Model Order Selection in Seasonal/Cyclical Long Memory Models*

Christian Leschinski and Philipp Sibbertsen‡

Institute of Statistics, Faculty of Economics and Management,

Leibniz University Hannover, D-30167 Hannover, Germany

Abstract

We propose an automatic model order selection procedure for k-factor GARMA processes. The procedure is based on sequential tests of the maximum of the periodogram and semiparametric estimators of the model parameters. As a byproduct, we introduce a generalized version of Walker's large sample g-test that allows to test for persistent periodicity in stationary ARMA processes. Our simulation studies show that the procedure performs well in identifying the correct model order under various circumstances. An application to Californian electricity load data illustrates its value in empirical analyses and allows new insights into the periodicity of this process that has been subject of several forecasting exercises.

JEL-Numbers: C22, C52

Keywords: Seasonal Long Memory · k-factor GARMA processes · Model selection · Electricity loads

---

*We are grateful to Liudas Giraitis and Uwe Hassler for their remarks on earlier versions of this paper and the participants of the NSVCM 2014 Workshop in Paderborn. The financial support of DFG is gratefully acknowledged.

‡Corresponding author:
Phone: +49-511-762-3783
Fax: +49-511-762-3923
E-Mail: sibbertsen@statistik.uni-hannover.de

1 Introduction

The increasing availability of high frequency data poses new challenges to the analysis of cyclical time series, because intraday data often exhibits periodic behavior and potentially contains multiple cycles such as daily and weekly seasonalities. Important examples of those series include intraday volatilities as discussed by Andersen and Bollerslev (1997), Bisaglia et al. (2003), Bordignon et al. (2008) and Rossi and Fantazzini (2014) as well as trading volumes and electricity data, where the aforementioned features are especially pronounced. Recent contributions such as Haldrup and Nielsen (2006), Soares and Souza (2006) and Diongue et al. (2009) stress the long memory properties of electricity time series and suggest that Gegenbauer models are useful to analyze these datasets, because they allow for different degrees of long memory at arbitrary periodic frequencies. An unresolved issue however, is how to select the number of cyclical components that have to be modeled. This is why we propose a model selection procedure that consistently estimates the required model order and demonstrate how it can be applied to the analysis of electricity load data.

Intuitively speaking, cyclical long memory is an intermediate case between short memory seasonal ARMA processes and the seasonally integrated model. While a time series exhibits long memory if it has a hyperbolically decaying autocorrelation function - the autocorrelation functions of cyclical long memory processes show sinusoidal patterns with hyperbolically decaying amplitude, so that the dependence between observations at distant periodic lags is relatively strong.

We will hereafter refer to the model class considered as $k-$factor GARMA models. These were proposed by Gray et al. (1989) and generalized by Giraitis and Leipus (1995) and Woodward et al. (1998). The k-factor GARMA model is given by

$$\phi(L)\prod_{j=1}^{k}(1-2\cos\gamma_j L+L^2)^{d_j}X_t=\theta(L)\varepsilon_t, \tag{1}$$

where $\phi(L)$ and $\theta(L)$ are the usual AR- and MA-polynomials, the lag-operator L is defined through $LX_t = X_{t-1}$ and ε_t is a white noise sequence. The filter $(1-2uL+L^2)^{-d}$ is the generating function of the orthogonal Gegenbauer polynomials denoted by $C_s^d(u)$:

$$(1-2uL+L^2)^{-d}=\sum_{s=0}^{\infty}C_s^d(u)L^s,$$

where the Gegenbauer polynomials are given by

$$C_s^d(u) = \sum_{g=0}^{\lfloor s/2 \rfloor} \frac{(-1)^g (2u)^{s-2g} \Gamma(d-g+s)}{g!(n-2g)!\Gamma(d)}.$$

In this representation the operator $\lfloor . \rfloor$ returns the integer part of its argument and $\Gamma(x)$ denotes the gamma function. The process defined by applying this filter to a white noise sequence v_t is a general linear process with the MA(∞)-representation shown below:

$$\begin{aligned} Y_t &= (1-2\cos\gamma L + L^2)^{-d} v_t \\ &= \sum_{s=0}^{\infty} \Psi_s v_{t-s}, \end{aligned} \tag{2}$$

where the coefficients Ψ_s in this representation are the Gegenbauer polynomials $C_s^d(\cos\gamma)$. The process is causal, invertible and has long memory, if

$$|d_j| < \begin{cases} 1/2, & \forall\ 0 < \gamma_j < \pi \\ 1/4, & \forall\ \gamma_j = 0 \text{ or } \gamma_j = \pi, \end{cases}$$

given that all roots of $\phi(L) = 0$ and $\theta(L) = 0$ lie outside of the unit circle and $\phi(L)$ and $\theta(L)$ have no roots in common (cf. Giraitis and Leipus (1995)). The spectral density of (1) is given by

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \frac{|\theta(e^{\mathbf{i}\lambda})|^2}{|\phi(e^{\mathbf{i}\lambda})|^2} \prod_{j=1}^{k} \left| 2\left(\cos\lambda - \cos\gamma_j\right)\right|^{-2d_j}. \tag{3}$$

As can be seen from (3) the spectral density has poles due to the long memory behavior at the cyclical frequencies γ_j with $j = 1, ..., k$.

Note that the k-factor GARMA model and the seasonal/cyclical long memory (SCLM) model of Robinson (1994) are equivalent given appropriate parameter restrictions. They generalize the ARFIMA class by allowing for poles in the spectral density at arbitrary frequencies γ_j and they nest most of the seasonal long memory models proposed in the literature such as the (rigid) SARFIMA model of Porter-Hudak (1990) or the flexible SARFIMA of Hassler (1994).

One method to choose the model order k and the locations γ_j of the poles is based on the LM tests of Robinson (1994) and Hassler et al. (2009) who test whether the sample supports a given specification of (1). The null hypotheses are of the form $H_0:$ $\theta = \theta_0$ versus $H_1: \ \theta \neq \theta_0$, where $\theta = \{\gamma_1, d_1, ..., \gamma_k, d_k\}'$. These procedures are useful in two situations. First, if the theory suggests a model for the process considered, or the researcher wants to test whether one of the nested special cases such as a (rigid) SARFIMA fits the data. Second, if the tests are applied on a grid of values for θ to obtain "model confidence sets" that contain the true model with a certain coverage probability as suggested by Hassler et al. (2009). This second application of the LM-tests allows to specify the location of periodic frequencies as well as their number, because θ implicitly contains the model order k. The model confidence set should thus contain models of the true model order k_0. This grid search procedure has become the most common specification method. Examples of its application include Gil-Alana (2002) and Gil-Alana (2007).

For larger k however, this model selection procedure suffers from severe dimensionality problems. Consider a sample of T observations. Assume that the grid for γ_j has as many points as there are periodogram ordinates and let n_d denote the number of values on the grid for the respective d_j. Then the number of points on the combined grid for a k-factor model is

$$n_d^k \prod_{j=1}^{k} \{\lfloor T/2 \rfloor - (j-1)\},$$

which is $O\Big(\{n_d \lfloor T/2 \rfloor\}^k\Big)$, so that the procedure quickly becomes unapplicable for models with a larger number of relevant cyclical frequencies. In these situations the only feasible procedure is to choose the model order k discretionary based on a visual inspection of the periodogram. Unfortunately, this often leads to misspecifications as demonstrated in our empirical application.

To overcome this problem we suggest a model order selection procedure based on iterative filtering and periodogram based tests for cyclical behavior in a time series. It is based on the observation that the residual series from a correctly specified model for X_t given in (1) with $k = k_0$ Gegenbauer filters has no poles in the periodogram whereas it still exhibits poles if the selected model order $k < k_0$ is too low. Consequently, we can apply k-factor Gegenbauer filters of increasing order k, until no significant periodicity can be detected anymore.

Section 2 discusses our model order selection procedure in more detail for the infeasible case in which the γ_j and d_j are known. To make the procedure feasible, we need a test that can be used after each filtering step to determine whether the residual process still contains significant periodicity. Such a test for periodicity of unknown frequency is suggested in Section 3. We also need estimators for γ_j and d_j that will be discussed in Section 4. Subsequently, we provide a Monte Carlo analysis of the finite sample properties of our model order selection procedure and apply it to a Californian electricity load series.

2 Infeasible Automatic Model Selection by Sequential Filtering

Consider the k-factor GARMA process in equation (1) and denote the stationary ARMA component by $u_t = \phi(L)^{-1}\theta(L)\varepsilon_t$, where $\varepsilon_t \overset{iid}{\sim} (0, \sigma_\varepsilon^2)$ is a white noise sequence with finite 8th order moments.[1] The fractional exponents are restricted to $0 < d_j < 1/2$ for γ_j in the open interval $(0, \pi)$ and $0 < d_j < 1/4$ for γ_j in the set $\{0, \pi\}$ so that the process exhibits stationary long memory.

Assume for now, that the true cyclical frequencies $\gamma_{j,0}$ and the true fractional exponents $d_{j,0}$ at those frequencies were known for all $j = 1, ..., k_0$. In addition to that, assume that for every finite sample from a weakly dependent linear process $Z_t = \sum_{j=0}^{\infty} a_j z_{t-j}$ with $z_t \overset{iid}{\sim} (0, \sigma_z^2)$ and $\sum_{j=0}^{\infty} |a_j| < \infty$ we are able to determine the probability $P(Z_t = u_t | Z_1, ..., Z_T)$ that Z_t contains no Gegenbauer component and is a pure short memory ARMA process. The main idea of our model order selection procedure is based on equation (5) below. Let $k^{(i)} \leq k_0$ denote a positive integer. If the filter

$$\Delta^{k^{(i)}} = \prod_{j=1}^{k^{(i)}} (1 - 2\cos\gamma_{j,0} L + L^2)^{d_{j,0}} \tag{4}$$

is applied to the series X_t, we obtain the residual process $\Delta^{k^{(i)}} X_t$ given by

$$\Delta^{k^{(i)}} X_t = \prod_{j=k^{(i)}+1}^{k} (1 - 2\cos\gamma_{j,0} L + L^2)^{-d_{j,0}} u_t. \tag{5}$$

If X_t was indeed generated by a $k^{(i)}$-factor GARMA process, then the right hand side of (5) reduces to the ARMA sequence u_t. If $k_0 > k^{(i)}$ on the other hand, the process at the right hand side is a $(k_0 - k^{(i)})$-factor Gegenbauer process, with $(k_0 - k^{(i)})$ poles in its spectral density. We can thus use a sequence of periodogram based tests for a significant periodicity in $\Delta^{k^{(i)}} X_t$ to determine the model order of the Gegenbauer process.

