STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 202x, pp 0–11.
Published online in International Academic Press (www.IAPress.org)

Optimal Tests for Distinguishing PAR Models from PSETAR Models

N. Bezziche, M. Merzougui *

LaPS Laboratory, Badji Mokhtar-Annaba University, 12, P.O.Box, 23000 Annaba, Algeria

Abstract This paper aims to detect nonlinearity in periodic autoregressive models. We introduce parametric and semiparametric local asymptotic optimal tests designed for distinguishing a periodic autoregressive model from a periodic self-exciting threshold autoregressive (SETAR) model. Leveraging the Local Asymptotic Normality (LAN) property specific to periodic SETAR models, we devise a parametric test that is locally asymptotically most stringent. Additionally, the utilization of kernel estimation for the density function allows the construction of an adaptive test for enhanced flexibility and accuracy in detecting nonlinearity. The performance of the proposed tests is assessed through simulation studies.

Keywords Periodic Self-Exciting Threshold Autoregressive, Local Asymptotic Normality, Local Asymptotic optimal test, Kernel estimation.

AMS 2010 subject classifications 62M10, 62M10

DOI: 10.19139/soic-2310-5070-2240

1. Introduction

The manifestation of periodicity in both mean and covariance features is a common observation in various seasonal time series. In response to this, numerous authors have proposed distinct periodic linear and nonlinear time series models, including noteworthy contributions by [8], [7], [6], [16], [2] and [1]. An extension of the Self-Exciting Threshold Autoregressive model, originally introduced by [18], emerged in the form of a periodic version thanks to the work of [13]. This adaptation proved successful in modeling monthly Fraser River flow time series with inherent nonlinear characteristics. The model incorporates a SETAR representation in each period, featuring a predetermined number of regimes (p) and an autoregressive (AR) formulation within each regime. For practical purposes, we assume a known number of regimes and thresholds, as the selection of thresholds typically relies on heuristic methods. Appendix B provides the R code implementing a grid-search technique to estimate thresholds in a periodic setting. This approach minimizes the sum of squared residuals (SSR) within the periodic autoregressive regimes, ensuring an optimal threshold selection for model performance. Simultaneously, we set the AR order to one for ease of calculation and presentation; however, vector representation is readily achievable if a different order is desired. Notably, when all parameters within a period are identical, the model reduces to the periodic autoregressive model (PAR). The Local Asymptotic Normality property, introduced by [11], serves as a pivotal tool in assessing the local optimality of estimation and testing procedures. Various researchers, including [17], [10], [5], [15], [4], [3] among others, have explored the LAN property in different model contexts. In this paper, we leverage the LAN property to construct a locally asymptotically most stringent test for linearity against a PSETAR(p,1,...,1) model. While [14] addresses the case of p=2, our contribution delves deeper to develop a locally adaptive test by estimating the innovation density using the kernel method. This refinement is particularly motivated by the need for more robust statistical methods to detect nonlinearity in periodic time series models. In this context, the current

*Correspondence to: M. Merzougui (Email: mouna.merzougui@univ-annaba.dz). Department of Mathematics, Badji Mokhtar-Annaba University, 12, P.O.Box, 23000 Annaba, Algeria.

ISSN 2310-5070 (online) ISSN 2311-004X (print)

paper builds on the LAN property to provide a more flexible and efficient test for nonlinearity. Specifically, our main contribution lies in extending previous parametric tests for linearity in PSETAR models by incorporating a nonparametric density estimation step, enhancing adaptability to a wider range of innovation distributions. The structure of the paper is organized as follows: Section 2 presents the central tool of our analysis, namely the LAN property for PSETAR(p,1,...,1). Section 3 focuses on the construction of parametric locally asymptotically most stringent tests. The adaptive test is elucidated subsequent to the estimation of the density in Section 4. Finally, Section 5 provides the simulation results, demonstrating the effectiveness of the proposed methodology.

2. Local Asymptotic Normality

2.1. Notation and Assumptions

The periodic Autoregressive PAR(1) model, characterized by a period S, is expressed as:

$$X_t = \phi_t X_{t-1} + \varepsilon_t, \quad t \in \mathbb{Z}, \tag{1}$$

where $\{\varepsilon_t; t \in \mathbb{Z}\} \sim \text{iid}(0, \sigma_t^2)$. Conversely, the process X_t adheres to a periodic self-exciting threshold autoregressive PSETAR$(p, 1, \dots, 1)$ model with a period S if it satisfies:

$$X_t = \sum_{i=1}^{p} \phi_{t,i} X_{t-1} I(c_{t,i-1} \leq X_{t-1} \leq c_{t,i}) + \varepsilon_t, \quad t \in \mathbb{Z}, \tag{2}$$

where $\{\varepsilon_t; t \in \mathbb{Z}\}$ is an iid$(0, \sigma_t^2)$ process, with an unknown density f. The autoregressive parameters $\phi_{t,1}, \dots, \phi_{t,p}$, the innovation variance σ_t^2, and the thresholds are periodic in time with a period S. We set $c_{t,0} = -\infty$ and $c_{t,p} = +\infty$, $\forall t$. For convenience, the threshold parameters are assumed to be known. This assumption is justified as their estimation is relatively subjective. In these models, the regime is specified by the values of the thresholds, and depending on the observed variable X_{t-1}, we derive X_t.

