and one-step estimator for $\theta = (0.2, 0.8, 0.6)$ when $n = 100$.

$n = 100$	B IE	B OS	RMSE IE	RMSE OS
H	0.3597	0.1846	0.4227	0.2036
$\phi(1)$	−0.0009	0.0010	0.0962	0.0530
$\phi(2)$	−0.0140	0.0028	0.0651	0.0471

Table 5. The bias and RMSE of initial estimator and one-step estimator for $\theta = (0.2, 0.8, 0.6)$ when $n = 1000$.

$n = 1000$	B IE	B OS	RMSE IE	RMSE OS
H	0.1351	0.0312	0.1644	0.0736
$\phi(1)$	−0.0014	−0.0009	0.0309	0.0088
$\phi(2)$	−0.0019	0.0014	0.0193	0.0104

An interesting aspect to consider is that when T is sufficiently large, even larger than n, but still finite, the effectiveness of this gradient descent approach may diminish. In such cases, alternative methods for optimizing the initial estimator should be explored.
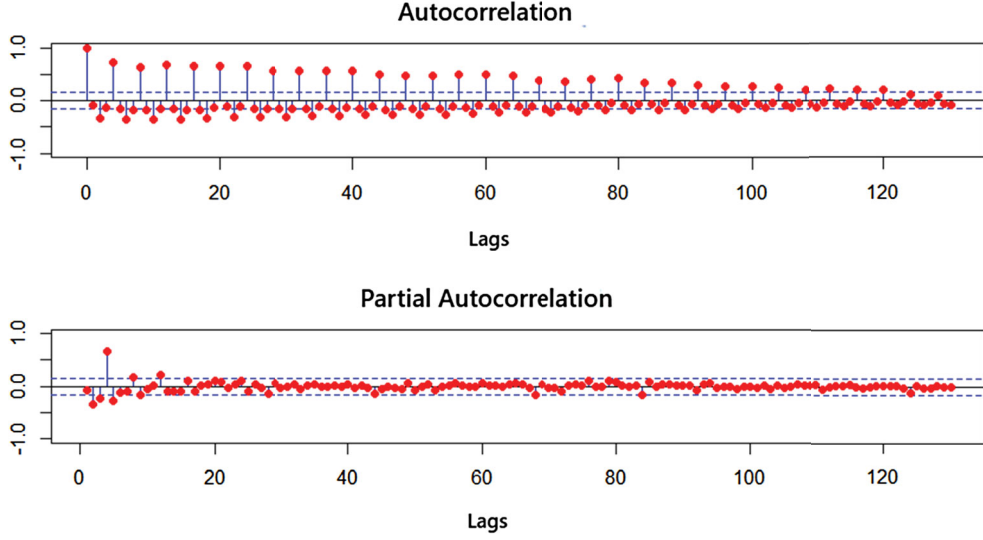
Table 6. The bias and RMSE of initial estimator and one-step estimator for $\theta = (0.2, 0.8, 0.6)$ when $n = 2000$.

$n = 2000$	B IE	B OS	RMSE IE	RMSE OS
H	0.0883	0.0274	0.1146	0.0735
$\phi(1)$	0.0011	0.0006	0.0305	0.0245
$\phi(2)$	−0.0032	0.0003	0.0186	0.0135

Figure 7. Autocorrelation and partial autocorrelation coefficient plots of the Colorado River runoff in Arizona.

6. Application on real data

In this section, we will conduct practical modelling and analysis to examine the application effectiveness of the seasonal autoregressive model driven by fractional Gaussian noise in real data.

The data on the Colorado River runoff in Arizona selected in this paper are from the public data of the United States Geological Survey. This dataset records the monthly river runoff of the Colorado River from 1922 to 2022, with the unit of cubic feet per second. To facilitate modelling, we average the data of each of the 12 months on a quarterly basis, obtaining the quarterly runoff data for the first, second, third and fourth quarters respectively. First, we calculate and obtain the autocorrelation function (ACF) plot and partial autocorrelation function (PACF) plot of the quarterly runoff data as follows.

From Figure 7, we can observe that the autocorrelation function shows a trailing pattern with a slow decay rate, while the partial autocorrelation function cuts off. Therefore, it is appropriate to consider using a fractional AR model with long memory properties.

To avoid data over crowding and considering that the river runoff around 1963 changed significantly for unknown reasons, we extract the runoff data from 1922 to 1962 and draw the following sample path plot.

As can be seen from Figure 8, the runoff of this river exhibits obvious seasonality. Considering the above two points, in this empirical analysis, we consider using the SFAR(1) model to simulate the above observations and compare it with the simulation of the seasonal autoregressive model driven by white noise.
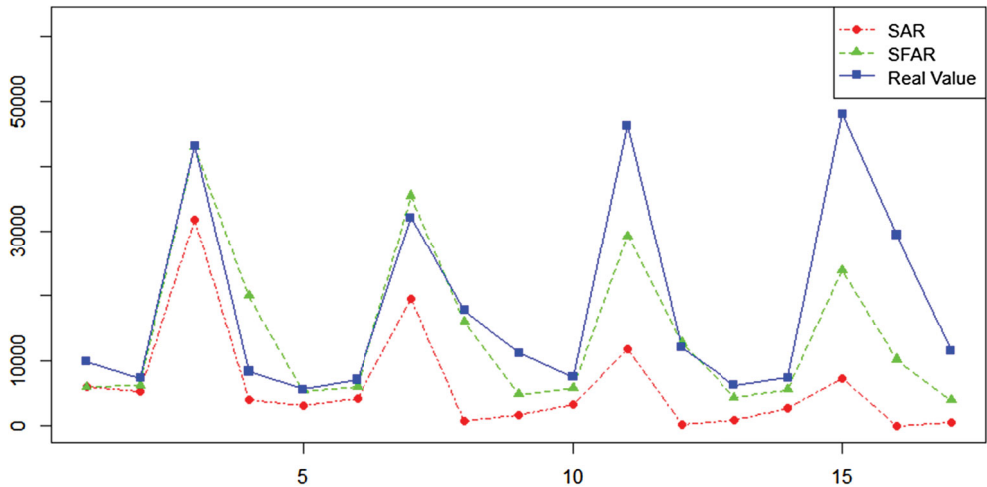
The seasonal autoregressive model driven by white noise is as follows:

$$X_{4n+u} = \alpha(4n + u)X_{4(n-1)+u} + \epsilon_{4n+u}, \quad u = 1, 2, 3, 4, \ n \in \mathbb{Z}^+, \tag{19}$$

where ϵ_{4n+u} is white noise, and $\alpha(1), \alpha(2), \alpha(3), \alpha(4)$ are the model coefficients, satisfying $\alpha(4n + u) = \alpha(u)$.
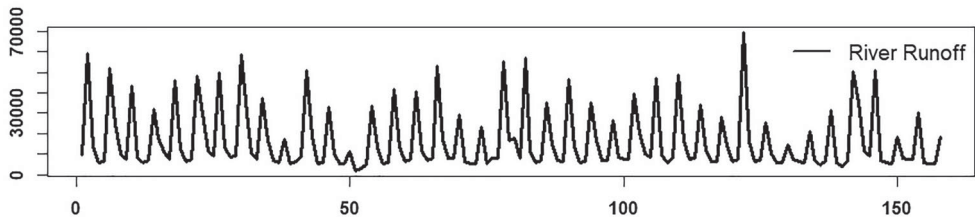
Figure 8. Sample path plot of the Colorado River runoff in Arizona from 1922 to 1962.

Figure 9. Model fitting plot.

Table 7. Fitting results of different models.

Model	SFAR	SAR
Parameters	$(\phi(1), \phi(2), \phi(3), \phi(4), H)$	$(\alpha(1), \alpha(2), \alpha(3), \alpha(4))$
Values	(0.96, 0.82, 0.80, 0.90, 0.60)	(0.80, 0.61, 0.16, 0.53)
RMSE	9439.37	16773.58
MAE	6107.16	12264.90

We utilized the data from 1922 to 1962 to derive the parameter estimations of the two models. Subsequently, we computed their RMSEb and MAE against the real data. Finally, we randomly simulated 20 data points within these 40 years using these two models. The results are shown in the following table and figure.

In the following figure, the values represent the parameters fitted by the SFAR and SAR models. From Figure 9 and Table 7, we can see that the seasonal autoregressive model driven by fractional noise has smaller RMSE and MAE values and better fitting performance. Therefore, the seasonal autoregressive model with long memory properties is more suitable for the study of the Colorado River runoff.

Disclosure statement

No potential conflict of interest was reported by the author(s).

References

Beran, J., Feng, Y., Ghosh, S., & Kulik, R. (2013). *Long-Memory Processes: Probabilistic Properties and Statistical Methods*. Springer.

Bisognin, C., & Lopes, S. R. C. (2009). Properties of seasonal long memory processes. *Mathematical and Computer Modelling*, 49(9–10), 1837–1851. https://doi.org/10.1016/j.mcm.2008.12.003

Brockwell, P. J., & Davis, R. A. (1991). *Time Series: Theory and Methods*. Springer Science & Business Media.

Brouste, A., Cai, C., & 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. https://doi.org/10.3103/S1066530714020021

Brouste, A., Soltane, M., & Votsi, I. (2020). One-step estimation for the fractional Gaussian noise at high-frequency. *ESAIM: Probability and Statistics*, 24, 827–841. https://doi.org/10.1051/ps/2020022

Carlin, J., & Dempster, A. (1989). Sensitivity analysis of seasonal adjustments: Empirical case studies. *Journal of the American Statistical Association*, 84(405), 6–20. https://doi.org/10.1080/01621459.1989.10478729

Chan, N. H., & Terrin, N. (1995). Inference for unstable long-memory processes with applications to fractional unit root autoregressions. *The Annals of Statistics*, 23(5), 1662–1683.

Chatfield, C., & Prothero, D. (1973). Box–Jenkins seasonal forecasting: Problems in a case study. *Journal of the Royal Statistical Society: Series A (General)*, 136(3), 295–315. https://doi.org/10.2307/2344994

Cohen, S., Gamboa, F., Lacaux, C., & Loubes, J. M. (2013). LAN property for some fractional type Brownian motion. *ALEA: Latin American Journal of Probability and Mathematical Statistics*, 10(1), 91–106. https://hal.science/hal-00638121.

Esstafa, Y. (2019). *Long-memory Time Series Models with Dependent Innovations* [Unpublished doctoral dissertation]. Université Bourgogne Franche-Comté.

Fox, R., & Taqqu, M. S. (1986). Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series. *The Annals of Statistics*, 14(2), 517–532. https://doi.org/10.1214/aos/1176349936

Franco, G. C., & Reisen, V. A. (2007). Bootstrap approaches and confidence intervals for stationary and non-stationary long-range dependence processes. *Physica A: Statistical Mechanics and its Applications*, 375(2), 546–562. https://doi.org/10.1016/j.physa.2006.08.027

Geweke, J., & Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, 4(4), 221–238. https://doi.org/10.1111/jtsa.1983.4.issue-4

Gloter, A., & Yoshida, N. (2021). Adaptive estimation for degenerate diffusion processes. *Electronic Journal of Statistics*, 15(1): 1424-1472.

Hariz, S. B., Brouste, A., Cai, C., & Soltane, M. (2024). Fast and asymptotically-efficient estimation in an autoregressive process with fractional type noise. *Journal of Statistical Planning and Inference*, 232, 106148. https://doi.org/10.1016/j.jspi.2024.106148

Harrison, P. J. (1965). Short-term sales forecasting. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 14(2-3), 102–139.

Hosking, J. R. (1984). Modeling persistence in hydrological time series using fractional differencing. *Water Resources Research*, 20(12), 1898–1908. https://doi.org/10.1029/WR020i012p01898

Hurst, H. E. (1951). Long-term storage capacity of reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1), 770–799. https://doi.org/10.1061/TACEAT.0006518

Hurvich, C. M., Deo, R., & Brodsky, J. (1998). The mean squared error of Geweke and Porter-Hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, 19(1), 19–46. https://doi.org/10.1111/jtsa.1998.19.issue-1

Kong, J., & Lund, R. (2023). Seasonal count time series. *Journal of Time Series Analysis*, 44(1), 93–124. https://doi.org/10.1111/jtsa.v44.1

Kutoyants, Y. A., & Motrunich, A. (2016). On multi-step MLE-process for Markov sequences. *Metrika*, 79(6), 705–724. https://doi.org/10.1007/s00184-015-0574-4

Le Cam, L. (1956). On the asymptotic theory of estimation and testing hypotheses. In *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, volume 1: Contributions to the Theory of Statistics* (Vol. 3, pp. 129–157).