Starting with $k^{(1)} = 0$ we test the null hypothesis

$$H_0: \; k_0 = k^{(i)} \text{ vs. } H_1: \; k_0 > k^{(i)} \tag{6}$$

repeatedly for $k^{(i)} = k^{(i-1)} + 1$ until the null hypothesis cannot be rejected anymore. That means, we test whether a model of order $k^{(i)}$ is adequate to describe the data or whether

[1] Note that this assumption is only required for the consistency of the semiparametric estimator of Hidalgo and Soulier (2004). All other methods employed do not require any additional moment conditions.

there are additional periodicities present that the $k^{(i)}$-factor model does not account for. The first $k^{(i)}$ for which the null of no significant periodicity cannot be rejected anymore is than selected as the model order:

$$\hat{k} = \min \left\{ k^{(i)} \text{ such that } P\left(\Delta^{k^{(i)}} X_t = u_t \mid X_1, ..., X_T\right) > \alpha \right\}, \tag{7}$$

where α is the desired significance level.

So far we have assumed to know $\gamma_{j,0}$, $d_{j,0}$ and $P(Z_t = u_t \mid Z_1, ..., Z_T)$. To make this procedure feasible we need estimators for $\gamma_{j,0}$ and $d_{j,0}$ that are consistent under the null hypothesis as well as under the alternative - otherwise the procedure could not be applied sequentially. The latter can be achieved using semiparametric estimators for the fractional exponent d_j as in Arteche and Robinson (2000) and semiparametric estimators of the location of the pole as in Yajima (1996) and Hidalgo and Soulier (2004). A test that allows to determine $P(Z_t = u_t \mid Z_1, ..., Z_T)$ is presented in the next section.

3 Testing for Periodicity of Unknown Frequency

Note that under the null hypothesis in equation (6) the filtered process $\Delta^{k^{(i)}} X_t$ is a short memory process. Define the periodogram of the weakly dependent linear process $Z_t = \sum_{j=0}^{\infty} a_j z_{t-j}$ with $z_t \overset{iid}{\sim} (0, \sigma_z^2)$ as

$$I(\lambda) = (2\pi T)^{-1} \left| \sum_{t=1}^{T} Z_t e^{-\mathbf{i}\lambda t} \right|^2, \quad \text{with} \quad \lambda \in [-\pi, \pi].$$

Here [.] denotes the closed interval. Further denote by I_κ the periodogram evaluated at the Fourier frequency $\lambda_\kappa = \frac{2\pi\kappa}{T}$ with $\kappa = 1, ..., n$ and $n = \lfloor (T-1)/2 \rfloor$.

Tests for periodicity at an unknown frequency are based on the well-known result that the periodogram ordinates of weakly dependent linear processes are approximately equal to $f(\lambda_\kappa)/2$ times an exponentially distributed random variable:

$$\frac{2I_\kappa}{f(\lambda_\kappa)} \overset{appr.}{\sim} \chi_2^2. \tag{8}$$

A detailed discussion of traditional periodogram based tests for periodicity can be found in Priestley (1981). We will focus on Walker's large sample g-test for the null hypothesis of a Gaussian *iid*-sequence ($H_0: Z_t = z_t$). In this case the relationship in (8) holds without an approximation error and we have $f(\lambda) = \sigma_z^2/(2\pi)$.

Walker's large sample g-statistic is based on the maximum of the periodogram and it is defined as

$$g_Z^* = \frac{4\pi \max(I_\kappa)}{\hat{\sigma}_z^2}, \tag{9}$$

where $\hat{\sigma}_z^2$ is a consistent estimator of σ_z^2. Due to the consistency of $\hat{\sigma}_z^2$, the distribution of g_Z^* is asymptotically the same as if σ_z^2 was known. Since the periodogram ordinates I_κ are independent for *iid*-sequences, the distribution of g_Z^* is given by

$$p(g_Z^* > \tilde{z}) = 1 - (1 - \exp(-\tilde{z}/2))^n.$$

Walker's g-test can be extended to allow for stationary ARMA processes under the null hypothesis, if $\hat{\sigma}_z^2$ in (9) is replaced by a consistent estimate of $f(\lambda)$. Then the modified G^*-statistic is defined as

$$G_Z^* = \max\left\{\frac{2I_\kappa}{\hat{f}(\lambda_\kappa)}\right\}. \tag{10}$$

Since the periodogram ordinates of the weakly dependent linear process Z_t divided by its spectral density $I_\kappa/f(\lambda_\kappa)$ are asymptotically independent at the Fourier frequencies, the modified test has asymptotically the same limiting distribution as Walker's original g-test.

Usually the spectral density is estimated through smoothed versions of the periodogram and in general we could use any consistent estimator from this class. For our purpose however, it is important that the estimate $\hat{f}(\lambda_\kappa)$ is very smooth in small samples and a single large spike in the periodogram has little impact on the estimated spectrum in its immediate neighborhood. Otherwise the G^*-test would have bad size and power properties. This is why we use a logspline spectral density estimate as suggested by Cogburn et al. (1974), who showed that this estimator can be interpreted as a kernel density estimate, too.

Let $P_{H_T} := \{\underline{a} = \omega_0 < \omega_1 < ... < \omega_{H_T} = \overline{b}\}$ denote a partition of $\left[\underline{a}, \overline{b}\right]$, with distinct knots ω_h and denote the number of segments by H_T. Then a function $S_\Delta : \left[\underline{a}, \overline{b}\right] \to \mathbb{R}$ is said to be a spline of degree ν if

$$S_{\Delta}(x,\nu)=\begin{cases}P_1(x), & \omega_0\leq x<\omega_1\\ P_2(x), & \omega_1\leq x<\omega_2\\ \vdots & \\ P_{H_T}(x), & \omega_{H_T-1}\leq x<\omega_{H_T},\end{cases}$$

with $P_h(x)=b_0+b_1(x-\omega_h)+b_2(x-\omega_h)^2+\ldots+b_\nu(x-\omega_h)^\nu$ and $P_h^{(j)}(\omega_{h+1})=P_{h+1}^{(j)}(\omega_h)$, where $P_h^{(j)}(x)$ denotes the j-th derivative of $P_h(x)$ and $j=1,...,\nu-1$. That is $S_\Delta(x,\nu)$ is a piecewise polynomial of degree ν with support on $[a,b]$ and $(\nu-1)$ continuous derivatives at the knots ω_h. If splines are used for spectral density estimation as suggested by Cogburn et al. (1974), the interval of interest is $\left[\underline{a},\overline{b}\right]=[0,\pi]$ and a periodic spline is obtained under the additional condition $P_1^{(1)}(0)=P_{H_T}^{(1)}(\pi)=0$.

The basis for the application of splines in spectral density estimation is the result in (8). For weakly dependent linear processes Z_t, the periodogram ordinates are asymptotically uncorrelated and $I_\kappa=f(\lambda_i)Q_\kappa$, where Q_κ is exponential distributed with mean one. For the logarithm of I_κ follows $\log I_\kappa=\varphi+q_\kappa$, where q_κ is the log of the exponential variable Q_κ and φ is the log-spectral density. This linearization allows to apply a smoothing spline function to estimate φ. The spectral density estimate $\hat{f}(\lambda)$ is then obtained after reversing the log-transformation. Kooperberg et al. (1995) derive the maximum likelihood estimator based on the basis spline representation and show that it consistently estimates the log-spectral density φ.
By using this method to estimate $f(\lambda)$, the modified G^*-test in equation (10) becomes feasible. Details of the estimation procedure can be found in Kooperberg et al. (1995). It is common in this literature to use cubic splines, where the degree of the local polynomials is $\nu=3$. This is also the case that we consider in our simulation studies and empirical application. The number of segments in the partition is determined according to $H_T=\lfloor 1+T^\zeta\rfloor$, with $0<\zeta<1$.
With these modifications Walker's g-statistic becomes applicable to test for periodicity in the filtered series $\Delta^{k^{(i)}}X_t$ from (5). To obtain consistency of the test against cyclical long memory effects we require the following assumption that guarantees the identifiability of the poles and is the same as in Hidalgo and Soulier (2004):

Assumption 1. *The fractional exponents d_j in (1) are bounded away from zero: $d_j>c_j>0$, where c_j is a small constant $\forall\ j=1,...,k$.*

For the modified G^*-test we then obtain the following result:

Proposition 1. *Let $\hat{f}(\lambda)$ be a consistent smoothed periodogram estimate of the spectral density. Then for X_t characterized by (1):*

1. $p(G_X^* > \tilde{z}) = 1 - (1 - \exp(-\tilde{z}/2))^n$ *for* $k = 0$.
2. $\lim_{T\to\infty} P\left(G_X^* = \infty\right) = 1$, *if* $k > 0$ *and Assumption 1 holds.*

Proofs of the main results are given in the appendix. Since we have $k_0 - k^{(i)} = 0$ under the null hypothesis and $k_0 - k^{(i)} > 0$ in each iteration step before, we can use this test to determine whether all significant periodicity has been removed after the i-th iteration step.

Remark: The G^*-test allows to test for very general forms of persistent periodicity in linear time series and is not restricted to the cyclical long memory case discussed here. To our knowledge this is a new contribution to the literature on testing for cyclical behavior at unknown frequencies, since traditional methods such as Walker's g-test or the exact test of Fisher (1929) assume *iid*-sequences under the null hypothesis.

