Generalised Linear Cepstral Models for the Spectrum of a Time Series

Tommaso Proietti[†]
University of Rome 'Tor Vergata'
and CREATES, Aarhus

Alessandra Luati[*]
University of Bologna

Abstract

The paper introduces the class of generalised linear models with Box-Cox link for the spectrum of a time series. The Box-Cox transformation of the spectral density is represented as a finite Fourier polynomial, with coefficients, that we term generalised cepstral coefficients, providing a complete characterisation of the properties of the random process. The link function depends on a power transformation parameter and encompasses the exponential model (logarithmic link), the autoregressive model (inverse link), and the moving average model (identity link). One of the merits of this model class is the possibility of nesting alternative spectral estimation methods under the same likelihood-based framework, so that the selection of a particular parametric spectrum amounts to estimating the transformation parameter. We also show that the generalised cepstral coefficients are a one to one function of the inverse partial autocorrelations of the process, which can be used to evaluate the mutual information between the past and the future of the process.

Key words and phrases: Generalised Linear Models; Box-Cox link; Whittle Likelihood; Mutual Information.

[†]Dipartimento di Economia e Finanza, Università di Roma 'Tor Vergata', Via Columbia 2, 00133 Rome, Italy.
E-mail: tommaso.proietti@uniroma2.it.
[*]Dipartimento di Scienze Statistiche 'Paolo Fortunati', Via Belle Arti 41, 40126 Bologna, Italy.
E-mail: alessandra.luati@unibo.it.

1 Introduction

The analysis of stationary processes in the frequency domain has a long tradition in time series analysis; the spectral density provides the decomposition of the total variation of the process into the contribution of periodic components with different frequency, as well as a complete characterization of the serial correlation structure of the process, so that it contains all the information needed for linear prediction and interpolation.

Several methods are available for estimating the spectrum. One of the most popular is Whittle estimation of an autoregressive spectrum, which leads to the solution of an empirical Yule-Walker set of equations (see, for instance, Percival and Walden, 1993, chapter 9). Alternatively, a class of nonparametric estimates is obtained by taking the Fourier transform of a smoothed sample autocovariance function: if a truncated or rectangular smoothing kernel is applied to the autocovariances, this is equivalent to fitting a finite order moving average model by the method of moments. A third popular approach is the exponential model proposed by Bloomfield (1973). The exponential model emerges by truncating the Fourier series expansion of the log-spectrum. The coefficients of the expansion are known as the cepstral coefficients and are in turn obtained from the discrete Fourier transform of the log-spectrum; their collection forms the cepstrum. This terminology was introduced by Bogert, Healy and Tuckey (1963), cepstral and cepstrum being anagrams of spectral and spectrum, respectively. The exponential model is a generalised linear model for observations (the periodogram), asymptotically distributed as an exponential random variable, adopting a logarithmic link for the mean, which is the spectral density itself.

This paper introduces the class of generalised linear models with Box-Cox link, according to which a linear model is formulated for the Box-Cox transformation of the spectral density. The link function depends on a power transformation parameter, and encompasses the exponential model, which corresponds to the case when the transformation parameter is equal to zero. Other important special cases are the inverse link, which leads to modelling the inverse spectrum and, in our setting, is equivalent to autoregressive estimation of the spectrum, and the identity link, which amounts to fitting a moving average model. The idea of developing power or cepstral correlation analysis as a direction for time series analysis dates back to Parzen (1992) and, in the context of speech recognition, to Kobayashi and Imai (1984). The rationale is finding a scale along which the transformed spectrum has a representation as a finite trigonometric polynomial of small order.

The coefficients of the trigonometric polynomial are related to the generalised autocovariances (Proietti and Luati, 2015) and are termed generalised cepstral coefficients. To enforce the constraints needed to guarantee the positivity of the spectral density, we propose a reparameterization of the model, based on a set of generalised inverse partial autocorrelations, which also provide a way of estimating the mutual information between past and future of a random process.