The model can be expressed in an alternative form utilizing indicator functions:

$$X_t = \phi_{t,1} X_{t-1}^1 + \cdots + \phi_{t,p} X_{t-1}^p + \varepsilon_t, \tag{3}$$

where

$$X_{t-1}^i = X_{t-1} I(c_{t,i-1} < X_{t-1} \leq c_{t,i}). \tag{4}$$

To provide greater specificity, let $t = S\tau + s$, leading to

$$X_{S\tau+s} = \phi_{s,1} X_{S\tau+s-1}^1 + \cdots + \phi_{s,p} X_{S\tau+s-1}^p + \varepsilon_{S\tau+s}, \quad s = 1, \dots, S, \quad \tau \in \mathbb{Z}. \tag{5}$$

Here, $H_f^{(n)}(\underline{\phi})$ represents a sequence of null hypotheses in which $\{X_t^{(n)}, t \in \mathbb{Z}\}$ is a sequence of realizations of a process satisfying the model (1), where $\underline{\phi} = (\phi_1, \phi_2, \dots, \phi_S)'$, and $H_f^{(n)}(\underline{\phi}^{(n)})$ denotes the sequence of alternative hypotheses under which the sequence $\{X_t^{(n)}, t \in \mathbb{Z}\}$ is a sequence of realizations of a process satisfying the periodic threshold autoregressive model (2), where

$$\underline{\phi}^{(n)} = (\underline{\phi}_1^{(n)'}, \underline{\phi}_2^{(n)'}, \dots, \underline{\phi}_S^{(n)'})' \in \mathbb{R}^{pS}, \tag{6}$$

$$\underline{\phi}_s^{(n)} = \left(\phi_s + \frac{1}{\sqrt{n}} h_{s,1}^{(n)}, \dots, \phi_s + \frac{1}{\sqrt{n}} h_{s,p}^{(n)} \right)', \quad s = 1, \dots, S. \tag{7}$$

Considering the notations:

$$\underline{\Phi} = (\phi_1, \dots, \phi_1; \dots; \phi_S, \dots, \phi_S)' \in \mathbb{R}^{pS}, \tag{8}$$

$$\underline{\tau}^{(n)} = (\underline{h}_1^{(n)'}, \dots, \underline{h}_p^{(n)'})', \quad \text{where} \quad \underline{h}_i^{(n)} = (h_{1,i}^{(n)}, h_{2,i}^{(n)}, \dots, h_{S,i}^{(n)})', \quad i = 1, \dots, p. \tag{9}$$

The terms $\underline{h}_i^{(n)}$, where $i = 1, \ldots, p$, can be construed as local periodic perturbations of the periodic parameters ϕ_s, such that $\|h\|$ is bounded. It is evident that if $\underline{h}_1^{(n)} = \cdots = \underline{h}_p^{(n)}$, then the model (2) simplifies to (1). Let $\underline{\nu}^{(n)}$ be the $pS \times pS$ matrix defined as

$$\underline{\nu}^{(n)} = \frac{1}{\sqrt{n}} K,$$

where the matrix K is specified as follows:

$$K = \begin{bmatrix} K_{1,1} & K_{1,2} & \cdots & K_{1,p} \\ K_{2,1} & K_{2,2} & \cdots & K_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ K_{S-1,1} & K_{S-1,2} & \cdots & K_{S-1,p} \\ K_{S,1} & K_{S,2} & \cdots & K_{S,p} \end{bmatrix},$$

where the $p \times S$-matrices $K_{k,l}$, for $k = 1, 2, \ldots, S$ and $l = 1, \ldots, p$, have general elements defined as:

$$(K_{k,l})_{i,j} = \begin{cases} 1, & \text{if } i = l, j = k, \\ 0, & \text{otherwise.} \end{cases}$$

With these notations, the alternative hypotheses can be expressed in the form:

$$H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)} \underline{\tau}^{(n)}).$$

Throughout this paper, we establish the following assumptions:

Assumption (A1).

The autoregressive parameters $\underline{\phi}$ satisfy the necessary and sufficient causality condition of (1), i.e.,

$$\left| \prod_{s=1}^{S} \phi_s \right| < 1.$$

Assumption (A2).

The innovation density $f(.)$ is presumed to satisfy the following conditions:

- $f(x) > 0, \quad \forall x \in \mathbb{R}$,
- $f(.)$ is absolutely continuous with respect to the Lebesgue measure μ,
- the Fisher information

$$I(f) = \int (\varphi_f(x))^2 f(x) dx$$

is finite, where $\varphi_f = -\frac{f'(.)}{f(.)}$,

- $\int x f(x) dx = 0$ and the variance is finite, i.e.,

$$\sigma_t^2 = E(X_t^2) < \infty.$$

2.2. *Local Asymptotic Normality*

Suppose, for simplicity of notation, that the size of the observed time series n is a multiple of S, i.e., $n = mS$, with $m \in \mathbb{N}^*$, and let $t = s + S\tau$, where $s = 1, \ldots, S$ and $\tau = 0, 1, \ldots, m-1$.

Denote by $Z_t^{(n)}(\underline{\phi})$ and $Z_t^{(n)}(\underline{\phi}^{(n)})$, $t \in \mathbb{Z}$, the calculated residuals under $H_f^{(n)}(\underline{\phi})$ and $H_f^{(n)}(\underline{\phi}^{(n)})$, respectively. Then, we have:

$$Z_{s,\tau}^{(n)}(\underline{\phi}_s^{(n)}) = Z_{s,\tau}^{(n)}(\underline{\phi}_s) - \frac{1}{\sqrt{n}} h_{s,1}^{(n)} X_{S\tau+s-1}^{(n)1} - \cdots - \frac{1}{\sqrt{n}} h_{s,p}^{(n)} X_{S\tau+s-1}^{(n)p}$$