Lieberman, O., Rosemarin, R., & Rousseau, J. (2012). Asymptotic theory for maximum likelihood estimation of the memory parameter in stationary Gaussian processes. *Econometric Theory*, *28*(2), 457–470. https://doi.org/10.1017/S0266466611000399

Marinucci, D., & Robinson, P. M. (2001). Semiparametric frequency domain analysis of fractional cointegration. *Journal of Econometrics*, *105*(1), 225–247. https://doi.org/10.1016/S0304-4076(01)00076-8

Robinson, P. M. (1995). Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, *23*(3), 1048–1072. https://doi.org/10.1214/aos/1176324636

Robinson, P. M. (2010). Long memory models. In *Macroeconometrics and Time Series Analysis* (pp. 163–168). Springer.

Soltane, M. (2024). Asymptotic efficiency in autoregressive processes driven by stationary Gaussian noise. *Stochastic Models*, *40*(1), 70–96. https://doi.org/10.1080/15326349.2023.2202227

Tsay, R. S. (2013). *Multivariate Time Series Analysis: With R and Financial Applications*. John Wiley & Sons.

Appendices

Appendix 1. Proofs of the main results

For clarity, we divide the technical results into two parts. The first part addresses the stationarity and spectral density of the SFAR(1) model, as well as the asymptotic properties of the initial estimator. The second part focuses on the asymptotic properties related to the one-step estimator.

A.1. Proof of Proposition 2.1

We will utilize the following lemma to demonstrate the stationarity of Y_n.

Lemma A.1: *For any $u \in \mathbb{N}$, the SFAR(1) model is defined by the recursive scheme*

$$X_{nT+u} = \phi(u)X_{(n-1)T+u} + \epsilon_{nT+u}^{H}, \tag{A1}$$

where ϵ_{nT+u}^{H} is a fractional Gaussian noise, $\phi(u) = \phi(u + nT) \leq 1$, $n \in \mathbb{N}$, and then $(X_{nT+u})_{n\in\mathbb{N}}$ is a stationary process.

Proof: We verify that the process satisfies the three conditions for weak stationarity individually.
(1) For any $u \in \mathbb{N}$, $E(X_{nT+u}) = \mu$ is a finite constant.

Because the equation $(1 - \phi(u))z = 0$ has a root outside the unit circle, the process $(X_{nT+u})_{n\in\mathbb{N}}$ is said to be an SFAR(1) process if it can be represented as follows:

$$X_{nT+u} = \sum_{j=0}^{\infty} \phi^{j}(u)\epsilon_{(n-j)T+u}^{H}. \tag{A2}$$

Without loss of generality, we assume $E(\epsilon_{n}^{H}) = 0$. For any time series $(X_{nT+u})_{n\in\mathbb{N}}$ under the monotone convergence theorem and the Cauchy-Schwarz inequality, we obtain

$$E|X_{nT+u}| \leq E \sum_{j=0}^{\infty} \left| \phi^{j}(u)\epsilon_{(n-j)T+u}^{H} \right| = \sum_{j=0}^{\infty} E|\phi^{j}(u)\epsilon_{(n-j)T+u}^{H}|$$

$$\leq \sum_{j=0}^{\infty} \left| \phi^{j}(u)|E|\epsilon_{(n-j)T+u}^{H} \right|$$

$$\leq C \sum_{j=0}^{\infty} |\phi^{j}(u)|. \tag{A3}$$

We know that $\phi^{j}(u) = o(\rho^{-j})$ as $j \to \infty$ for $1 < \rho < \frac{1}{\phi(u)}$. Consequently, $\phi^{j}(u)$ is absolutely summable, i.e., $\sum_{j=0}^{\infty} |\phi^{j}(u)| < \infty$. Thus $E|X_{nT+u}| < \infty$ as shown in Equation (A3). By the monotone convergence theorem, $\sum_{j=0}^{\infty} \phi^{j}(u)$ is absolutely convergent almost surely.

Considering that

$$\left| \sum_{j=0}^{k} \phi^j(u)\epsilon_{(n-j)T+u}^{H} \right| \leq \sum_{j=0}^{k} \left| \phi^j(u)\epsilon_{(n-j)T+u}^{H} \right|$$

with the dominated convergence theorem,

$$E(X_{nT+u}) = \lim_{k \to \infty} E\left(\sum_{j=0}^{k} \phi^j(u)\epsilon_{(n-j)T+u}^{H} \right) = 0.$$

(2) For any $u \in \mathbb{N}$, $E(X_{nT+u})^2 \leq \infty$.

From Equation (A2), we derive

$$E(X_{nT+u})^2 = E\left(\sum_{j=0}^{\infty} \phi^j(u)\epsilon_{(n-j)T+u}^{H} \right)^2$$

$$= E\left(\sum_{s=0}^{\infty} \sum_{k=0}^{\infty} \phi^s(u)\phi^k(u)\epsilon_{(n-s)T+u}^{H}\epsilon_{(n-k)T+u}^{H} \right). \tag{A4}$$

By applying the conclusion above, we obtain

$$E|X_{nT+u}|^2 = \sum_{s=0}^{\infty} \sum_{k=0}^{\infty} |\phi^s(u)\phi^k(u)|E\left| \epsilon_{(n-s)T+u}^{H}\epsilon_{(n-k)T+u}^{H} \right|,$$

and the covariance of $\epsilon_{(n-s)T+u}^{H}$ and $\epsilon_{(n-k)T+u}^{H}$ is

$$\rho\left(\epsilon_{(n-s)T+u}^{H}, \epsilon_{(n-k)T+u}^{H} \right) = \frac{1}{2}\left(|(s-k)+1|^{2H} - 2|(s-k)T|^{2H} + |(s-k)T-1|^{2H} \right).$$

Since $\phi^j(u)$ is absolutely summable, it is also square summable. Additionally, as $s - k \to \infty$, $\rho(\epsilon_{(n-s)T+u}^{H}, \epsilon_{(n-k)T+u}^{H}) \to 0$, implying that there exists a constant M such that $E|\epsilon_{(n-s)T+u}^{H}\epsilon_{(n-k)T+u}^{H}| \leq M$. Based on the above discussion and Equation (A4), we have established that $E(X_{nT+u})^2 \leq \infty$.