4 Local Semiparametric Estimators of Cyclical Frequencies and Memory Parameters

With the modified G^*-test at hand, we now turn to the estimation of the cyclical frequency γ_j and the memory parameter d_j after a significant periodic effect is detected. To ensure the consistency of these estimators the following assumptions are required in addition to Assumption 1:

Assumption 2. *The cyclical frequencies γ_j are bounded away from each other: $\gamma_{j-1} + \epsilon_j < \gamma_j$, where $\epsilon_j > 0$ is a small constant for all $j = 2, ..., k$.*

Assumption 3. *For the bandwidth parameter m and some trimming parameter l we have $\frac{m}{T} + \frac{l}{m}\log(m) + \frac{\log(m^3)}{l} \to 0$, as $T \to \infty$.*

Here m and l are the bandwidth and trimming parameters used by the generalized local Whittle estimator that will be discussed below. Assumption 3 is a modification of Assumption B.4 in Arteche and Robinson (2000) for the case of symmetric poles. Assumption 2 is required to ensure that the effects from two neighboring poles can be distinguished and that the spectral density has the typical shape of a long-memory process in the neighborhood of every cyclical freqency γ_j. This is required in addition to Assumption 1 for an estimation of the location of the cyclical frequencies with the method of Hidalgo and Soulier (2004). The local Whittle estimator of Arteche and

Robinson (2000) is consistent for the k-factor GARMA process (1) under Assumptions 1 and 3.

If a significant periodicity is detected, we have to estimate the frequency at which it occurs. The maximum of the periodogram was shown to be a consistent semiparametric estimator for the cyclical frequencies $\gamma_{j,0}$ in (1) by Yajima (1996) and Hidalgo and Soulier (2004).

Due to the sequential nature of our model selection procedure, we only estimate the location of one of the remaining $(k^{(i)} - k_0)$ poles in every iteration of the procedure. To estimate $\gamma_{j,0}$ we use

$$\hat{\gamma}_j = \arg\max_{\lambda_\kappa} I_{k^{(i)}}(\lambda_\kappa), \tag{11}$$

where $I_{k^{(i)}}(\lambda_\kappa)$ denotes the periodogram of the residual process $\Delta^{k^{(i)}} X_t$. Hidalgo and Soulier (2004) prove the consistency of the maximum of the periodogram as an estimator of the cyclical frequencies in (1) under Assumptions 1 and 2.

To estimate the fractional exponents $d_{j,0}$ we use a generalized local Whittle approach similar to that suggested by Arteche and Robinson (2000). Since they consider an asymmetric model that allows for different fractional exponents at each side of the pole, they estimate the exponents separately with a generalized local Whittle estimator using m frequencies on the respective side of the pole. The poles are assumed to be known, but the results of Hidalgo and Soulier (2004) show that the estimation of the cyclical frequencies leaves the semiparametric estimators asymptotically unaffected. The estimators are defined as $\tilde{d}_{j,a} = \arg\min R_a(d_{j,a})$ for $a = 1, 2$, where $R_a(d_{j,a}) = \log \tilde{C}_a(d_{j,a}) - \frac{2d_{j,a}}{m_a - l} \sum_{\kappa=l+1}^{m_a} \log \lambda_\kappa$, $\tilde{C}_1(d_{j,1}) = \frac{1}{m_1 - l} \sum_{\kappa=l+1}^{m_1} \lambda_\kappa^{2d_{j,1}} I(\gamma_j + \lambda_\kappa)$, $\tilde{C}_2(d_{j,2}) = \frac{1}{m_2 - l} \sum_{\kappa=l+1}^{m_2} \lambda_\kappa^{2d_{j,2}} I(\gamma_j - \lambda_\kappa)$ and l is a trimming parameter.

For γ_j close to 0 or π, there can be less than m Fourier frequencies on the respective side of γ_j that is close to the boundary. Therefore, we conduct the estimations using $m_a \le m$ periodogram ordinates on the respective side of γ_j. The pooled estimator $\hat{d}_j$ is a weighted average of the $\tilde{d}_{j,a}$:

$$\hat{d}_j = \frac{m_1}{m_1 + m_2} \tilde{d}_{j,1} + \frac{m_2}{m_1 + m_2} \tilde{d}_{j,2}.$$

Note that the estimator of Arteche and Robinson (2000) is restricted to $\gamma_j \in (0, \pi)$. As Hidalgo and Soulier (2004) point out, the power law of the spectral density is $|\lambda - \gamma_j|^{-\beta d_j}$, where $\beta = 2$ if $\gamma_j \in \{0, \pi\}$ and $\beta = 1$ otherwise. For this reason we define

$$\hat{d}_j^* = \hat{d}_j / \hat{\beta}, \quad \text{where} \tag{12}$$
$$\hat{\beta} = 1 + I_{(\hat{\gamma}_j < \lambda_{\kappa^*} \text{ or } \hat{\gamma}_j > \pi - \lambda_{\kappa^*})}$$

and the function $I_{(\hat{\gamma}_j<\lambda_{\kappa^*} \text{ or } \hat{\gamma}_j>\pi-\lambda_{\kappa^*})}$ is an indicator function that takes on the value 1, if $\hat{\gamma}_j$ is one of the κ^* Fourier frequencies closest to 0 or to π. If $\frac{\kappa^*}{T}+\frac{1}{\kappa^*}\to 0$, we have $\lambda_{\kappa^*}\to 0$ and $\lambda_{n-\kappa^*}\to\pi$, so that $\hat{\beta}$ is a consistent estimator for the power law coefficient β. For our filtering procedure we determine the bandwidth $m=\lfloor 1+T^{\xi}\rfloor$ by changing the bandwidth parameter $0<\xi<1$.

The consistency and asymptotic normality of the pooled estimator for $d_j\in(-1/2,1/2)$ follow immediately from Theorem 2 in Arteche and Robinson (2000) und Assumptions 2 and 3. For the case of a pole at 0, Velasco (1999) shows, that the consistency extends to the interval $d_j\in(-1/2,1)$ under conditions very similar to those in Robinson (1995a) and Arteche and Robinson (2000). This suggests that our procedure could also be applied within this interval. For $d_j\in(-0.5,2)$ alternatively the exact local Whittle approach of Shimotsu and Phillips (2005) can be used.

5 Feasible Automatic Model Order Selection

With the results from the previous sections we no longer need to assume that the true cyclical frequencies $\gamma_{j,0}$, the fractional exponents $d_{j,0}$ and $P\left(\Delta^{k^{(i)}}X_t=u_t \mid X_1,...,X_T\right)$ were known as we did in Section 2.

If we replace $P\left(\Delta^{k^{(i)}}X_t=u_t \mid X_1,...,X_T\right)$ in (7) by $p(G_X^*>\tilde{z})$ from Proposition 1, we obtain the feasible estimator:

$$\hat{k}=\min\left\{k^{(i)} \text{ such that } p(G_X^*>\tilde{z})>\alpha\right\}. \tag{13}$$

The feasible model order selection procedure is carried out in the following steps:

Step 1: Apply the filter $\Delta^{k^{(i)}}$ to the time series X_t.

Step 2: Test whether there are any significant poles in the spectrum of $\Delta^{k^{(i)}}X_t$ using the modified G^*-test in (10). Proceed to Step 3 if the null hypothesis $H_0: k_0=k^{(i)}$ is rejected - otherwise go to Step 5.

Step 3: Estimate γ_j and d_j using the estimators defined in (11) and (12).

Step 4: Set $i=i+1$ and go back to Step 1.

Step 5: Estimate k with the estimator given in (13).

Due to the consistency of the modified G^*-test established in Proposition 1, it asymptotically has a power of 1 if $k^{(i)}<k_0$, so that the test rejects until $k^{(i)}=k_0$. If $k^{(i)}=k_0$, then there is a probability of α that a type I error occurs and our procedure selects a model order of $\hat{k}>k_0$. To achieve consistency for $\hat{k}$, we follow Bai (1997) and make the size dependent on the sample size T so that $\alpha_T \to 0$, but very slowly. Under these conditions, we can establish the results in Proposition 2 below:

Proposition 2. *Suppose that X_t is given in (1) and that the size α_T converges to zero slowly ($\alpha_T \to 0$ yet $\lim_{T\to\infty} \inf T\alpha_T > 0$). Let G^*, $\hat{\gamma}_j$, $\hat{d}_j$ and $\hat{k}$ be defined as in equations (10), (11), (12) and (13). Then under Assumptions 1 to 3 and using the logspline spectral density estimate $\hat{f}(\lambda)$ we have $P(\hat{k} = k_0) \to 1$, as $T \to \infty$.*

As for Proposition 1, a proof for Proposition 2 can be found in the appendix. Of course $\alpha_T = \alpha$ can be kept constant in empirical applications. After the selection of appropriate long memory dynamics, the correct ARMA model orders p and q can be selected using an information criterion if the selected model is re-estimated parametrically using an approximate Whittle likelihood procedure. Giraitis and Leipus (1995) prove the consistency of this estimator. The semiparametric estimation results can be used as starting values for the numerical optimization.

As usual for semiparametric estimators there is the problem of an optimal bandwidth choice. In this case we have to select ζ that determines the number of knots in the smoothing spline via $H_T = \lfloor 1 + T^{\zeta} \rfloor$ and ξ that determines the usual bandwidth $m = \lfloor 1 + T^{\xi} \rfloor$. In the case of only a single pole at the origin the MSE-optimal bandwidth m could be selected with the procedure of Henry and Robinson (1996) and Henry (2001). For multiple poles however, no such results are available. An additional complication in the k-factor GARMA model is, that the parameter estimates might be negatively effected in small samples if the selected bandwidth is too large, so that Assumption 2 is not sufficient to inhibit that the effects from neighbouring poles interfere within the selected bandwidth. Hassler and Olivares (2013) show that a conservative deterministic bandwidth selection outperforms data-driven approaches in most situations.

In addition to that, simulation results not reported here show, that $\hat{k} > k_0$ mainly occurs in the presence of short memory dynamics if $\hat{f}(\lambda)$ is not flexible enough to remove peaks in the spectrum that originate from the short memory component. This typically leads to a selection of several cyclical frequencies $\hat{\gamma}_j$ within a very narrow frequency interval. We thus recommend to repeat the model selection procedure using a grid of different values for ξ and ζ and then to select the largest model with clearly distinct cyclical frequencies $\hat{\gamma}_j$.