$$= Z_{s,\tau}^{(n)}(\underline{\phi}_s) - \frac{1}{\sqrt{n}}\underline{\tau}_s^{*(n)'}\underline{X}_{s-1+S\tau}^{(n)} = Z_{s,\tau}^{(n)}(\underline{\phi}_s) - \gamma_{s,\tau}^{(n)},$$

where $\underline{\tau}^{*(n)} = K\underline{\tau}^{(n)}$ with $\underline{\tau}_s^{*(n)} = (h_{s,1}^{(n)}, \ldots, h_{s,p}^{(n)})$, and

$$\underline{X}_{s-1+S\tau}^{(n)} = (X_{S\tau+s-1}^{(n)1}, \ldots, X_{S\tau+s-1}^{(n)p})'.$$

for $s = 1, \ldots, S$ and $S \in \mathbb{Z}$. The logarithm of the likelihood ratio, $\Lambda_f^{(n)}(\underline{\phi}^{(n)}) = \Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)})$ is given by:

$$\Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)}) = \sum_{s=1}^{S}\sum_{\tau=0}^{m-1}\left[\log(f_{\sigma_s}(Z_{s,\tau}^{(n)}(\underline{\phi}_s) - \gamma_{s,\tau}^{(n)})) - \log(f_{\sigma_s}(Z_{s,\tau}^{(n)}(\underline{\phi}_s)))\right] + o_P(1).$$

We denote $\Gamma_s(\underline{\phi})$ as the variance-covariance matrix of $\underline{X}_{s-1+S\tau}$, where $s = 1, \ldots, S$ and $\tau \in \mathbb{Z}$:

$$\Gamma_s(\underline{\phi}) = \text{Cov}(\underline{X}_{s-1+S\tau}, \underline{X}'_{s-1+S\tau}), \quad s = 1, \ldots, S, \quad \tau \in \mathbb{Z}.$$

This matrix captures the statistical relationships among the components of $\underline{X}_{s-1+S\tau}$ for various values of s and τ, leading to the result:

$$\Gamma_s(\underline{\phi}) = \begin{pmatrix} E(X_{s-1}^1)^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & E(X_{s-1}^p)^2 \end{pmatrix} = \begin{pmatrix} \Gamma_{s,1}(\underline{\phi}) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Gamma_{s,p}(\underline{\phi}) \end{pmatrix}.$$

and

$$\Gamma(\underline{\phi}) = \begin{pmatrix} \frac{\Gamma_1(\underline{\phi})}{\sigma_1^2} & 0 & \cdots & 0 \\ 0 & \frac{\Gamma_2(\underline{\phi})}{\sigma_2^2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \frac{\Gamma_S(\underline{\phi})}{\sigma_S^2} \end{pmatrix}_{pS \times pS}$$

and we define the sequence

$$\underline{\delta}^{(n)}(\underline{\phi}) = \begin{pmatrix} \delta_{1,1}^{(n)}(\underline{\phi}_1) \\ \vdots \\ \delta_{1,p}^{(n)}(\underline{\phi}_1) \\ \vdots \\ \delta_{S,1}^{(n)}(\underline{\phi}_S) \\ \vdots \\ \delta_{S,p}^{(n)}(\underline{\phi}_S) \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1}\varphi_{\sigma_1}(Z_{1,\tau}^{(n)}(\underline{\phi}_1))X_{S\tau}^{(n)1} \\ \vdots \\ \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1}\varphi_{\sigma_1}(Z_{1,\tau}^{(n)}(\underline{\phi}_1))X_{S\tau}^{(n)p} \\ \vdots \\ \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1}\varphi_{\sigma_S}(Z_{S,\tau}^{(n)}(\underline{\phi}_S))X_{S-1+S\tau}^{(n)1} \\ \vdots \\ \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1}\varphi_{\sigma_S}(Z_{S,\tau}^{(n)}(\underline{\phi}_S))X_{S-1+S\tau}^{(n)p} \end{pmatrix}$$

Proposition 2.1

Assuming the validity of assumptions A1 and A2, under $H_f^{(n)}(\underline{\phi})$ as $n \to \infty$, we assert the following:

1. $\Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)}) = \underline{\tau}^{(n)'}\underline{\Delta}_f^{(n)}(\underline{\phi}) - \frac{1}{2}\underline{\tau}^{(n)'}\Gamma^{\Delta_f^{(n)}}(\underline{\phi})\underline{\tau}^{(n)} + o_p(1)$,
2. $\underline{\Delta}_f^{(n)}(\underline{\phi}) \xrightarrow{d} N_{pS}(0, \Gamma^{\Delta_f^{(n)}}(\underline{\phi}))$.

where the square matrix $\Gamma^{\Delta_f^{(n)}}(\underline{\phi}) = \frac{I(f_1)}{S} K'\Gamma(\underline{\phi})K$, and the central sequence $\underline{\Delta}_f^{(n)}(\underline{\phi})$ is given by:

$$\underline{\Delta}_f^{(n)}(\underline{\phi}) = K'\underline{\delta}^{(n)}(\underline{\phi}),$$

where

$$\underline{\Delta}_f^{(n)}(\underline{\phi}) = (\delta_{1,1}^{(n)}, \delta_{2,1}^{(n)}, \dots, \delta_{S,1}^{(n)}; \dots; \delta_{1,p}^{(n)}, \delta_{2,p}^{(n)}, \dots, \delta_{S,p}^{(n)})'.$$

Proof. See the appendix.