(3) For any $k, s \in \mathbb{N}$, $E(X_{kT+u} - \mu)E(X_{sT+u} - \mu) = \gamma_{(k-s)T}$, which means that the autocovariance of X_{kT+u} and X_{sT+u} depends only on the time interval $(k - s)T$.

Without loss of generality, we assume $\mu = 0$ and the covariance of $(X_{nT+u})_{n\in\mathbb{N}}$ is rewritten as

$$E(X_{kT+u}X_{sT+u}) = E\left(\sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \phi^i(u)\phi^j(u)\epsilon_{(s-i)T+u}^{H}\epsilon_{(k-j)T+u}^{H} \right)$$

$$= \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \phi^i(u)\phi^j(u)\frac{1}{2}(|(s-k+j-i)T+1|^{2H}$$

$$- 2|(s-k+j-i)T|^{2H} + |(s-k+j-i)T-1|^{2H}).$$

Then for any $q \in \mathbb{N}$, it follows directly from the above equation that

$$E(X_{(k+q)T+u}X_{(s+q)T+u}) = E(X_{kT+u}X_{sT+u}).$$

Thus we have shown that $(X_{nT+u})_{n\in\mathbb{N}}$ is stationary. Since Y_n is a combination of X_{nT+u} in a cyclic manner, its stationarity naturally follows. ∎

A.2. Proof of Proposition 2.2

Since the stationarity of $(X_{nT+u})_{n\in\mathbb{N}}$ has been established, we can determine its spectral density using Theorem 4.4 (Brockwell & Davis, 1991). $(X_{nT+u})_{n\in\mathbb{N}}$ satisfies the recursion

$$X_{nT+u} = X_{(n-1)T+u} + \epsilon^H_{nT+u}$$

$$= \sum_{j=0}^{\infty} \phi^j(u)\epsilon^H_{(n-j)T+u}$$

$$= \sum_{j=0,T,2T,\ldots} \phi^{\frac{j}{T}}(u)\epsilon^H_{nT+u-j}.$$

From the above expression, we obtain the transfer function $H(\mathrm{e}^{-\mathrm{i}\lambda})$ as follows:

$$H(\mathrm{e}^{-\mathrm{i}\lambda}) = \sum_{k=0}^{\infty} \phi^k(u)\mathrm{e}^{-\mathrm{i}\lambda kT}.$$

Thus the spectral density function $f_{H,\phi(u)}(\lambda)$ is given by

$$f_{H,\phi(u)}(\lambda) = |H(\mathrm{e}^{-\mathrm{i}\lambda})|^2 f_{\epsilon^H_n}(\lambda) = \frac{f_{\epsilon^H_n}(\lambda)}{1 - 2\phi(u)\cos(\lambda T) + \phi^2(u)}.$$

A.3. Proof of Proposition 2.3

Because the stationarity of $(Y_n)_{n\in\mathbb{N}}$ has been proved above, Y_n has the following expression:

$$Y_n = \sum_{u=1}^{T} X_{nT+u}$$

$$= \sum_{j=0}^{\infty} \phi^j(1)\epsilon^H_{(n-j)T+1} + \sum_{j=0}^{\infty} \phi^j(2)\epsilon^H_{(n-j)T+2} + \cdots + \sum_{j=0}^{\infty} \phi^j(T)\epsilon^H_{(n-j)T+T}$$

$$= \sum_{k=0}^{\infty} \tilde{h}_k \epsilon^H_{(n+1)T-k},$$

and the coefficient of transfer function has the form of

$$\tilde{h}_k = \begin{cases} \phi^{\frac{k}{T}}(T), & \text{if } k = 0, T, 2T, \ldots, \\ \phi^{\frac{k-1}{T}}(T-1), & \text{if } k = 1, 1+T, 1+2T, \ldots, \\ \quad\vdots \\ \phi^{\frac{k-(T-1)}{T}}(1), & \text{if } k = T-1, 2T-1, 3T-1, \ldots. \end{cases}$$

Then the transfer function is given by

$$H(\mathrm{e}^{-\mathrm{i}\lambda}) = \sum_{j=kT}^{\infty} \phi^{\frac{j}{T}}(T)\mathrm{e}^{-\mathrm{i}j\lambda} + \sum_{j=1+kT}^{\infty} \phi^{\frac{j-1}{T}}(T-1)\mathrm{e}^{-\mathrm{i}j\lambda} + \cdots + \sum_{j=(k+1)T-1}^{\infty} \phi^{\frac{j-T+1}{T}}(1)\mathrm{e}^{-\mathrm{i}j\lambda}.$$

To simplify the notation, we denote

$$\Phi_{\phi(T-p)}(\lambda) = \sum_{j=kT+p}^{\infty} \phi^{\frac{j-p}{T}}(T-p)\mathrm{e}^{-\mathrm{i}j\lambda} = \frac{\mathrm{e}^{-\mathrm{i}p\lambda}}{1 - \phi(T-p)\mathrm{e}^{-\mathrm{i}\lambda T}}, \quad p = 0, 1, \ldots, T-1.$$

Then we have

$$H(e^{-i\lambda}) = \left(\sum_{P=0}^{T-1} \Phi_{\phi(T-p)}(\lambda) \right),$$

and the spectral density of $(Y_n)_{n \in \mathbb{N}}$ has the representation

$$g_{H,\phi} = |H(e^{-i\lambda})|^2 f_{\epsilon_n^H}^H(\lambda) = \left| \sum_{P=0}^{T-1} \Phi_{\phi(T-p)}(\lambda) \right|^2 f_{\epsilon_n^H}^H(\lambda).$$

Remark A.1: $|\cdot|$ denotes the modulus of $H(e^{-i\lambda})$.

A.4. Proof of Theorem 2.4

The first part of the proof is to establish the consistency of $\widehat{H}_n$, while the second part is to verify the consistency of $\hat{\phi}_n(u)$.