Situations with $\hat{k} < k_0$ can be identified by plotting the autocorrelation function of the filtered process $\Delta^{\hat{k}} X_t$. If the selected model is specified correctly, the hyperbolically decaying sinusoidal pattern should be removed. This is the strategy we will follow in our empirical application in Section 7.

6 Monte Carlo Study

In this section we conduct Monte Carlo experiments to evaluate the performance of our model order selection procedure in various situations. To separate the performance of the sequential filtering procedure from that of the modified G^*-test, we first use Walker's g-test to detect significant periodic behavior. In this setup we start by considering the case of a single cyclical frequency and investigate how the performance of the selection procedure depends on the location of the pole and the magnitude of the long memory parameter. Then, we analyze how the procedure performs if the number of poles is increased.

Subsequently, we replace Walker's g-test with the modified G^*-test and analyze the robustness of the selection procedure if short memory dynamics and conditional heteroscedasticity are present. All results shown are obtained with M=5000 Monte Carlo repetitions. Simulations from the Gegenbauer process are generated using its AR(∞)-representation with a truncation after 1000 lags and a burnin period of 1000 observations. In all experiments the shorter series uses only the first $T_1 < T_2$ observations of the longer series. The DGP is always the k-factor GARMA process from equation (1). The trimming parameter is kept constant at $l = 1$ as in Arteche and Robinson (2000) and the bandwidth parameter is kept at $\xi = 0.7$ if not stated differently. Hereafter, we will use the term "power" to refer to the ability of the procedure to identify the true model order k_0.

6.1 Power Depending on the Location of the Pole and the Number of Cyclical Frequencies

First, we consider the 1-factor Gegenbauer process $(1-2\cos\gamma L + L^2)^d X_t = \varepsilon_t$ that does not contain any short memory dynamics. In this case the filtered process $\Delta^{k^{(i)}} X_t$ equals the white noise sequence ε_t under the null hypothesis in every step of the filtering procedure. Consequently, we can use Walker's original large sample statistic (9) instead of the modified G^*-statistic (10) to determine whether the filtered process contains significant periodicity. We allow the fractional exponent d that determines the shape of the pole to increase in steps of 0.05 from 0 to 0.45 and we shift the location of the pole in 14 steps from $\gamma = 0$ to $\gamma = \pi$.[2] The sample sizes considered are given by $T \in \{500, 1000, 2000, 5000\}$. Table 2 in the appendix gives a detailed overview of the results obtained in these simulations. A graphical summary of the results is depicted in Figure 1.

Since we are interested in the effects of γ, T and d, we always keep one of the parameters

[2]Because of the change of the power law parameter β discussed above the results for $\gamma = 0$ and $\gamma = \pi$ are simulated with half of the d reported.

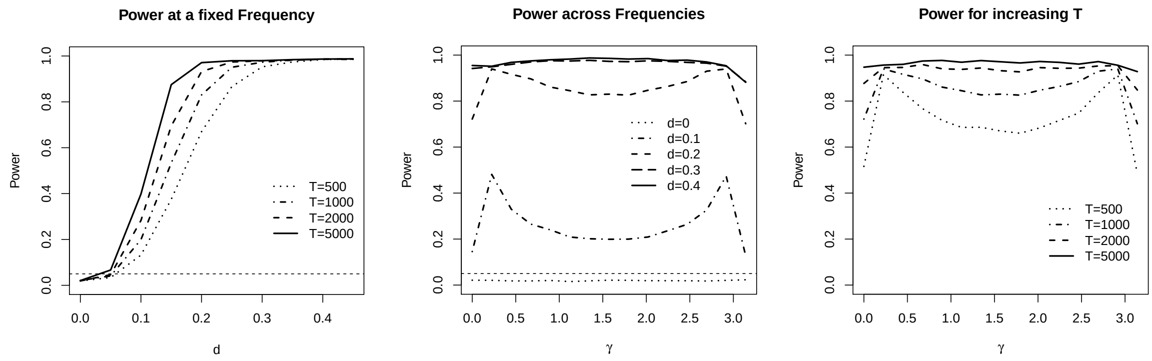

Figure 1: The DGP is $(1-2\cos\gamma L+L^2)^d X_t=\varepsilon_t$. On the left the cyclical frequency is fixed to $\gamma=\pi/2$ and T is increased. In the middle the sample size is fixed to $T=1000$ and d is increased. In the graph on the right $d=0.2$ is fixed and T is increased.

fixed and show how the power depends on the other two. The graph on the left shows power curves if the pole is fixed to the frequency $\gamma=\pi/2$. It can be seen, that the power of the selection procedure is increasing in the magnitude of the fractional exponent d and the sample size T. For $d\leq 0.05$ however, it barely increases. This is because the location of the pole can only be identified consistently if the fractional exponent is bounded away from zero as stated in Assumption 1. Note that for the reasons discussed in Section 5, the maximal power that the selection procedure can achieve in practical applications is $1-\alpha$. It is clear to see that the procedure is slightly conservative in the size case where $d=0$.

The graph in the middle shows the power for a fixed sample size of $T=1000$ across the spectrum of possible periodic frequencies γ and for different values of d. One can see, that the conservativeness of the procedure and the non-existing power for small d occur independent of the location of the periodic frequency. For smaller d the power is increasing with increasing distance of γ from $\pi/2$. Only for $\gamma\in\{0,\pi\}$ there is a drop in power again. This is caused by the fact that the fractional exponents at frequencies further away from the boundaries are estimated using $2m$ Fourier frequencies - that is m frequencies on either side of the pole. At the boundaries however, we can only use $m_a<m$ Fourier frequencies on one side of the pole so that the estimator for the fractional exponent has a higher variance at these frequencies.

In the graph on the right in Figure 1 we keep d fixed at 0.2 and the curves show the power across frequencies for increasing T, so that one can see that the increasing power in T remains intact across all frequencies and the power does not depend on the cyclical frequency γ asymptotically.

Since the purpose of our procedure is to select the true model order k_0 if there are multiple cyclical frequencies, we now allow the number of cyclical frequencies k to take values from

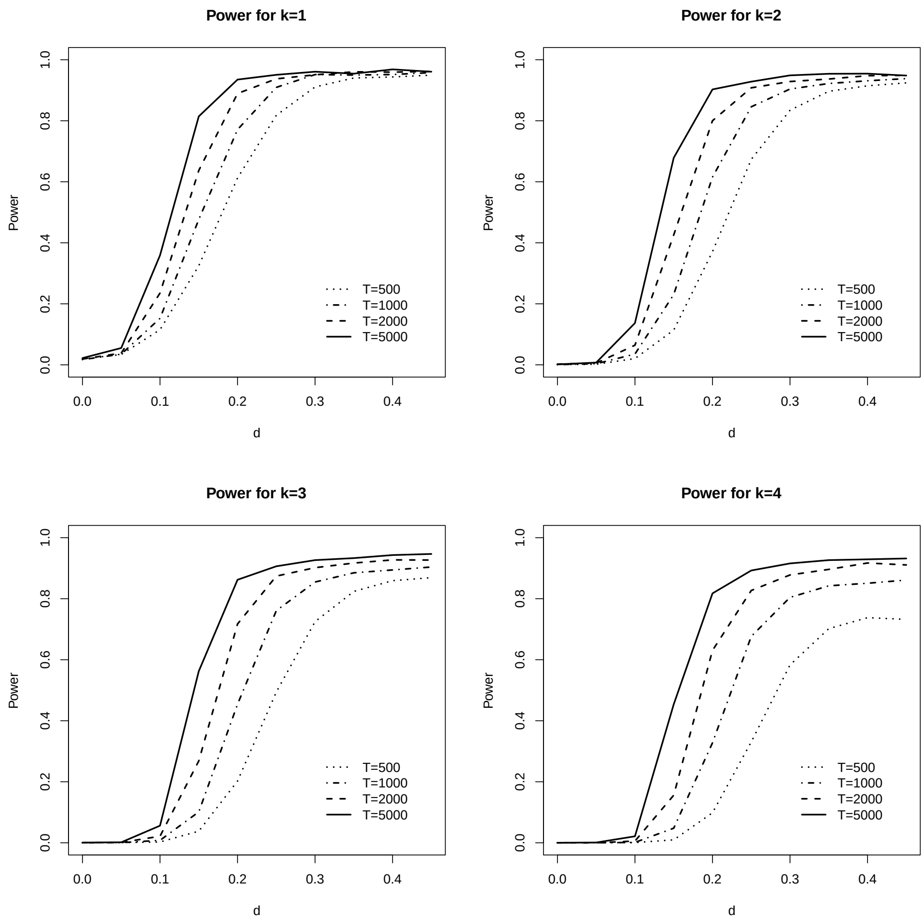

Figure 2: The DGP is $\prod_{j=1}^{k}(1-2\cos\gamma_j L + L^2)^{d_j} X_t = \varepsilon_t$ with $\gamma_j = (\pi j)/(k+1)$. Power curves are shown for $k \in \{1,2,3,4\}$ and increasing sample sizes.

1 to 4. The DGP is given by $\prod_{j=1}^{k}(1-2\cos\gamma_j L + L^2)^{d_j} X_t = \varepsilon_t$, but the specification of the test, the sample sizes and the values of the fractional exponents considered are the same as before. We set $\gamma_j = (\pi j)/(k+1)$, so that the seasonal frequencies are equally spaced and the outer minimal and maximal periodic frequencies have the distance $\pi/(k+1)$ to the boundaries. All long memory parameters are kept equal at $d_j = d$. To avoid interference effects from neighboring poles we now use $\xi = 0.55$. The results of this experiment are shown in Figure 2. If d and T remain constant, we can observe that the power is lower for higher k. For $k = 4$ we also observe that the power starts to become non-monotonic in d for smaller sample sizes. This is due to the interference effect from neighboring poles discussed in Section 5. Nevertheless, asymptotically the procedure selects the right model order with probability $1-\alpha_T$. So the experiments show that the sequential-g procedure also works well for larger k.