Corollary 2.1

Under the LAN property, the following results hold:

1. $\underline{\Delta}_f^{(n)}(\underline{\phi}) \xrightarrow{d} N(\Gamma^{\Delta_f^{(n)}}(\underline{\phi})\underline{\tau}^{(n)}, \Gamma^{\Delta_f^{(n)}}(\underline{\phi}))$ under $H_f^{(n)}(\underline{\phi}^{(n)})$,
2. $\Lambda_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)}) \xrightarrow{d} N\left(\frac{1}{2}\underline{\tau}^{(n)\prime}\Gamma^{\Delta_f^{(n)}}(\underline{\phi})\underline{\tau}^{(n)}, \underline{\tau}^{(n)\prime}\Gamma^{\Delta_f^{(n)}}(\underline{\phi})\underline{\tau}^{(n)}\right)$ under $H_f^{(n)}(\underline{\phi}^{(n)})$,
3. The hypotheses $H_f^{(n)}(\underline{\phi})$ and $H_f^{(n)}(\underline{\phi}^{(n)})$ are contiguous,
4. The central sequence $\underline{\Delta}^{(n)}(\underline{\phi})$ verifies the local asymptotic linearity property:

$$\underline{\Delta}^{(n)}(\underline{\phi}^{(n)}) - \underline{\Delta}^{(n)}(\underline{\phi}) = -\frac{I(f_1)}{S}\Gamma^{\Delta_f^{(n)}}(\underline{\phi})\underline{\tau}^{(n)} + o_p(1).$$

These results bring greater understanding to the asymptotic dynamics of the system being studied.

3. Optimal parametric Test

The LAN structure indicates the convergence of the local experiment:

$$\xi_f^{(n)}(\underline{\phi}) = \left\{P_{\underline{\Phi}+\underline{\nu}^{(n)}\underline{\tau}^{(n)}}; \underline{\tau}^{(n)} \in \mathbb{R}^{pS}, \sup_n \underline{\tau}^{(n)\prime}\underline{\tau}^{(n)} < \infty, \underline{\tau}^{(n)} \to \underline{\tau} \text{ as } n \to \infty\right\}$$

to the Gaussian shift experiment of dimension pS:

$$E_f = N\left(\Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \underline{\sigma})\underline{\tau}, \Gamma^{\Delta_f^{(n)}}(\underline{\phi}, \underline{\sigma})\right), \quad \underline{\tau} \in \mathbb{R}^{pS}.$$

Let

$$\eta = \Gamma^{\Delta}(\underline{\phi}, \underline{\sigma}) \begin{bmatrix} \underline{h}_1 \\ \vdots \\ \underline{h}_p \end{bmatrix}$$

The test problem of the null hypothesis $H_f^{(n)}(\underline{\phi})$: $\text{PAR}_S(1)$ model versus the local alternative $H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}^{(n)})$: local S- PSETAR$(p, 1, \dots, 1)$ model can be framed as follows: testing

$$\eta \in M(\Gamma^{\Delta}(\underline{\phi})\Omega) \quad \text{versus} \quad \eta \notin M(\Gamma^{\Delta}(\underline{\phi})\Omega),$$

where the $pS \times S$ matrix Ω is defined as

$$\Omega = \begin{pmatrix} I_{S\times S} \\ \vdots \\ I_{S\times S} \end{pmatrix},$$

and $M(\Omega)$ represents the subspace of $\mathbb{R}^{pS}$ generated by Ω. Following [12], the most stringent size-α test, denoted as φ, rejects $\eta \in M(\Gamma^{\Delta}(\phi)\Omega)$ if:

$$Q_f^{(n)}(\underline{\phi}) = \underline{\Delta}_f^{(n)'}(\underline{\phi})\left[\Gamma^{\Delta}(\underline{\phi})^{-1} - \Omega(\Omega'\Gamma^{\Delta}(\underline{\phi})\Omega)^{-1}\Omega'\right]\underline{\Delta}_f^{(n)}(\underline{\phi}) > \chi^2_{S(p-1),1-\alpha}.$$

The test statistic depends on the unknown parameter $\underline{\phi}$, but replacing it with a $\sqrt{n}$-consistent estimator does not affect the asymptotic behavior of the test statistic. Therefore, we introduce the following hypothesis.

Assumption (A3):

i. The estimator $\underline{\phi}^{(n)}$ is $\sqrt{n}$-consistent.
ii. $\underline{\phi}^{(n)}$ is locally asymptotically discrete.

Remark 3.1: The $\sqrt{n}$-consistency of the estimator $\underline{\phi}^{(n)}$ ensures that, as the sample size increases, the estimator converges in probability to the true parameter value. This property is crucial for maintaining the validity of statistical inference procedures, such as hypothesis testing, particularly in the asymptotic framework. Furthermore, the local asymptotic discreteness of $\underline{\phi}^{(n)}$ indicates that in the vicinity of the true parameter value, the estimator assumes distinct and well-defined values with increasing precision. This local behavior enhances the reliability of the estimator in capturing the underlying characteristics of the model.

The following proposition establishes the locally asymptotically optimal test.