(1) Consistency of $\widehat{H}_n$

This proof is based on Lemma 5.5 (Hariz et al., 2024) and the corollary (Hurvich et al., 1998). We can express $f_{H,\phi(u)}$ in the following form:

$$\tilde{f}_{H,\phi(u)}(\lambda) = (1 - \cos \lambda T)^{2d} f_{H,\phi(u)}(\lambda)$$

$$= (1 - \cos \lambda T)^{2d} (1 - 2\phi(u) \cos \lambda T + \phi^2(u))^{-1} f_{\epsilon_n^H}^H(\lambda)$$

$$= C_H (1 - \cos \lambda T)^{2d+1} (1 - 2\phi(u) \cos \lambda T + \phi^2(u))^{-1} \sum_{j \in \mathbf{Z}} \frac{1}{|\lambda + 2j\pi|^{2H+1}}.$$

According to Hurvich et al. (1998), it can be concluded that

$$\hat{d}_n - d = -\frac{1}{2S_m} \sum_{j=1}^{m} (a_j - \bar{a}_m) \log(\tilde{f}_{H,\phi(u)}) - \frac{1}{2S_m} \sum_{j=1}^{m} (a_j - \bar{a}_m) \epsilon_j,$$

where ϵ_j is the error defined in Equation (3) of Hurvich et al. (1998). According to Theorem 1 from the aforementioned sources, we have

$$\hat{d}_n \xrightarrow[n \to \infty]{\mathbb{P}} d.$$

Hence, it is evident that

$$\widehat{H}_n \xrightarrow[n \to \infty]{\mathbb{P}} H.$$

(2) Consistency of $\hat{\phi}_n(u)$

Assuming

$$\widetilde{\Phi}_i^j(u) = \left(\epsilon_{u+iT}^H, \epsilon_{u+(i+1)T}^H, \ldots, \epsilon_{u+jT}^H \right), \quad i \le j,$$

we can derive the following expression:

$$\hat{\phi}_n(u) - \phi(u) = \frac{\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u)}{\Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u)}. \tag{A5}$$

We apply the Taylor expansion of the matrix $\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)$ at H to the the numerator, yielding

$$\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u) = \widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}(u)$$

$$+ \widetilde{\Phi}_2^{n^*}(u)A_{nT}^{(1)}(H)\Phi_1^{n-1}(u)(\widehat{H}_n - H)$$

$$+ \frac{1}{2}\widetilde{\Phi}_2^{n^*}(u)A_{nT}^{(2)}(H)\Phi_1^{n-1}(u)(\widehat{H}_n - H)^2$$

$$+ \frac{1}{6}\widetilde{\Phi}_2^{n^*}(u)A_{nT}^{(3)}(\overline{H}_n)\Phi_1^{n-1}(u)(\widehat{H}_n - H)^3.$$

Thanks to the work of Hariz et al. (2024), we have the following three conclusions:

$$\frac{1}{n}\widetilde{\Phi}_2^{n^*}(u)A_{nT}^{(1)}(H)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} k_{H,\phi(u)}^{(1)}, \tag{A6}$$

$$\frac{1}{n}\widetilde{\Phi}_2^{n^*}(u)A_{nT}^{(2)}(H)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} k_{H,\phi(u)}^{(2)}, \tag{A7}$$

$$n^{-\frac{3}{2}}\widetilde{\Phi}_2^{n^*}(u)A_{nT}^{(3)}(\overline{H}_n)\Phi_1^{n-1}(u) = O_{\mathbb{P}}(1), \tag{A8}$$

where $k_{H,\phi(u)}^{(1)}$, $k_{H,\phi(u)}^{(2)}$ are constants, $\overline{H}_n \in B(H, |\widehat{H}_n - H|)$, and $A_{nT}^{(1)}(H)$, $A_{nT}^{(2)}(H)$, $A_{nT}^{(3)}(H)$ are

$$A_{nT}^{(1)}(H) = -\Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}\Gamma_{(n-1)T}^{-1}(H),$$

$$A_{nT}^{(2)}(H) = \Gamma_{(n-1)T}^{-1}(H)\frac{\partial^2\Gamma_{(n-1)T}(H)}{\partial^2 H}\Gamma_{(n-1)T}^{-1}(H)$$
$$+ 2\Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}\Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}\Gamma_{(n-1)T}^{-1}(H),$$

$$A_{nT}^{(3)}(H) = -\Gamma_{(n-1)T}^{-1}(H)\frac{\partial^3\Gamma_{(n-1)T}(H)}{\partial^3 H}\Gamma_{(n-1)T}^{-1}(H)$$
$$- 3\Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}\Gamma_{(n-1)T}^{-1}(H)\frac{\partial^2\Gamma_{(n-1)T}(H)}{\partial^2 H}\Gamma_{(n-1)T}^{-1}(H)$$
$$- 3\Gamma_{(n-1)T}^{-1}(H)\frac{\partial^2\Gamma_{(n-1)T}(H)}{\partial^2 H}\Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}\Gamma_{(n-1)T}^{-1}(H)$$
$$- 6\Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}\Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}$$
$$\times \Gamma_{(n-1)T}^{-1}(H)\frac{\partial\Gamma_{(n-1)T}(H)}{\partial H}\Gamma_{(n-1)T}^{-1}(H).$$

It has been demonstrated in Esstafa (2019) that

$$\frac{1}{n}\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} 0. \tag{A9}$$

The combination of Equations (A6), (A7), (A8) and (A9) allows us to deduce that

$$\frac{1}{n}\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} 0.$$

Next, we consider the asymptotic properties of the denominator of Equation (A5). We can similarly expand the denominator using a Taylor series around H, resulting in

$$\Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u) = \Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}(u)$$
$$+ \Phi_1^{n-1^*}(u)A_{nT}^{(1)}(H)\Phi_1^{n-1}(u)(\widehat{H}_n - H)$$
$$+ \frac{1}{2}\Phi_1^{n-1^*}(u)A_{nT}^{(2)}(H)\Phi_1^{n-1}(u)(\widehat{H}_n - H)^2.$$

Similarly, this part of the proof aligns with Lemma 1 (Esstafa, 2019) and satisfies

$$\frac{1}{n}\Phi_1^{n-1^*}(u)A_{nT}^{(1)}(H)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} k_{H,\phi(u)}^{(3)}, \tag{A10}$$

$$n^{-\frac{3}{2}}\Phi_1^{n-1^*}(u)A_{nT}^{(2)}(H)\Phi_1^{n-1}(u) = O_{\mathbb{P}}(1), \tag{A11}$$

$$\frac{1}{n}\Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{1-\phi^2(u)}, \tag{A12}$$

where $k_{H,\phi(u)}^{(3)}$ is a constant. The denominator in Equation (A5) converges in probability as follows:

$$\frac{1}{n}\Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{1-\phi^2(u)}.$$

Combining the above equations, we find that when the numerator of Equation (A5) is multiplied by $\frac{1}{n}$, it approaches 0, while the denominator, also multiplied by $\frac{1}{n}$, converges to a constant. Furthermore, since convergence in probability implies convergence in distribution, we conclude that

$$\hat{\phi}_n(u) \xrightarrow[n\to\infty]{\mathbb{P}} \phi(u).$$

This establishes a clear relationship between the asymptotic behavior of the numerator and denominator, leading to the convergence of the estimated function.