The empirical applications, dealing with the Southern Oscillation Index time series, as well as a Monte

Carlo simulation from a process characterised by a high dynamic range, illustrate the flexibility of the class of generalised linear spectral models and show that the appropriate spectral estimation model (exponential, autoregressive, moving average, etc.) can be selected in a likelihood based framework.

The paper is structured as follows: section 2 introduces the class of generalised linear cepstral models and discusses its time series properties. Section 3 deals with the estimation based on the maximisation of the Whittle likelihood and derives the asymptotic properties of the corresponding estimator. Illustrations are provided in section 4. Finally, in section 5, we outline some conclusions and directions for further research.

2 Generalised Linear Cepstral Models for the Spectrum

Let $\{y_t\}_{t\in T}$ be a stationary zero-mean real-valued stochastic process, $T = \{0, \pm 1, \pm 2, \dots\}$, with covariance function $\gamma_k = \int_{-\pi}^{\pi} e^{\imath\omega k} dF(\omega)$, $k = 0, \pm 1, \pm 2, \dots$ where $F(\omega)$ is the spectral distribution function of the process and $\imath$ is the imaginary unit. We assume that the spectral density function of the process exists, $F(\omega) = \int_{-\pi}^{\omega} f(\lambda)\mathrm{d}\lambda$, that the process is regular (Doob, 1953, p. 564), i.e. $\int_{-\pi}^{\pi} \ln f(\omega)\mathrm{d}\omega > -\infty$ and that $\int_{-\pi}^{\pi} f(\omega)^{\lambda}\, \mathrm{d}\omega < \infty$ for all $\lambda \in \mathbb{R}$. We let $\Gamma_n = \{\gamma_{|s-t|}, s, t = 1, \dots, n\}$ denote the autocovariance matrix of y_t of order n (a Toeplitz matrix). The cepstrum of the process (Bogert, Healy and Tukey, 1963) is defined as the sequence of cepstral coefficients

$$c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln[2\pi f(\omega)]\cos(\omega k)\mathrm{d}\omega, k = 0, 1, \dots \tag{1}$$

which characterise the Fourier series: $\ln[2\pi f(\omega)] = c_0 + 2\sum_{j=1}^{\infty} c_j \cos(\omega j)$. The Wold representation of $\{y_t\}_{t\in T}$ is written as $y_t = \psi(B)\xi_t$, where $\{\xi_t\}_{t\in T}$ is a white noise process with zero mean and finite variance σ^2, $\xi_t \sim \mathrm{WN}(0, \sigma^2)$, so that $2\pi f(\omega) = \sigma^2|\psi(e^{-\imath\omega})|^2$ where $\psi(z) = 1 + \psi_1 z + \psi_2 z^2 + \dots$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$.

Let us consider the Box-Cox transform of the spectral generating function $2\pi f(\omega)$, with transformation parameter $\lambda \in \mathbb{R}$,

$$g_\lambda(\omega) = \begin{cases} \frac{[2\pi f(\omega)]^{\lambda} - 1}{\lambda}, & \lambda \neq 0, \\ \ln[2\pi f(\omega)], & \lambda = 0. \end{cases}$$

We assume that $g_\lambda(\omega)$ can be represented by a finite trigonometric polynomial:

$$g_\lambda(\omega) = c_{\lambda,0} + 2\sum_{k=1}^{K} c_{\lambda,k}\cos(\omega k). \tag{2}$$

The representation (2) is linear in the coefficients $c_{\lambda k}$, the inverse Fourier tranform of the $g_\lambda(\omega)$ function:

$$c_{\lambda k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g_\lambda(\omega)\cos(\omega k)\mathrm{d}\omega, \quad k = 0, 1, \dots, K.$$

When λ equals 0, we obtain Bloomfield (1973) exponential model as a special case. For $k = 1, \dots K$, the coefficients c_{0k} are equal to the cepstral coefficients given in (1). We will henceforth refer to (2) as the generalised cepstral coefficient at lag k, and to $\{c_{\lambda k}, k = 1, \dots, K\}$ as the generalised cepstrum.