Proposition 3.1 Suppose assumptions A1-A3 hold. Then

i) The test statistic

$$\hat{Q}_f^{(n)}(\underline{\hat{\phi}}^{(n)}) = \frac{S}{I(f_1)}\sum_{s=1}^{S}\left(\frac{\hat{\sigma}_s^2}{\sum_{i=1}^{p}\hat{\Gamma}_{s,i}}\right)\left[\sum_{j=1}^{p}\left(\frac{\sum_{i=1}^{p}\hat{\Gamma}_{s,i} - \hat{\Gamma}_{s,j}}{\hat{\Gamma}_{s,j}}\delta_{s,j}^2(\underline{\hat{\phi}}^{(n)}) - \sum_{\substack{i=1\\ i\neq j}}^{p}\delta_{s,i}(\underline{\hat{\phi}}^{(n)})\delta_{s,j}(\underline{\hat{\phi}}^{(n)})\right)\right]$$

is asymptotically $\chi^2_{S(p-1)}$ under $H_f^{(n)}(\underline{\phi})$.

ii) It has asymptotic power:

$$1 - \mathcal{F}(\chi^2_{1-\alpha}; S, \nu), \quad \text{under} \quad H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau}),$$

where

$$\nu = \frac{1}{S}I(f)\sum_{s=1}^{S}\left(\frac{1}{\sum_{i=1}^{p}\Gamma_{s,i}}\right)\sum_{j=1}^{p}\left[\Gamma_{s,j}\left(\sum_{i=1}^{p}\Gamma_{s,i} - \Gamma_{s,j}\right)h_{s,j}^2 - \sum_{\substack{i=1\\ i\neq j}}^{p}\Gamma_{s,i}\Gamma_{s,j}h_{s,i}h_{s,j}\right],$$

and $\mathcal{F}(\chi^2_{1-\alpha}; r, \nu)$ denotes the noncentral chi-square distribution function with r degrees of freedom and noncentrality parameter ν.

iii) It is the locally asymptotically most stringent test against $H_f^{(n)}(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau})$.

Proof. See appendix.

Remark 3.2: In the case where $p = 2$, the obtained result aligns with the findings presented by Merzougui (2016).

4. Adaptive test

In a nonparametric context, the innovation density f remains unspecified. An effective approach for estimating it involves employing the Kernel method. This section explores a semi-parametric model, acknowledging the unknown but symmetric nature of the density function. This symmetry assumption significantly simplifies the development of the adaptive test. Thus, we introduce:

Assumption (A4): The innovation density function f is symmetric and possesses a finite fourth moment.

Following the same construction and notation of Kreiss (1987), we define:

i) $g(x;\eta) = \frac{1}{\sqrt{2\pi\eta^2}} \exp\left(-\frac{x^2}{2\eta^2}\right), \quad x \in \mathbb{R},$
ii) $f_\eta(x) = \int_{-\infty}^{+\infty} g(x-y;\eta) f(y)\, dy,$
iii) $f_{\eta,\tau}(x,\underline{\phi}) = \frac{1}{2(m-1)} \sum_{\tau_0=0}^{m-1} \left[g(x+z_{s,\tau_0},\eta) + g(x-z_{s,\tau_0},\eta)\right], \quad \tau = 0,\dots,m-1.$

The estimator of $\varphi_f(.)$ is:

$$\hat{\varphi}_{n,\tau}(x,\underline{\phi}) = \begin{cases} -\frac{1}{2}\left(\frac{\hat{f}'_{\eta(n),\tau}(x,\underline{\phi})}{\hat{f}_{\eta(n),\tau}(x,\underline{\phi})}\right) & \text{if } \hat{f}_{\eta(n),\tau}(x,\underline{\phi}) \geq d_n, |x| \leq g_n, |\hat{f}'_{\eta(n),\tau}(x,\underline{\phi})| \leq c_n \hat{f}_{\eta(n),\underline{\tau}}(x,\underline{\phi}) \\ 0 & \text{otherwise} \end{cases}$$

where $c_n \to \infty, g_n \to \infty, \eta(n) \to 0, d_n \to 0.$
The estimator of the Fisher Information $I(f_1)$ is:

$$\hat{I}_n = \frac{1}{S}\sum_{s=1}^{S}\left(\frac{1}{m}\sum_{\tau=0}^{m-1}\hat{\varphi}^2_{n,\tau}\left(Z^{(n)}_{s,\tau}(\underline{\phi}_s), \underline{\phi}_s\right)\right).$$

Finally, the central sequence is estimated by:

$$\tilde{\delta}^{(n)}_s(\underline{\phi}_s) = \frac{1}{\sqrt{n}}\sum_{\tau=0}^{m-1}\hat{\varphi}_{n,\tau}(x,\underline{\phi}_s)X(s-1+d\tau), \quad s = 1,2,\dots,S.$$

The following proposition establishes the adaptive test for a classical PAR(1) model against a PSETAR($p,1,\dots,1$) model.

Proposition 4.1 Under the assumptions (A1)-(A4), we have, under $H^{(n)}_f(\underline{\phi})$, the test statistic

$$\tilde{Q}^{(n)}_f(\underline{\hat{\phi}}^{(n)}) = \frac{S}{\hat{I}(f_1)}\sum_{s=1}^{S}\left(\frac{\hat{\sigma}^2_s}{\sum_{i=1}^{p}\hat{\Gamma}_{s,i}}\right)\left[\sum_{j=1}^{p}\left(\frac{\sum_{i=1}^{p}\hat{\Gamma}_{s,i}-\hat{\Gamma}_{s,j}}{\hat{\Gamma}_{s,j}}\tilde{\delta}^2_{s,j}(\underline{\hat{\phi}}^{(n)}) - \sum_{\substack{i=1\\i\neq j}}^{p}\tilde{\delta}_{s,i}(\underline{\hat{\phi}}^{(n)})\tilde{\delta}_{s,j}(\underline{\hat{\phi}}^{(n)})\right)\right],$$

which follows a chi-square distribution with $S(p-1)$ degrees of freedom, denoted as $\chi^2_{S(p-1),1-\alpha}$.

i) It is a locally adaptive test of $H^{(n)}_f(\underline{\phi})$ versus $H^{(n)}_f(\underline{\Phi} + \underline{\nu}^{(n)}\underline{\tau})$.
ii) The asymptotic power of this test is the same as the one obtained in the parametric case.