A.5. Proof of Theorem 2.5

According to Theorem 2 (Hurvich et al., 1998), without loss of generality, we can assume $m = [n^\delta]$ for some $\frac{1}{2} < \delta < \frac{2}{3}$. Let $\mathcal{L}$ denote convergence in distribution. We thus have

$$\sqrt{m}(\widehat{H}_n - H) \xrightarrow[n\to\infty]{\mathcal{L}} \mathcal{N}(0, U_{K,\delta}),$$

where $U_{K,\delta}$ is a constant related to K and δ. Building on the results from Equation (A5), we establish that

$$\sqrt{m}(\hat{\phi}_n(u) - \phi(u)) = \sqrt{m}\frac{\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}(u)}{\Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u)} + R_n^{(1)}. \tag{A13}$$

According to the proof of consistency and some results on Esstafa (2019), the denominator of the first term on the right side of the above equation satisfies

$$\frac{1}{n}\Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} \frac{1}{1-\phi^2(u)},$$

the nominator converges to a normal distribution

$$\frac{1}{\sqrt{n}}\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}(u) \xrightarrow[n\to\infty]{\mathbb{P}} \mathcal{N}(0, \frac{1}{1-\phi^2(u)}),$$

when $n \to \infty$, and the reminder $R_n^{(1)}$ converges to 0. Thus we can rewrite Equation (A13) as follows:

$$\sqrt{m}(\hat{\phi}_n(u) - \phi(u)) = \sqrt{m}\frac{\frac{1}{n}\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}}{\frac{1}{n}\Phi_1^{n-1^*}(u)\Gamma_{(n-1)T}^{-1}(\widehat{H}_n)\Phi_1^{n-1}(u)} + R_n^{(1)}$$

$$= \frac{\sqrt{m}}{\sqrt{n}}(1-\phi^2(u))\frac{1}{\sqrt{n}}\widetilde{\Phi}_2^{n^*}(u)\Gamma_{(n-1)T}^{-1}(H)\Phi_1^{n-1}(u) + R_n^{(1)}.$$

By Slutsky theorem, we can conclude that $\sqrt{m}(\hat{\phi}_n(u) - \phi(u))$ converges to a normal distribution.

Lastly, we aim to present these results in the form of a joint normal distribution. Drawing on the findings from Hariz et al. (2024), the asymptotic distribution of $\sqrt{n}(\hat{\phi}_n(u) - \phi(u))$ can be expressed as a constant multiple of the asymptotic distribution of $(\widehat{H}_n - H)$. Moreover, according to the Cramer–Wold theorem, the asymptotic distribution of $\sum_{u=1}^{T}(\hat{\phi}_n(u) - \phi(u))$ still adheres to an asymptotic normal distribution. Thus the vector

$$\left((\hat{\phi}_n(1) - \phi(1)), (\hat{\phi}_n(2) - \phi(2)), \ldots, (\hat{\phi}_n(T) - \phi(T)), \ldots, \widehat{H}_n - H\right)$$

converges to a Gaussian vector, tending towards a joint normal distribution. The covariance matrix of this vector is

$$\widetilde{\Sigma}_\theta = U_\theta U_\theta^\top, \tag{A14}$$

where

$$U_\theta = \left(1, \frac{C_{H,\phi(1)}^{(1)}}{(1-\phi^2(1))}, \ldots, \frac{C_{H,\phi(T)}^{(1)}}{(1-\phi^2(T))}\right)^\top,$$

and $C_{H,\phi(u)}^{(1)}$ is the constants related to $\phi(u)$ and H.

A.6. Proof of Theorem 3.1

To prove Theorem 3.1, we need to establish the following three lemmas and verify whether $g_{H,\underline{\phi}(\lambda)}$ is regular. The regularity conditions of $g_{H,\underline{\phi}(\lambda)}$ will be demonstrated in the auxiliary results.

Lemma A.2: *Let $\theta_0 \in \Theta$, $\delta > 0$, such that for any $\theta \in B(\theta_0, \delta)$, it holds that*

$$\|\mathcal{I}(\theta) - \mathcal{I}(\theta_0)\| \leq K\|\theta - \theta_0\|,$$

where K is some constant.

Proof: Without loss of generality, let $B(\theta_0, \delta)$ be a convex set in $\mathbb{R}^3$. For ease of notation, $g_{H,\phi}(\lambda)$ can be denoted as $g_\theta(\lambda)$. According to the relevant conclusions in Cohen et al. (2013) and the discussion of regularity conditions for $g_{H,\underline{\phi}}(\lambda)$, it is known that for any $k, j \in \{1, 2, \ldots, d\}$, the following inequality holds:

$$\left| \frac{1}{4\pi} \left(\int_{-\pi}^{\pi} \frac{\partial \log g_\theta(\lambda)}{\partial \theta_k} \frac{\partial \log g_\theta(\lambda)}{\partial \theta_j} \, d\lambda \right) - \frac{1}{4\pi} \left(\int_{-\pi}^{\pi} \frac{\partial \log g_{\theta_0}(\lambda)}{\partial \theta_{0,k}} \frac{\partial \log g_{\theta_0}(\lambda)}{\partial \theta_{0,j}} \, d\lambda \right) \right| \leq K\|\theta - \theta_0\|,$$

where K is defined as

$$K = \sup_{\theta \in B(\theta_0, \delta)} \left(\frac{\partial}{\partial \theta_i} \left(\int_{-\pi}^{\pi} \frac{\partial \log g_\theta(\lambda)}{\partial \theta_k} \frac{\partial \log g_\theta(\lambda)}{\partial \theta_j} \, d\lambda \right) \right)_{1 \leq i \leq d},$$

which is related to k and j. Furthermore, since the conditions (A1) and (A2) (Cohen et al., 2013) hold, it follows that $K < \infty$, hence the lemma holds. ∎