The spectral model with Box-Cox link and mean function

$$f(\omega) = \begin{cases} \frac{1}{2\pi}\left[1+\lambda g_\lambda(\omega)\right]^{\frac{1}{\lambda}}, & \lambda \neq 0, \\ \frac{1}{2\pi}\exp[g_\lambda(\omega)], & \lambda = 0, \end{cases} \tag{3}$$

will be referred to as a generalised linear cepstral model for the spectrum with parameter λ and order K.

Remark 1 The rationale of generalised linear cepstral models is that there exists a value of the transformation parameter λ, such that $g_\lambda(\omega)$ can be expressed in terms of a finite and parsimonious set of Fourier coefficients. The next subsection shows that $\{c_{\lambda k}, k = 1, \ldots, K\}$ contain all the necessary information for prediction and feature extraction of y_t.

Remark 2 The dynamic range of the spectrum is defined as $10 \log_{10}((\max_\omega f(\omega))/(\min_\omega f(\omega)))$, see Percival and Walden (1993, section 6.3). It is related in the limit to the condition number of the autocovariance matrix Γ_n, as n increases. The objective of the power transform of the spectrum is to achieve a low dynamic range for $[2\pi f(\omega)]^\lambda$, so that using only a few terms of the Fourier expansion, as in (2), is suitable. When the process is characterised by a high dynamic range, due to the presence of spectral peaks, a negative transformation parameter provides a parsimonious model.

2.1 Time series properties

It is immediate from (3) that λ and the corresponding generalised cepstrum $\{c_{\lambda k}, k = 1, \ldots, K\}$ uniquely characterise the spectral properties of the random process $\{y_t\}_{t \in T}$.

For $\lambda = 0$, $c_{0,0} = \ln \sigma^2$, from the Szegö-Kolmogorov formula for the prediction error variance,

$$\sigma^2 = \exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln[2\pi f(\omega)]\mathrm{d}\omega\right\};$$

the coefficients of the Wold representation are obtained by the recursive formula $\psi_j = j^{-1}\sum_{r=1}^{j} r c_{0r}\psi_{j-r}$, $j = 1, 2, \ldots, \psi_0 = 1$, see e.g. Pourahmadi (1983).

For $\lambda \neq 0$, the generalised cepstral coefficients $c_{\lambda k}$ are related to the generalised autocovariance function, introduced by Proietti and Luati (2015),

$$\gamma_{\lambda k} = \frac{1}{2\pi}\int_{-\pi}^{\pi} [2\pi f(\omega)]^\lambda \cos(\omega k)\mathrm{d}\omega, \tag{4}$$

by the following relationships:

$$c_{\lambda 0} = \frac{1}{\lambda}(\gamma_{\lambda 0} - 1), \quad c_{\lambda k} = \frac{1}{\lambda}\gamma_{\lambda k}, \quad k \neq 0. \tag{5}$$

The coefficients of the Wold and autoregressive representation of the process can be obtained by a spectral factorisation that naturally arises after a reparameterisation of the generalised cepstral coefficients. In fact, $[2\pi f(\omega)]^\lambda = 1 + \lambda g_\lambda(\omega) = \gamma_{\lambda 0} + 2\sum_{k=1}^K \gamma_{\lambda k}\cos(\omega k)$, so that, by the representation theorem of Fejér and Riesz for nonnegative trigonometric polynomials (Grenander and Szegö, 1958, p. 20-21), we can write:

$$[2\pi f(\omega)]^\lambda = \sigma_\lambda^2 b_\lambda(e^{-\imath\omega}) b_\lambda(e^{\imath\omega}), \quad b_\lambda(e^{-\imath\omega}) = 1 + b_{\lambda 1}e^{-\imath\omega} + \cdots + b_{\lambda K}e^{-\imath\omega K}. \tag{6}$$