Proof. See appendix.

5. Simulation results

To implement the adaptive test, the bandwidth parameter η is chosen as $\eta = n^{-(1/3)}$, where n denotes the sample size. The performance of the proposed test is then assessed through a series of simulation studies using two types of models:

1. **Model M.1**: A Linear PAR(1) model with a period $S = 4$, aimed at calculating the empirical levels of the adaptive test.
2. **Models M.2, M.3, and M.4**: Three Periodic PSETAR($p,1,\dots,1$) models, with periods $S = 2,4$. These models are used to simulate time series of varying sizes ($n = 100, 200, 300, 400$). The true parameter values of model **M.2** were chosen to assess the sensitivity of the test to weak nonlinearity, i.e., when the parameters within the same cycles are not highly distinct.

For each data-generating process, we perform 1000 Monte Carlo replications and report the frequencies where nonlinearity is correctly identified. The innovation processes are assumed to follow two probability densities:

1. **Standard normal distribution:**
$$f_1(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$$
2. **Kreiss density (1987):** A mixture of two normal distributions:
$$f_2(x) = \frac{0.5\sqrt{10}}{\sqrt{2\pi}} \exp\left(-\frac{\left(\sqrt{10}x - 3\right)^2}{2}\right) + \frac{0.5\sqrt{10}}{\sqrt{2\pi}} \exp\left(-\frac{\left(\sqrt{10}x + 3\right)^2}{2}\right)$$

Results for Model M.1

Model M.1: A Linear $\text{PAR}_4(1)$ model with parameters $\underline{\phi} = (0.4, 0.7, 0.2, 0.9)'$. The empirical levels of the adaptive test ($\alpha = 5\%$) are shown in Table 1. These results indicate that the empirical levels of the test are consistent and acceptable.

Table 1. Empirical levels for Model M.1

f/n	100	200	300	400
f_1	0.0370	0.0330	0.0270	0.0260
f_2	0.0420	0.0290	0.0310	0.0300

Results for Models M.2, M.3, and M.4

To evaluate the empirical power of the adaptive test, we consider three PSETAR(2,1,1) models:

- **Model M.2**: $\underline{\phi} = (0.35, 0.45; 0.75, 0.65)'$, $S = 2$
- **Model M.3**: $\underline{\phi} = (0.3, 0.5; 0.8, 0.6)'$, $S = 2$
- **Model M.4**: $\underline{\phi} = (0.6, 0.3; -0.6, -0.3; 0.4, -0.8; 0.7, 0.3)'$, $S = 4$

The empirical power results ($\alpha = 5\%$) are presented in Tables 2, 3, and 4.

Table 2. Empirical power for Model M.2

f/n	100	200	300	400
f_1	0.4790	0.4810	0.5020	0.5530
f_2	0.4930	0.5220	0.5820	0.6250

Table 3. Empirical power for Model M.3

f/n	100	200	300	400
f_1	0.5380	0.6810	0.7730	0.8310
f_2	0.8600	0.8750	0.9380	0.9620

The results in Table 1 demonstrate that the empirical levels of the adaptive test align well with the nominal level ($\alpha = 5\%$). Furthermore, the results in Tables 2–4 indicate that the power of the test increases with the sample size n. Notably, the test performs better under the Kreiss density (f_2), which is far from normality, compared to the standard normal density (f_1). These findings also reveal that the adaptive tests can detect nonlinearity with substantial power, even in scenarios where the model is nearly linear (e.g., Model M.2).

Table 4. Empirical power for Model M.4

f/n	100	200	300	400
f_1	0.5620	0.9640	0.9970	0.9990
f_2	0.9270	0.9800	0.9970	0.9990

6. Conclusion

In this paper, we have developed an adaptive test for detecting nonlinearity in periodic time series models, leveraging the Local Asymptotic Normality property to establish a locally most stringent test for linearity. The results from our simulation studies demonstrate that the proposed test maintains a well-calibrated empirical level ($\alpha = 0.05$) and exhibits increased power as the sample size grows. The adaptive test shows enhanced performance in the presence of non-normal densities, and is capable of detecting nonlinearity even in nearly linear models. While the adaptive test shows promising results in simulation settings, future work could explore its application to real-world data to demonstrate its practical utility. This would provide valuable insights into the test's effectiveness in realistic scenarios and further highlight its potential for periodic time series datasets.

Acknowledgement

We sincerely thank the referees for their valuable comments and suggestions, which have greatly contributed to clarifying and improving the quality of this work.