Lemma A.3: *For any $\theta \in \Theta^l$, it follows from the distribution of the parameter θ that*

$$\frac{\Delta l_n(\theta)}{\sqrt{n}} + \sqrt{n}\mathcal{I}(\theta) = O_{\mathbb{P}}(1).$$

Proof: The Lemma 3.6 (Cohen et al., 2013) implies that, from the distribution of θ, we have

$$E\left(\frac{\Delta l_n(\theta)}{n}\right) \to -\mathcal{I}(\theta).$$

To determine the convergence rate of the above expression, Lemma 3 and Lemma 4 (Lieberman et al., 2012) yield the following conclusion:

$$E\left(\frac{\Delta l_n(\theta)}{n}\right) + \mathcal{I}(\theta) = O(n^{-1+\delta}),$$

where δ is a positive real number. Therefore,

$$E\left(\frac{\Delta l_n(\theta)}{\sqrt{n}}\right) + \sqrt{n}\mathcal{I}(\theta) = O(n^{-\frac{1}{2}+\delta}).$$

Furthermore, by utilizing Lemma 3.6 (Cohen et al., 2013) once again, we obtain

$$\text{Var}\left(\frac{\Delta l_n(\theta)}{\sqrt{n}}\right) = O(1).$$

Thus the proof is concluded. ∎

Lemma A.4: *Let $\{\overline{\theta}_n\}_n$ be a stochastic sequence satisfying $\overline{\theta}_n - \theta = o_{\mathbb{P}}(1)$. Then, according to the distribution of parameter θ, for any $k > 0$, it holds that*

$$\frac{\Delta l_n(\overline{\theta}_n)}{n} - \frac{\Delta l_n(\theta)}{n} = O_{\mathbb{P}}(n^k(\overline{\theta}_n - \theta)).$$

Proof: Let $C_{k,\theta}$ be a compact convex set depending on k and θ, and $\overline{\theta}_n \in C_{k,\theta}$. According to the proof of Lemma 3.7 (Cohen et al., 2013), we have

$$\sup_{\overline{\theta}_n \in C_{k,\theta}} \left| \frac{\partial^3}{\partial^{i_1}\theta_1 \partial^{i_2}\theta_2 \cdots \partial^{i_d}\theta_d} \frac{l_n(\overline{\theta}_n)}{n^{1+k}} \right| = O_{\mathbb{P}}(1),$$

where $(i_1, i_2, \ldots, i_d) \in \{0, 1, 2, 3\}^d$, satisfying $i_1 + i_2 + \cdots + i_d = 3$. In conclusion, for a finite positive random variable K, we have

$$P\left(\left\| \frac{\Delta l_n(\overline{\theta}_n)}{n} - \frac{\Delta l_n(\theta)}{n} \right\| \leq K n^k (\|\overline{\theta}_n - \theta\|) \right) \geq P(\overline{\theta}_n \in C_{k,\theta}),$$

which implies $\frac{\Delta l_n(\overline{\theta}_n)}{n} - \frac{\Delta l_n(\theta)}{n} = O_{\mathbb{P}}(n^k(\overline{\theta}_n - \theta))$ holds. ∎

According to the hypothesis of this theorem, we can deduce

$$\sqrt{n}(\tilde{\theta}_n - \theta) = \sqrt{n}(\hat{\theta}_n - \theta) + \mathcal{I}^{-1}(\hat{\theta}_n)\frac{\nabla l_n(\hat{\theta}_n)}{\sqrt{n}}. \tag{A15}$$

Applying mean-value theorem to $\nabla l_n(\theta)$, we have

$$\nabla l_n(\hat{\theta}_n) = \nabla l_n(\theta) + (\hat{\theta}_n - \theta) \int_0^1 \triangle l_n(\theta + v(\hat{\theta}_n - \theta)) \, dv, \tag{A16}$$

Substituting Equations (A16) to (A15), we produce

$$\sqrt{n}(\tilde{\theta}_n - \theta) = \sqrt{n}(\hat{\theta}_n - \theta)\mathcal{I}^{-1}(\hat{\theta}_n)\left(\mathcal{I}(\hat{\theta}_n) + \frac{\int_0^1 \triangle l_n(\theta(v)) \, dv}{n} \right) + \mathcal{I}^{-1}(\hat{\theta}_n)\frac{\nabla l_n(\theta)}{\sqrt{n}}, \tag{A17}$$

where $\theta(v) = \theta + v(\hat{\theta}_n - \theta)$, $v < 1$. Next, we will discuss the consistency and asymptotic normality of one-step estimator.

(1) Consistency of $\tilde{\theta}_n$

Observing Equation (A17), the first and second terms on the right-hand side can be expressed as

$$A_n = \sqrt{n}(\hat{\theta}_n - \theta)\mathcal{I}^{-1}(\hat{\theta}_n)\left(\mathcal{I}(\hat{\theta}_n) + \frac{\int_0^1 \triangle l_n(\theta(v)) \, dv}{n} \right)$$

$$= \sqrt{n^\delta}(\hat{\theta}_n - \theta)\mathcal{I}^{-1}(\hat{\theta}_n)\sqrt{n^{1-\delta}}\left(\mathcal{I}(\hat{\theta}_n) + \frac{\int_0^1 \triangle l_n(\theta(v)) \, dv}{n} \right)$$

and

$$B_n = \mathcal{I}^{-1}(\hat{\theta}_n)\frac{\nabla l_n(\theta)}{\sqrt{n}} = \mathcal{I}^{-1}(\theta)\frac{\nabla l_n(\theta)}{\sqrt{n}} + \left(\mathcal{I}^{-1}(\hat{\theta}_n) - \mathcal{I}^{-1}(\theta) \right)\frac{\nabla l_n(\theta)}{\sqrt{n}}.$$

First, we analyse the properties of A_n and derive the following equation:

$$\mathcal{I}(\hat{\theta}_n) + \frac{\int_0^1 \triangle l_n(\theta(v)) \, dv}{n} = (\mathcal{I}(\hat{\theta}_n) - \mathcal{I}(\theta_n))$$

$$+ \left(\mathcal{I}(\theta_n) + \frac{\Delta l_n(\theta)}{n} \right) + \frac{1}{\sqrt{n}}\int_0^1 \left(\frac{\Delta l_n(\theta(v))}{\sqrt{n}} - \frac{\Delta l_n(\theta)}{\sqrt{n}} \right) dv, \tag{A18}$$

based on Equation (A18) and Lemmas 6.2, 6.3, and 6.4. The convergence order of $\frac{A_n}{\sqrt{n}}$ is

$$\frac{A_n}{\sqrt{n}} = n^{-\frac{\delta}{2}} \left(O_\mathbb{P}(n^{-\frac{\delta}{2}}) + O_\mathbb{P}(n^{-\frac{1}{2}}) + O_\mathbb{P}(n^{k-\frac{\delta}{2}}) \right).$$

Thus, when $k - \delta < 0$, we have $\frac{A_n}{\sqrt{n}} \xrightarrow[n \to \infty]{\mathbb{P}} 0$.