6.2 The Influence of Short Memory Dynamics and Conditional Heteroscedasticity

So far we have only considered the restricted version of the model selection procedure that is based on Walker's original g-test given in (9). Under the null hypothesis, this test assumes that the process is *iid*. Since short memory dynamics are present in most applications, we now consider the case in which the residuals from (5) show stationary ARMA behavior.

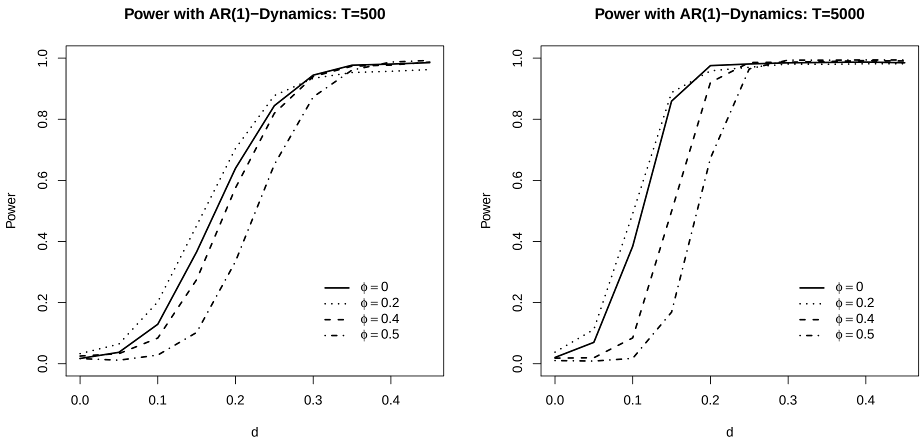

Figure 3: Power curves of the sequential G^*-procedure for $(1-\phi L)(1-2\cos\frac{\pi}{2}L+L^2)^d X_t = \varepsilon_t$ with increasing autoregressive parameters ϕ.

It is well known that the semiparametric estimators of the fractional exponents d_j are asymptotically unaffected by the presence of short memory dynamics. The critical part of our procedure is the performance of the g-test. As discussed in Section 5, the procedure tends to find several spurious seasonal peaks in a very narrow interval if the spectral density is assumed to be that of an *iid*-process but in fact short memory dynamics are present. This is why we will now investigate the size and power properties of the procedure if the modified G^*-test from (10) is used that allows for the presence of stationary ARMA dynamics under the null hypothesis.

The simulated DGP is $(1-\phi L)(1-2\cos\frac{\pi}{2}L+L^2)^d X_t = \varepsilon_t$. This means the short memory component of the process takes the form of an AR(1). We consider the case of one pole at $\gamma = \pi/2$ with increasing fractional exponents and increasing persistence of the short memory process. The bandwidth parameter that determines the number of knots used for the logspline estimate is set to $\zeta = 0.1$ and the spline used is a cubic spline. Figure 3 illustrates the key findings. The complete results are given in Table 3 in the appendix. We can compare the size and power results for $\phi = 0$ with those in Table 2 obtained with the restricted g-test if the pole is at $\gamma = \pi/2$. As expected there are power losses if the unrestricted modified G^*-test is applied and no short memory dynamics are present ($\phi = 0$), but these are very small in magnitude. This means that Walker's large sample g-test can be replaced by the modified G^*-test at almost no costs. This can also be seen by comparing the solid curves in Figure 3 with the dotted and the solid curve in the graph on the left in Figure 1. These are virtually identical.

Figure 3 also shows, that the power of the test actually increases if there are weak AR(1) dynamics present (for $\phi \leq 0.3$). This effect could appear because part of the flexibility of the spline is used to fit the spectral density of the AR-part at the origin, so that there is less flexibility to fit the cyclical pole. For stronger AR-dynamics the power starts to drop for every given d and T and the model order is only correctly identified for larger memory parameters. There are two intuitive explanations for this effect. First, a smaller proportion of the variance of the process is attributed to the cyclical behavior if the AR-dynamics are stronger, so that it is harder to identify the poles from the periodogram. Second, it becomes harder to use the periodogram to distinguish peaks in the spectral density due to short memory behavior from poles due to long memory behavior if ϕ increases. Nevertheless, the sequential-G^* procedure still works well as long as the AR-dynamics remain moderate ($\phi \leq 0.5$). This is less restrictive than it might appear, because the persistence of the k-factor GARMA process is first and foremost caused by its long memory components and not by the ARMA component alone.

Since many of the potential applications of cyclical long memory models are in the area of financial high frequency data, it is also of interest to investigate whether our selection procedure remains valid under conditional heteroscedasticity. Robinson and Henry (1999) show that the consistency, the asymptotic normality and even the asymptotic variance of the Gaussian semiparametric estimator for the fractional exponents remain unaffected by conventional ARCH/GARCH type conditional heteroscedasticity and also by long memory conditional heteroscedasticity. Additional simulation studies showed that these results carry over to our model selection procedure, even though this also depends on the semiparametric estimator of Hidalgo and Soulier (2004) and the modified G^*-test. In fact, the power is virtually identical even with just 500 observations.[3]

7 Modeling Californian Electricity Loads with k-factor GARMA Models

As discussed in the introduction, GARMA models are useful in high frequency datasets that potentially exhibit multiple seasonalities. One such example are electricity load

[3]These results can be obtained from the authors upon request.

series. In particular the forecasting of the latter has attracted continued attention in the literature, because electricity demand is very volatile and it has to be matched by supply in real time. Since different means of electricity generation have very different marginal costs, precise forecasts are of major importance for electricity producers to schedule production capacities accordingly. Recently, GARMA models have been used to generate such forecasts by Soares and Souza (2006) who apply a 1-factor GARMA model to forecast electricity demand in the area of Rio de Janeiro (Brazil) and Diongue et al. (2009) who use a 3-factor GIGARCH model to forecast German spot market electricity prices.

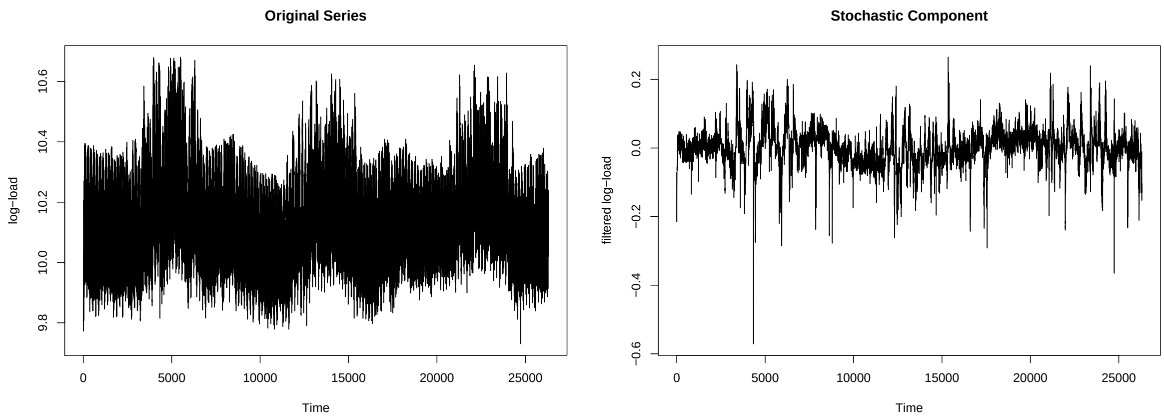

Figure 4: Californian system wide log-load series from 2000 to 2002 before and after deterministic seasonality is removed.

In general, two different approaches can be distinguished in the literature on electricity load forecasting. The global approach is to fit a relatively complex model to the hourly series, whereas local approaches forecast all 24 hourly series with separately estimated simpler models.

By applying our model selection procedure to electricity load series from the Californian power system, we complement this literature with an in-sample perspective on model selection in the global series.

The data is downloaded from UCEI[4] and covers the period from 2000 to 2002, which gives us 26304 hourly observations. A similar dataset was considered in Weron and Misiorek (2008) who focus on prices in the period 1999-2000. To remove deterministic seasonality upfront, we consider the residual series obtained from regressing the log of each hourly series separately on a trend and dummy variables for the day of the week and the month of the year. Insignificant dummies are discarded stepwise using a general-to-specific procedure. A similar approach was suggested in Haldrup and Nielsen (2006). This simple deterministic model already achieves an R^2 of approximately 90.49 percent,

[4]www.ucei.berkeley.edu

which implies that a large proportion of the seasonality in the series is deterministic. The original series as well as the residual series from this regression are plotted in Figure 4. It is obvious, that the variance of the series is strongly reduced. The left side of Figure 5 depicts the autocorrelation function of the residual series. It shows clearly that the series exhibits long memory and periodicity.

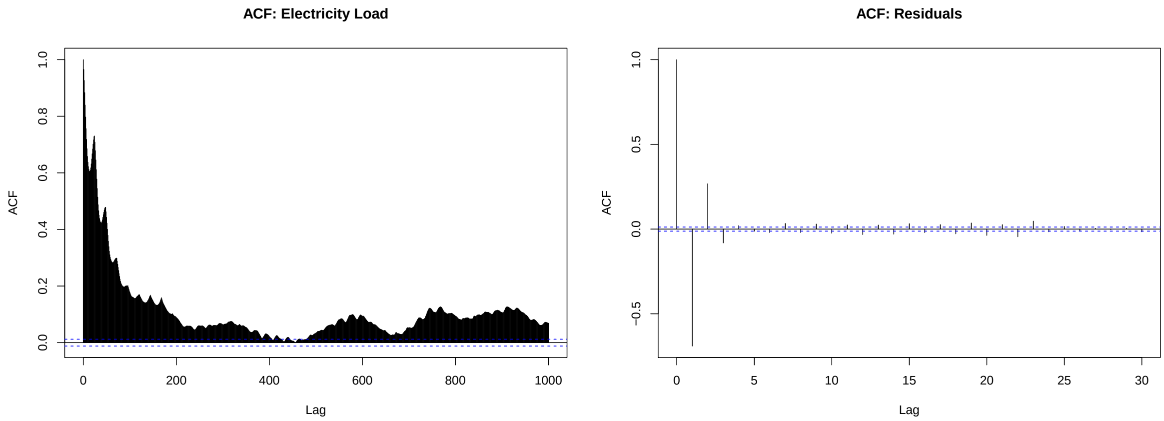

Figure 5: Autocorrelation function of the original series (left) and the residual series of 14-factor Gegenbauer process (right).