REFERENCES

1. A. Aknouche, B. Almohaimeed, and S. Dimitrakopoulos, *Periodic autoregressive conditional duration.* Journal of Time Series Analysis, 43, 5-29, 2022.
2. A. Aknouche, and H. Guerbyenne, *On some probabilistic properties of double periodic AR models.* Statistics and Probability Letters, 79, 407-413, 2009.
3. A. Amimour, *Adaptive Test for Periodic ARFIMA Models.* Thailand Statistician. 21(4): 802-811, 2023.
4. N. Azouagh, and S. El Melhaoui, *Detection of EXPAR nonlinearity in the Presence of a Nuisance Unidentified Under the Null Hypothesis.* Sankhya B: The Indian Journal of Statistics, Springer; Indian Statistical Institute, vol. 83(2), pages 397-429, 2021.
5. M. Bentarzi, and L. Djeddou, *Adaptive Estimation of Periodic First-Order Threshold Autoregressive Model.* Communications in Statistics - Simulation and Computation, 43(7), 1611-1630, 2014.
6. A.Bibi, and A. Gautier, *Stationarity and asymptotic inference of some periodic bilinear models. Comptes Rendus de l'Académie des Sciences - Series I - Mathematics, 341(9), 679-682.(2).* Comptes Rendus de l'Académie des Sciences - Series I - Mathematics, 341(9), 679-682.(2), 2005.
7. T. Bollerslev, and E. Ghysels, *Periodic Autoregressive Conditional Heteroskedasticity.* Journal of Business and Economic Statistics, 14(2), 139-152, 1996.
8. E. G. Gladyshev, *Periodically correlated random sequences.* Soviet Mathematics, 2, 385-388, 1961.
9. J. Hájek, and Z. Šidák, *Theory of Rank Tests.* Academic Press, New York.
10. J. P. Kreiss, *On Adaptive Estimation in Stationary ARMA Processes.* The Annals of Statistics, 15(1), 112-133, 1987.
11. L. LeCam, *Asymptotic Methods in Statistical Decision Theory.* Springer-Verlag, New York, 1986.
12. L. LeCam, *Locally Asymptotically Normal Families of Distributions.* University of California Publications in Statistics, 3, 37-98, 1960.
13. P. A. W. Lewis, and B. K. Ray, *Nonlinear modelling of periodic threshold autoregressions using TSMARS.* Journal of Time Series Analysis, 23(4), 459-471, 2002.
14. M. Merzougui, *Optimal Test for PAR(1) dependence against PSETAR(2,1,1) Models with specified Threshold.* Communications in Statistics - Theory and Methods, 45(4), 872-886, 2016.
15. M. Sadoun, and M. Bentarzi, *Efficient estimation in periodic INAR(1) model: parametric case.* Communication in Statistics-Simulation and Computation 49(8):1-21, 2019.
16. Q. Shao, *Mixture periodic autoregressive time series model.* , Statistics and Probability Letters, 76(6), 609-618, 2006.
17. A. R. Swensen, *(1985). The Asymptotic Distribution of the likelihood ratio for autoregressive time series with a regression trend.* Journal of Multivariate Analysis, 16, 54-70, 1985.
18. H. Tong, *On threshold models. In Pattern Recognition and Signal Processing.* (Ed. C. H. Chen). Amsterdam: Sijhoff and Noordhoff, 1978.

A. Appendix

A.1. Proof of Proposition 2.1

The proof of this LAN result is established based on Swensen's Conditions (1985). It is a straightforward extension of the case p=2 presented in [14]. The extension ensures that the LAN property holds for the periodic threshold autoregressive model with a higher-dimensional parameter space.

A.2. Proof of Proposition 3.1

Considering the decomposition of the central sequence,

$$\underline{\Delta}_f^{(n)}(\underline{\phi}) = \begin{bmatrix} \underline{\Delta}_{I,f}^{(n)} \\ \underline{\Delta}_{II,f}^{(n)} \end{bmatrix},$$

and the corresponding decomposition of $\Gamma^{\Delta}(\underline{\phi})$, we can express the quadratic form (3) as

$$Q_f^{(n)}(\underline{\phi}) = \frac{S}{I(f_1)} \sum_{s=1}^{S} \left(\frac{\sigma_s^2}{\sum_{i=1}^{p} \Gamma_{s,i}} \right) \left[\sum_{j=1}^{p} \left(\frac{\sum_{i=1}^{p} \Gamma_{s,i} - \Gamma_{s,j}}{\Gamma_{s,j}} \delta_{s,j}^2(\underline{\phi}) - \sum_{\substack{i=1 \\ i \neq j}}^{p} \delta_{s,i}(\underline{\phi}) \delta_{s,j}(\underline{\phi}) \right) \right].$$

Certainly, these results are derived under the assumption that the parameters $\underline{\phi}$ are known. However, we demonstrate that these findings still hold when these parameters are unknown. Specifically, let

$$\underline{\tau}^{(n)} = (\underline{\nu}^{(n)})^{-1}(\hat{\underline{\phi}}^{(n)} - \underline{\Phi})$$

for any $\underline{\nu}^{(n)}$-convergent estimator $\hat{\underline{\phi}}^{(n)}$ of the unknown parameter $\underline{\Phi}$. By substituting this expression into the asymptotic linearity expression given by ii) of Proposition 3.3, we obtain

$$\underline{\Delta}_f^{(n)}(\hat{\underline{\phi}}^{(n)}) - \underline{\Delta}_f^{(n)}(\underline{\phi}) = -\Gamma^{\Delta}(\underline{\phi}, \underline{\sigma})(\underline{\nu}^{(n)})^{-1}(\hat{\underline{\phi}}^{(n)} - \underline{\Phi}) + o_p(1),$$

hence we have

$$(\Gamma^{\Delta}(\underline{\phi}, \underline{\sigma}))^{-1/2}(\underline{\Delta}_f^{(n)}(\hat{\underline{\phi}}^{(n)}) - \underline{\Delta}_f^{(n)}(\underline{\phi})) = (\Gamma^{\Delta}(\underline{\phi}, \underline{\sigma}))^{1/2}(\underline{\nu}^{(n)})^{-1}(\hat{\underline{\phi}}^{(n)} - \underline{\Phi}) + o_p(1).$$