Second, we consider the property of B_n and it has the form of

$$B_n = \mathcal{I}^{-1}(\hat{\theta}_n) \frac{\nabla l_n(\theta)}{\sqrt{n}} = \mathcal{I}^{-1}(\theta) \frac{\nabla l_n(\theta)}{\sqrt{n}} + (\mathcal{I}^{-1}(\hat{\theta}_n) - \mathcal{I}^{-1}(\theta)) \frac{\nabla l_n(\theta)}{\sqrt{n}}.$$

According to Hariz et al. (2024) and Theorem 1 in Lieberman et al. (2012), we have

$$\frac{\nabla l_n(\theta)}{\sqrt{n}} \xrightarrow[n \to \infty]{\mathbb{P}} 0.$$

When $\mathcal{I}n(\cdot)$ is a non-degenerate continuous function, as indicated by the above equation, it can be observed that both the first and second terms of B_n tend to 0. Consequently, $\frac{B_n}{\sqrt{n}}$ converges in probability to 0, and naturally, it also converges in distribution to 0.

Combining the above results, we can conclude the consistency of $\tilde{\theta}_n$.

(2) Asymptotic normality of $\tilde{\theta}_n$

According to the results of Hariz et al. (2024), the equation

$$\mathcal{I}^{-1}(\theta) \frac{\nabla l_n(\theta)}{\sqrt{n}} + (\mathcal{I}^{-1}(\hat{\theta}_n) - \mathcal{I}^{-1}(\theta)) \frac{\nabla l_n(\theta)}{\sqrt{n}} \tag{A19}$$

converges in probability to a bounded limit as $n \to \infty$. Simultaneously, the second term on the right-hand side of Equation (A19) converges to 0. By applying the Slutsky theorem, we can verify the asymptotic normality of $\tilde{\theta}_n$.

Appendix 2. Auxiliary results

Lemma A.5: *Under the hypothesis on the parametric space, we have the following results hold.*

(1) *For any $H \in [0, 1]$ and $j \in \{0, 1, 2, 3\}$, consider the expression $\frac{\partial}{\partial \lambda} \frac{\partial^j}{\partial^j H} g_{H,\underline{\phi}}(\lambda)$. Here $g_{H,\underline{\phi}}(\lambda)$ denotes the spectral density of Y_n as specified in Equation (11).*

(2) *For any $j \in \{0, 1, 2, 3\}$ the functions $\frac{\partial^j}{\partial^j H} g_{H,\underline{\phi}}(\lambda)$ are symmetric with respect to λ.*

(3) *For any $\delta > 0$ and all $(H, \lambda) \in [0, 1] \times [-\pi, \pi] \setminus \{0\}$, the following bounds apply.*
   (a) $C_{1,\delta} |\lambda|^{1-2H+\delta} \le g_{H,\underline{\phi}}(\lambda) \le C_{2,\delta} |\lambda|^{1-2H-\delta}.$
   (b) $|\frac{\partial}{\partial \lambda} g_{H,\underline{\phi}}(\lambda)| \le C_{3,\delta} |\lambda|^{-2H-\delta}.$
   (c) *For any $j \in \{0, 1, 2, 3\}$, $|\frac{\partial^j}{\partial^j H} g_{H,\underline{\phi}}(\lambda)| \le C_{4,\delta} |\lambda|^{-2H-\delta},$*
       where $C_{i,\delta}$ represents constants for $i = 1, 2, 3, 4$.

These results establish that $g_{H,\underline{\phi}}(\lambda)$ satisfies the regularity conditions.

Proof: We start from Assertion (3)(a), which states that

$$g_{H,\underline{\phi}}(\lambda) = C_H \left| \sum_{p=0}^{T-1} \frac{\mathrm{e}^{-\mathrm{i}p\lambda}}{1 - \phi(T)\mathrm{e}^{-\mathrm{i}\lambda T}} \right|^2 (1 - \cos(\lambda)) \sum_{j \in Z} \frac{1}{|\lambda + 2j\pi|^{2H+1}},$$

where $C_H = \frac{1}{2\pi} \Gamma(2H + 1) \sin(\pi H)$ and $\Gamma(\cdot)$ denotes the Gamma function. According to Lemma 5.4 in Hariz et al. (2024), we have

$$K_{1,\delta} |\lambda|^{1-2H+\delta} \le C_H(1 - \cos(\lambda)) \sum_{j \in Z} \frac{1}{|\lambda + 2j\pi|^{2H+1}} \le K_{2,\delta} |\lambda|^{1-2H-\delta}$$

and

$$\frac{\sqrt{2}T}{1 - \phi_{\max}^2(u)} \le \left| \sum_{p=0}^{T-1} \frac{e^{-ip\lambda}}{1 - \phi(T)e^{-i\lambda T}} \right|^2 \le \frac{\sqrt{2}T}{1 - \phi_{\min}^2(u)},$$

where $\phi_{\max}^2(u)$ and $\phi_{\min}^2(u)$ are the maximum and minimum values of $\phi(u)$, respectively. Thus Assertion (3)(a) has been proved and Assertion (3)(b) follows straightforwardly from Assertion (3)(a).

Next, we discuss Assertion (3)(c), which can be obtained directly from lemma 5.4 in Hariz et al. (2024). The partial derivative of $g_{H,\underline{\phi}}(\lambda)$ does not depend on $|\sum_{p=0}^{T-1} \frac{e^{-ip\lambda}}{1-\phi(T)e^{-i\lambda T}}|^2$, and the modulus is bounded. ∎