If one would discretionary fit a model based on a visual inspection of the periodogram shown on the left side of Figure 6, one would probably consider a two factor model with a pole at the origin and one in the neighborhood of $\lambda = 0.25$ which approximately corresponds to the daily frequency $2\pi/24$. Such a process could be approximated well by an ARFIMA process, since the effect at $\lambda = 0.25$ seems to be of a small magnitude. In sharp contrast to that, our model selection procedure indicates that a 14-factor GARMA process should be used to model the series. The intuition for this unexpected result becomes obvious, if one considers the residual series $\Delta^{k^{(i)}} X_t$ from (5) after filtering out the effect of the pole at the origin. The periodogram of the filtered process is depicted on the right hand side of Figure 6. The poles that become visible here are not visible in the original periodogram on the left hand side, because of the magnitude of the Fourier frequency closest to the origin. To include these effects, it is necessary to increase the model order.

As discussed in Section 5, the bandwidth choice has a critical influence on the results of the selection procedure. This is why we repeat the analysis using a grid of values for ξ and ζ, that determine the number of frequencies m used in the estimation of the long memory parameters d_j and the number of segments H_T used in the logspline estimate of the spectral density. We allow ξ to increase from 0.5 to 0.65 in steps of

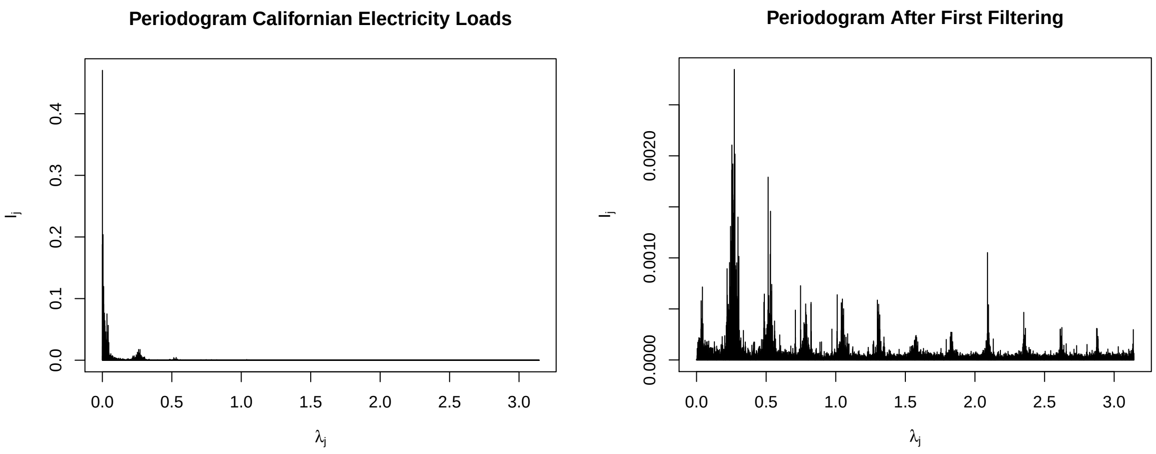

Figure 6: Periodogram the Californian load series (left) and the residual series $\Delta^{\hat{k}^{(i)}} X_t$ obtained after removing the non-cyclical long memory effects (right).

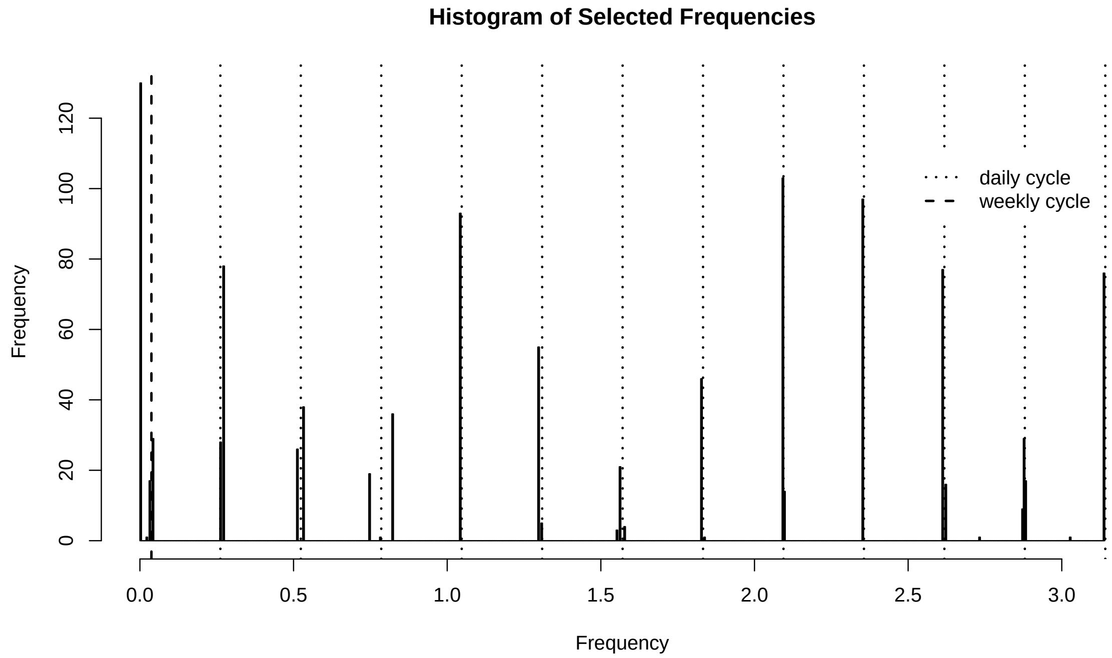

Figure 7: Histogram of the frequencies selected by the sequential-G^* procedure applied on a grid of bandwidth parameters.

0.01 and determine ζ so that the number of segments varies from 2 to 8. For every parameter constellation we store all estimated cyclical frequencies $\hat{\gamma}_j$. The histogram in Figure 7 shows which frequencies are selected and how often they are found to be significant. It can be seen, that not every of the cyclical frequencies is selected with every combination of bandwidth parameters. However, 14 frequencies are selected in the majority of parameter constellations. These are the zero frequency that corresponds to a non-seasonal long memory component, $2\pi/(24\times7)$ - the frequency of a weekly cycle and $2\pi\upsilon/(24)$, with $\upsilon = 1,...,12$ which is the frequency of a daily cycle and its harmonics. For better comparison, these theoretical frequencies are superimposed as dashed and dotted vertical lines in Figure 7.

These findings indicate, that the data is best explained by the 14-factor GARMA model with these respective frequencies, which confirms the initial expectation that there are multiple cycles in the data generating process. To fit the 14-factor GARMA model we estimate the relevant cyclical frequencies by the local modes in the histogram. These estimates $\hat{\gamma}_j$ along with the cycle length in days $\hat{T}_j$ and the estimated fractional exponents $\hat{d}_j$ are given in Table 1. If one considers the estimated fractional exponents $\hat{d}_j$, one observes that the exponent at the origin is larger than 0.25 which implies non-stationary but mean reverting long memory. The memory parameter of the daily frequency is 0.5550 which is also in the non-stationary but mean reverting region.[5] The other exponents lie between 0.11 and 0.34 and tend to decrease for higher frequencies.

j	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$\hat{d}_j$	0.3373	0.2401	0.5550	0.3469	0.1314	0.3237	0.2441	0.2111	0.1902	0.2042	0.1979	0.1803	0.1241	0.0578
$\hat{\gamma}_j$	0.0025	0.0425	0.2625	0.5325	0.8225	1.0425	1.2975	1.5625	1.8275	2.0925	2.3525	2.6125	2.8775	3.1375

Table 1: Fitted 13-factor Gegenbauer model. Standard errors for d_j are 0.0233 for γ_1 and γ_{14}, 0.0200 for γ_2, and 0.0166 for all other periodic frequencies.

The autocorrelation function of the residuals from the 14-factor model is plotted on the right hand side of Figure 5. Even though there is still some unmodeled short term dependence in the data, the total degree of dependence is greatly reduced. The R^2 obtained is as high as 84.97 percent, which means that the deterministic model and the Gegenbauer model together explain approximately 98.57 percent of the variation of the original series.

Based on the visual inspection of the periodogram we would have chosen a 2-factor GARMA model with non-seasonal long memory and a daily cycle. In contrast to that, our model order selection procedure finds a 14-factor model that also includes cycles at the weekly frequency and the harmonics of the daily frequency. These different model choices have very different implications. In the case of the 2-factor model no weekly cycle is detected and one would conclude that the weekly seasonality in the original series is best modeled by deterministic dummies. Furthermore, considering the local series for every hour of the day separately can be done with a relatively small loss of information if there are no intraday cycles. Our findings on the other hand imply, that there is

[5] The non-stationary part of the parameter space for the d_j was excluded from our analysis in the preceding sections. However, we discussed in Section 4 that the results can be expected to carry over to the mean reverting region where $d_j \in (0.5, 1)$. This is also supported by simulations.

indeed a stochastic weekly cycle and that there are several significant intraday cycles at the harmonic frequencies that lead to a considerable loss of information when local models are considered. In fact, the exemptional fit of the 14-factor model for the global series shows that this series can indeed be modeled directly if the daily load profile is accounted for by a suitable seasonal model.

8 Conclusion

We introduced an automatic model order selection procedure for k-factor GARMA models that is based on tests of the maximum of the periodogram and semiparametric estimations of the locations of the poles as well as the fractional exponents and can be used for general model selection in cyclical long-memory processes. As a byproduct we suggest a modified test for persistent periodicity in stationary ARMA models. Our simulation studies show that the procedure performs well if there is a single pole, with several poles, under additional short memory dynamics and under conditional heteroscedasticity.