Multiplying the left side by the matrix

$$I - (\Gamma^{\Delta}(\underline{\phi}))^{1/2}\Omega(\Omega'\Gamma^{\Delta}(\underline{\phi})\Omega)^{-1}\Omega'(\Gamma^{\Delta}(\underline{\phi}))^{1/2}$$

and considering the continuity of $\Gamma^{\Delta}(\underline{\phi}, \underline{\sigma})$, we obtain, under $H_f^{(n)}(\underline{\phi})$ and hence under $H_f^{(n)}(\underline{\phi}^{(n)})$,

$$\left[I - (\Gamma^{\Delta}(\hat{\underline{\phi}}^{(n)}))^{1/2}\Omega(\Omega'\Gamma^{\Delta}(\hat{\underline{\phi}}^{(n)})\Omega)^{-1}\Omega'(\Gamma^{\Delta}(\hat{\underline{\phi}}^{(n)}))^{1/2} \right] (\Gamma^{\Delta}(\hat{\underline{\phi}}^{(n)}))^{-1/2}(\underline{\Delta}_f^{(n)}(\hat{\underline{\phi}}^{(n)}) - \underline{\Delta}_f^{(n)}(\underline{\phi})) = o_p(1).$$

This leads to

$$Q_f^{(n)}(\hat{\underline{\phi}}^{(n)}) = Q_f^{(n)}(\underline{\phi}) + o_p(1),$$

under $H_f^{(n)}(\underline{\phi})$ and hence under $H_f^{(n)}(\underline{\phi}^{(n)})$. Therefore, the test statistic $Q_f^{(n)}(\hat{\underline{\phi}}^{(n)})$ follows the central $\chi^2(S)$ and the noncentral $\chi^2(S;\nu)$ under $H_f^{(n)}(\underline{\phi})$ and $H_f^{(n)}(\underline{\phi}^{(n)})$, respectively, where

$$\nu = \underline{\tau}'\Gamma^{\Delta}(\underline{\phi})' \left[\Gamma^{\Delta}(\underline{\phi})^{-1} - \Omega(\Omega'\Gamma^{\Delta}(\underline{\phi})\Omega)^{-1}\Omega' \right] \Gamma^{\Delta}(\underline{\phi})\underline{\tau},$$

which gives

$$\nu = \frac{1}{S} I(f) \sum_{s=1}^{S} \left(\frac{1}{\sum_{i=1}^{p} \Gamma_{s,i}} \right) \sum_{j=1}^{p} \left(\Gamma_{s,j} \left(\sum_{i=1}^{p} \Gamma_{s,i} - \Gamma_{s,j} \right) h_{s,j}^2 - \sum_{\substack{i=1 \\ i \neq j}}^{p} \Gamma_{s,i} \Gamma_{s,j} h_{s,i} h_{s,j} \right).$$

Consequently, the proof of ii) follows immediately.

A.3. Proof of Proposition 4.1

The equivalence of the statistics in Propositions 3.1 and 4.1, in probability, is demonstrated through the same approach used by Kreiss (1987). Specifically, we establish that

$$\hat{I}_n = I(f_1) + o_p(1)$$

and

$$\tilde{\underline{\Delta}}_f^{(n)}(\hat{\underline{\phi}}^{(n)}) - \underline{\Delta}_f^{(n)}(\hat{\underline{\phi}}^{(n)}) = o_p(1).$$

B. Threshold Search on Monthly Time Series Data

This appendix has been included at the request of the referee to provide further details on the threshold estimation method in application. The code below illustrates the grid search procedure used to determine the optimal threshold for each month of a generic monthly time series dataset.

```
# Load the data
# Replace "temp.txt" with the path to your data file
# Example: General monthly time series
yy1 <- read.table("temp.txt")
yy <- ts(yy1, start = c(1901, 1), frequency = 12) # Define the series as monthly
yy <- log(yy) # Apply a logarithmic transformation
yy <- matrix(yy, nrow = 12) # Reshape into a 12-row matrix for each month

# Initialize parameters
d <- 12 # Number of months
m <- ncol(yy) # Number of years

# Function to compute RSS for a given threshold
compute_rss_month <- function(threshold, data) {
  n <- length(data)
  Y <- data[2:n] # Dependent variable
  X_lag <- data[1:(n - 1)] # Lagged independent variable

  # Split the data based on the threshold
  below_threshold <- which(X_lag <= threshold)
  above_threshold <- which(X_lag > threshold)

  if (length(below_threshold) > 1 && length(above_threshold) > 1) {
    # Fit separate linear models
    model_below <- lm(Y[below_threshold] ~ X_lag[below_threshold])
    model_above <- lm(Y[above_threshold] ~ X_lag[above_threshold])
```

```
    # Calculate RSS
    rss_below <- sum(residuals(model_below)^2)
    rss_above <- sum(residuals(model_above)^2)

    return(rss_below + rss_above)
  } else {
    return(NA) # Not enough data in one group
  }
}

# Apply grid search for threshold for each month
results <- lapply(1:d, function(month) {
  data <- yy[month, ] # Extract data for the current month

  # Define the threshold grid (10th to 90th percentiles)
  threshold_grid <- seq(quantile(data, 0.1), quantile(data, 0.9), length.out = 100)

  # Calculate RSS for each threshold
  rss_values <- sapply(threshold_grid, compute_rss_month, data = data)

  # Find the optimal threshold
  optimal_threshold <- threshold_grid[which.min(rss_values)]

  list(
    month = month,
    optimal_threshold = optimal_threshold,
    rss_values = rss_values,
    threshold_grid = threshold_grid
  )
})

# Plot all graphs for each month
par(mfrow = c(3, 4)) # 3 rows and 4 columns for subplots
for (res in results) {
  plot(res$threshold_grid, res$rss_values, type = "l",
       main = paste("RSS (Month", res$month, ")"),
       xlab = "Threshold", ylab = "RSS")
  abline(v = res$optimal_threshold, col = "red", lty = 2) # Mark the optimal threshold
}
```