Our procedure allows an easier application of k-factor GARMA models in empirical analyses and prevents the use of false model specifications that are based on discretionary decisions after a visual analysis of the periodogram. As the example of the Californian electricity load series in Section 7 shows, the periodogram can be very misleading if it is used as a tool for the selection of the model order.

We also gain new insights into the behavior of electricity load series. It turns out that the stochastic variation of the Californian electricity loads can be modeled very well by a 14-factor GARMA process. The fit achieved suggests that it can indeed be a good strategy to model the global series directly instead of fitting separate models for every hour of the day. This is especially important for short term forecasts since the 1-hour-ahead forecast of a local model only uses data up to 23 hours ago and does not utilize the information contained in the most recent observations.

These insights demonstrate the potential of k-factor GARMA models for the modeling of periodic time series like the electricity load example considered here or the intraday trading volume and volatility series discussed in the introduction. The semiparametric estimators used for the selection procedure require low computational effort since they only utilize periodogram ordinates local to the respective pole, which makes them easy to apply in large high frequency datasets.

Appendix

T=500															
d/γ	**0**	**.22**	**.45**	**.67**	**.9**	**1.12**	**1.35**	**1.57**	**1.8**	**2.02**	**2.24**	**2.47**	**2.69**	**2.92**	**π**
0	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.01
.05	0.03	0.06	0.05	0.04	0.04	0.05	0.04	0.03	0.04	0.04	0.04	0.04	0.05	0.06	0.02
.1	0.10	0.36	0.24	0.18	0.16	0.15	0.14	0.13	0.14	0.15	0.16	0.20	0.24	0.35	0.08
.15	0.25	0.74	0.58	0.47	0.40	0.39	0.38	0.38	0.38	0.40	0.43	0.47	0.58	0.74	0.26
.2	0.52	0.91	0.84	0.77	0.72	0.69	0.69	0.67	0.66	0.68	0.72	0.75	0.84	0.91	0.48
.25	0.74	0.95	0.93	0.92	0.90	0.88	0.86	0.87	0.87	0.88	0.89	0.91	0.94	0.94	0.69
.3	0.87	0.95	0.95	0.96	0.96	0.95	0.95	0.95	0.95	0.95	0.95	0.95	0.96	0.94	0.78
.35	0.93	0.95	0.96	0.98	0.97	0.97	0.97	0.97	0.97	0.98	0.98	0.97	0.96	0.94	0.82
.4	0.95	0.95	0.96	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.96	0.95	0.83
.45	0.96	0.95	0.97	0.98	0.98	0.98	0.98	0.99	0.98	0.98	0.98	0.98	0.97	0.95	0.81
T=1000															
d/γ	**0**	**.22**	**.45**	**.67**	**.9**	**1.12**	**1.35**	**1.57**	**1.8**	**2.02**	**2.24**	**2.47**	**2.69**	**2.92**	**π**
0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
.05	0.03	0.09	0.07	0.05	0.05	0.05	0.04	0.04	0.05	0.05	0.05	0.05	0.07	0.09	0.04
.1	0.15	0.48	0.33	0.26	0.24	0.21	0.20	0.20	0.20	0.21	0.24	0.26	0.33	0.47	0.13
.15	0.41	0.86	0.73	0.65	0.57	0.55	0.52	0.53	0.54	0.55	0.59	0.65	0.74	0.86	0.39
.2	0.72	0.94	0.92	0.90	0.86	0.85	0.83	0.83	0.83	0.85	0.86	0.88	0.93	0.94	0.70
.25	0.90	0.94	0.96	0.96	0.96	0.95	0.96	0.95	0.95	0.96	0.96	0.96	0.96	0.95	0.85
.3	0.94	0.95	0.96	0.97	0.98	0.97	0.98	0.97	0.97	0.98	0.97	0.97	0.96	0.95	0.88
.35	0.96	0.95	0.97	0.98	0.98	0.97	0.98	0.98	0.98	0.98	0.97	0.97	0.97	0.95	0.88
.4	0.95	0.95	0.97	0.97	0.98	0.98	0.99	0.99	0.98	0.98	0.98	0.98	0.97	0.95	0.88
.45	0.96	0.95	0.97	0.98	0.98	0.98	0.99	0.99	0.98	0.99	0.98	0.98	0.97	0.94	0.87
T=2000															
d/γ	**0**	**.22**	**.45**	**.67**	**.9**	**1.12**	**1.35**	**1.57**	**1.8**	**2.02**	**2.24**	**2.47**	**2.69**	**2.92**	**π**
0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
.05	0.03	0.12	0.08	0.06	0.06	0.06	0.05	0.05	0.06	0.06	0.07	0.07	0.09	0.12	0.04
.1	0.22	0.61	0.45	0.36	0.33	0.28	0.27	0.28	0.28	0.29	0.32	0.37	0.45	0.62	0.21
.15	0.59	0.91	0.85	0.80	0.75	0.71	0.69	0.69	0.70	0.73	0.75	0.79	0.86	0.93	0.58
.2	0.88	0.95	0.95	0.96	0.94	0.94	0.94	0.93	0.93	0.95	0.94	0.94	0.95	0.95	0.85
.25	0.95	0.96	0.96	0.97	0.97	0.97	0.98	0.97	0.97	0.98	0.97	0.97	0.97	0.96	0.90
.3	0.95	0.96	0.97	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.97	0.97	0.95	0.92
.35	0.96	0.95	0.97	0.98	0.98	0.98	0.99	0.98	0.98	0.98	0.98	0.97	0.98	0.96	0.91
.4	0.97	0.96	0.97	0.98	0.98	0.98	0.99	0.99	0.98	0.99	0.98	0.97	0.97	0.95	0.91
.45	0.97	0.95	0.98	0.98	0.98	0.98	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.95	0.90
T=5000															
d/γ	**0**	**.22**	**.45**	**.67**	**.9**	**1.12**	**1.35**	**1.57**	**1.8**	**2.02**	**2.24**	**2.47**	**2.69**	**2.92**	**π**
0	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
.05	0.06	0.16	0.11	0.08	0.07	0.07	0.07	0.07	0.07	0.07	0.08	0.08	0.11	0.15	0.06
.1	0.35	0.78	0.60	0.51	0.46	0.42	0.41	0.40	0.41	0.43	0.46	0.52	0.62	0.79	0.36
.15	0.82	0.94	0.92	0.93	0.90	0.89	0.89	0.88	0.87	0.90	0.90	0.91	0.94	0.94	0.81
.2	0.95	0.96	0.96	0.97	0.98	0.97	0.98	0.97	0.97	0.97	0.97	0.96	0.97	0.96	0.93
.25	0.97	0.96	0.97	0.98	0.97	0.98	0.98	0.98	0.97	0.98	0.98	0.97	0.97	0.96	0.94
.3	0.96	0.96	0.97	0.97	0.98	0.98	0.98	0.98	0.98	0.99	0.98	0.98	0.98	0.97	0.94
.35	0.96	0.97	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.96	0.95
.4	0.97	0.96	0.98	0.98	0.98	0.98	0.99	0.99	0.98	0.99	0.98	0.98	0.98	0.96	0.95
.45	0.97	0.96	0.98	0.98	0.98	0.99	0.98	0.99	0.99	0.99	0.99	0.98	0.98	0.96	0.95

Table 2: shows how often the sequential-g procedure selects the right model order $k = 1$ for different d and at different frequencies γ.

T=500						
d*/*ϕ	**0**	**.1**	**.2**	**.3**	**.4**	**.5**
0	0.02	0.02	0.03	0.03	0.03	0.02
.05	0.04	0.06	0.06	0.06	0.03	0.01
.1	0.13	0.18	0.20	0.17	0.08	0.03
.15	0.37	0.43	0.45	0.41	0.28	0.10
.2	0.64	0.69	0.70	0.70	0.58	0.33
.25	0.84	0.86	0.88	0.87	0.82	0.65
.3	0.94	0.94	0.93	0.93	0.94	0.87
.35	0.98	0.97	0.95	0.95	0.97	0.96
.4	0.98	0.97	0.96	0.97	0.98	0.99
.45	0.99	0.97	0.96	0.97	0.99	0.99

T=1000						
d*/*ϕ	**0**	**.1**	**.2**	**.3**	**.4**	**.5**
0	0.02	0.03	0.03	0.03	0.03	0.01
.05	0.04	0.07	0.09	0.05	0.03	0.01
.1	0.19	0.26	0.27	0.21	0.08	0.02
.15	0.50	0.57	0.61	0.54	0.33	0.11
.2	0.81	0.85	0.86	0.84	0.69	0.42
.25	0.95	0.95	0.94	0.95	0.92	0.77
.3	0.98	0.97	0.96	0.97	0.98	0.95
.35	0.98	0.97	0.97	0.98	0.99	0.99
.4	0.99	0.98	0.97	0.98	0.99	0.99
.45	0.99	0.98	0.98	0.99	0.99	1.00

T=200						
d*/*ϕ	**0**	**.1**	**.2**	**.3**	**.4**	**.5**
0	0.02	0.03	0.03	0.03	0.03	0.01
.05	0.05	0.08	0.10	0.05	0.02	0.01
.1	0.26	0.34	0.36	0.24	0.08	0.01
.15	0.67	0.74	0.75	0.67	0.38	0.13
.2	0.92	0.94	0.93	0.92	0.81	0.51
.25	0.97	0.97	0.96	0.98	0.97	0.87
.3	0.98	0.97	0.97	0.98	0.99	0.98
.35	0.98	0.98	0.97	0.98	0.99	0.99
.4	0.98	0.98	0.98	0.99	0.99	0.99
.45	0.99	0.98	0.98	0.99	0.99	0.99

T=5000						
d*/*ϕ	**0**	**.1**	**.2**	**.3**	**.4**	**.5**
0	0.02	0.03	0.04	0.02	0.02	0.01
.05	0.07	0.11	0.11	0.05	0.02	0.01
.1	0.38	0.48	0.49	0.29	0.08	0.02
.15	0.86	0.89	0.89	0.79	0.50	0.17
.2	0.98	0.96	0.96	0.97	0.92	0.67
.25	0.98	0.97	0.97	0.98	0.99	0.96
.3	0.98	0.97	0.98	0.98	0.99	0.99
.35	0.99	0.98	0.98	0.99	0.99	0.99
.4	0.99	0.98	0.98	0.99	0.99	0.99
.45	0.98	0.98	0.98	0.99	0.99	0.99

Table 3: shows how often the sequential-G^* procedure selects the right model order in presence of AR(1)-dynamics with the autoregressive parameter ϕ. The DGP has k=1 at frequency $\gamma = \pi/2$.

Proof of Proposition 1:

1. For the linear process Z_t with $\sum_{j=0}^{\infty}|a_j| < \infty$ and the smoothed periodogram estimator with a "lag window" of M^* covariances we have $\hat{f}(\lambda) = f(\lambda) + O(\sqrt{M^*/T})$ as a direct consequence of Theorem 10.4.1 in Brockwell and Davis (2009), so that we have pointwise consistency for $\frac{1}{M^*} + \frac{M^*}{T} \to \infty$ and the asymptotic distribution of G^* only depends on that of $I(\lambda_\kappa)$. Theorem 10.3.2 in Brockwell and Davis (2009) further implies that $\left(\frac{I(\lambda_1)}{f(\lambda_1)}, ..., \frac{I(\lambda_n)}{f(\lambda_n)}\right) \xrightarrow{d} (Q_1, ..., Q_n)$, where $Q_1, ..., Q_n$ are independent exponential variables, so that $p(G_X^* > \tilde{z}) = 1 - (1 - \exp(-\tilde{z}/2))^n$ still holds under the null hypothesis of a short memory process.

2. To prove the consistency of the G^*-test, first consider $\max(I_\kappa)$. From Hidalgo and Soulier (2004) we have that $\hat{\gamma}_j = \arg\max_{\lambda_\kappa}(I(\lambda_\kappa)) \to \gamma_j$ with rate $T\log^{\tilde{b}}(T)$, for any $\tilde{b} > c_j/2$, where c_j is the constant that bounds away d_j from 0 in Assumption 1. It is known that for long memory processes $\lim_{T\to\infty} E\left(\frac{I_\kappa}{f(\lambda_\kappa)}\right) \neq 1$. However, Theorem 1 of Robinson (1995b) states that $\lim_{T\to\infty} E(I(\lambda_\kappa)) \propto f(\lambda_\kappa)$. In addition to that, Lemma 2 of Hidalgo and Soulier (2004) states, that $\frac{I_\kappa}{\sqrt{E(I_\kappa)}} = U_\kappa$ converges weakly to a distribution without mass at 0. Together this implies that $I(\lambda_\kappa) \propto U_\kappa f(\lambda_\kappa)$. Since $f(\gamma_j) = \infty$ this implies that $\lim_{T\to\infty} P(I(\hat{\gamma}_j) < \tilde{c}) = 0 \ \forall \ \tilde{c} < \infty$.

   Now consider the logspline estimate $\hat{f}(\lambda)$. If $\hat{f}(\lambda)$ remained a consistent estimate, it would converge with rate $\sqrt{H_T/T}$ which is much more slowly than $T\log^{\tilde{b}}(T)$. This means that asymptotically $\hat{f}(\hat{\gamma}_j)$ remains bounded relative to $I(\hat{\gamma}_j)$, so that $\lim_{T\to\infty} P\left(\max\left(\frac{I_\kappa}{f(\lambda_\kappa)}\right) = \frac{\max(I_\kappa)}{f(\lambda_\kappa)}\right) = 1$ and consequently $\lim_{T\to\infty} P(G^* = \infty) = 1$. □

Proof of Proposition 2:

The consistency of the model selection procedure can be established in analogy to that of the sequential structural break test in Bai (1997).
Consider the event $k^{(i)} < k_0$ and denote by c_T the critical value associated with the significance level α_T. If the estimated number of poles is less than the true number, there exists at least one pole in the spectrum. Due to the result in Proposition 1, we know that in this situation $P(G^* > c_T) \to 1$ as $T \to \infty$. The null hypothesis of no seasonal long memory will thus be rejected with probability tending to 1 and the next iteration of the model selection procedure begins. This implies that $P(\hat{k} < k_0)$ converges to zero as the sample size increases.
Alternatively, consider the event $\hat{k} > k_0$. This event can arise if there is a type I error in the $k_0 + 1$th iteration of the G^*-test or if there are multiple consecutive type I errors with the first one occurring in the $k_0 + 1$th iteration. We thus have $P(\hat{k} = k_0 + r) \to \alpha_T^r$, $\forall\ r \geq 1$. Consistency is achieved since $\alpha_T \to 0$, so that $P(\hat{k} = k + r) \to 0\ \forall\ r \neq 0$. From $P(\hat{k} < k_0) \to 0$ and $P(\hat{k} > k_0) \to 0\ \forall\ r = 0, 1, 2, ...$ follows immediately $P(\hat{k} = k_0) \to 1$. $\square$

References

Andersen, T. G. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4(2):115–158.

Arteche, J. and Robinson, P. M. (2000). Semiparametric inference in seasonal and cyclical long memory processes. *Journal of Time Series Analysis*, 21(1):1–25.

Bai, J. (1997). Estimating multiple breaks one at a time. *Econometric Theory*, 13(03):315–352.

Bisaglia, L., Bordignon, S., and Lisi, F. (2003). k-Factor GARMA models for intraday volatility forecasting. *Applied Economics Letters*, 10(4):251–254.

Bordignon, S., Caporin, M., and Lisi, F. (2008). Periodic long-memory GARCH models. *Econometric Reviews*, 28(1-3):60–82.

Brockwell, P. J. and Davis, R. A. (2009). *Time series: theory and methods*. Springer.

Cogburn, R., Davis, H. T., et al. (1974). Periodic splines and spectral estimation. *The Annals of Statistics*, 2(6):1108–1126.

Diongue, A. K., Guegan, D., and Vignal, B. (2009). Forecasting electricity spot market prices with a k-factor gigarch process. *Applied Energy*, 86(4):505–510.

Fisher, R. A. (1929). Tests of significance in harmonic analysis. *Proceedings of the Royal Society of London. Series A*, 125(796):54–59.

Gil-Alana, L. A. (2002). Seasonal long memory in the aggregate output. *Economics Letters*, 74(3):333–337.

Gil-Alana, L. A. (2007). Testing the existence of multiple cycles in financial and economic time series. *Annals of Economics & Finance*, 8(1).

Giraitis, L. and Leipus, R. (1995). A generalized fractionally differencing approach in long-memory modeling. *Lithuanian Mathematical Journal*, 35(1):53–65.

Gray, H. L., Zhang, N.-F., and Woodward, W. A. (1989). On generalized fractional processes. *Journal of Time Series Analysis*, 10(3):233–257.

Haldrup, N. and Nielsen, M. Ø. (2006). A regime switching long memory model for electricity prices. *Journal of Econometrics*, 135(1):349–376.

Hassler, U. (1994). (mis)specification of long memory in seasonal time series. *Journal of Time Series Analysis*, 15(1):19–30.

Hassler, U. and Olivares, M. (2013). Semiparametric inference and bandwidth choice under long memory: experimental evidence. *İstatistik, Journal of the Turkish Statistical Association*, 6(1):27–41.

Hassler, U., Rodrigues, P. M., and Rubia, A. (2009). Testing for general fractional integration in the time domain. *Econometric Theory*, 25:1793–1828.

Henry, M. (2001). Robust automatic bandwidth for long memory. *Journal of Time Series Analysis*, 22(3):293–316.

Henry, M. and Robinson, P. (1996). Bandwidth choice in gaussian semiparametric estimation of long range dependence. In *Athens Conference on Applied Probability and Time Series Analysis*, pages 220–232. Springer.

Hidalgo, J. and Soulier, P. (2004). Estimation of the location and exponent of the spectral singularity of a long memory process. *Journal of Time Series Analysis*, 25(1):55–81.

Kooperberg, C., Stone, C. J., and Truong, Y. K. (1995). Rate of convergence for logspline spectral density estimation. *Journal of Time Series Analysis*, 16(4):389–401.

Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *Journal of the American Statistical Association*, 85(410):338–344.

Priestley, M. B. (1981). *Spectral analysis and time series.* Academic press.

Robinson, P. M. (1994). Efficient tests of nonstationary hypotheses. *Journal of the American Statistical Association*, 89(428):1420–1437.

Robinson, P. M. (1995a). Gaussian semiparametric estimation of long range dependence. *The Annals of Statistics*, 23:1630–1661.

Robinson, P. M. (1995b). Log-periodogram regression of time series with long range dependence. *The Annals of Statistics*, 23:1048–1072.

Robinson, P. M. and Henry, M. (1999). Long and short memory conditional heteroskedasticity in estimating the memory parameter of levels. *Econometric Theory*, 15(03):299–336.

Rossi, E. and Fantazzini, D. (2014). Long memory and Periodicity in Intraday Volatility. *Journal of Financial Econometrics*.

Shimotsu, K. and Phillips, P. C. (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics*, 33(4):1890–1933.

Soares, L. J. and Souza, L. R. (2006). Forecasting electricity demand using generalized long memory. *International Journal of Forecasting*, 22(1):17–28.

Velasco, C. (1999). Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis*, 20(1):87–127.

Weron, R. and Misiorek, A. (2008). Forecasting spot electricity prices: A comparison of parametric and semiparametric time series models. *International Journal of Forecasting*, 24(4):744–763.

Woodward, W. A., Cheng, Q. C., and Gray, H. L. (1998). A k-factor GARMA long-memory model. *Journal of Time Series Analysis*, 19(4):485–504.

Yajima, Y. (1996). Estimation of the frequency of unbounded spectral densities. Proceedings of The Business and Economic Statistics Section, American Statistical Association.