According to (6), when $\lambda \neq 0$ the generalised cepstral coefficients are obtained as

$$c_{\lambda 0} = \frac{1}{\lambda}\left[\sigma_\lambda^2(1 + b_{\lambda 1}^2 + \cdots + b_{\lambda K}^2) - 1\right], \quad c_{\lambda k} = \frac{1}{\lambda}\sigma_\lambda^2 \sum_{j=k}^K b_{\lambda j} b_{\lambda, j-k}. \tag{7}$$

Moreover, $\sigma_\lambda^{2/\lambda}$ is the prediction error variance of the process, as

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \ln\left(2\pi f(\omega)\right) \mathrm{d}\omega = \frac{1}{\lambda}\ln \sigma_\lambda^2.$$

The coefficients $b_{\lambda k}$ can be uniquely determined by imposing the condition that the roots of the polynomial $b_\lambda(z) = 1 + b_{\lambda,1}z + \cdots + b_{\lambda K}z^K$ lie outside the unit circle, which in turn can be enforced by adopting a reparameterization due to Barndorff-Nielsen and Schou (1973) and Monahan (1984). Given K coefficients $\varsigma_{\lambda k}$, $|\varsigma_{\lambda k}| < 1$, $k = 1, \ldots, K$, that in the present setting are interpretable as generalised partial inverse autocorrelations (Bhansali, 1983), the coefficients of the polynomial $b_\lambda(z)$ are obtained from the last iteration of the Durbin-Levinson recursion

$$b_{\lambda j}^{(k)} = b_{\lambda j}^{(k-1)} + \varsigma_{\lambda k} b_{\lambda, k-j}^{(k-1)}, \quad b_{\lambda k}^{(k)} = \varsigma_{\lambda k}, \tag{8}$$

for $k = 1, \ldots, K$, and $j = 1, \ldots, k-1$, so that $b_{\lambda j} = b_{\lambda j}^{(K)}$. The role of the coefficients $\varsigma_{\lambda k}$ shall be further discussed in subsection 2.2.

The coefficients of the Wold representation can be obtained recursively as follows (Gould, 1974):

$$\psi_j = j^{-1}\sum_{r=1}^{j}\left(r\frac{\lambda+1}{\lambda} - j\right) b_{\lambda j}\psi_{j-r}, \quad j > 0, \quad \psi_0 = 1.$$

The coefficients of the infinite AR representation may be similarly derived. In sum, all the relevant information for prediction is available from $K+1$ bits of information.

Remark 3 For $\lambda = 1$, $c_{1k} = \gamma_k = \mathrm{E}(y_t y_{t-k}), k > 0$, the autocovariance function of the process is obtained. In the case $\lambda = -1$ and $k \neq 0$, $c_{-1,k} = -\gamma i_k$, where γi_k is the inverse autocovariance of y_t (Cleveland, 1972). The intercept $c_{\lambda 0}$ for $\lambda = -1, 0, 1$, is related to important characteristics of the stochastic process, as $1/(1 - c_{-1,0})$ is the interpolation error variance, $\exp(c_{0,0}) = \sigma^2$, the prediction error variance, and $c_{1,0} + 1 = \gamma_0$ is the unconditional variance of y_t.

Remark 4 Fractionally integrated processes (Giraitis et al., 2012) arise as limiting cases. Consider for instance the case $b_\lambda(z) = 1 - z$, so that $[2\pi f(\omega)]^\lambda = \sigma_\lambda^2|1 - e^{-\imath\omega}|^2$ is the spectrum of a first order non-invertible moving average process. If $\lambda = -d^{-1}, d \in (0, 0.5)$, then a fractionally integrated process arises, whose spectral density is unbounded at the origin: $f(\omega) = \frac{\sigma^2}{2\pi}|2\sin(\omega/2)|^{-2d}$. More generally, if $\lambda < -2$ and $b_\lambda(z)$ can be factorised as $b_\lambda(z) = (1-z)b_\lambda^*(z)$, with $b_\lambda^*(z) = 1 + b_{\lambda,1}^* z + \cdots + b_{\lambda,K-1}^* z^{K-1}$, $b_\lambda^*(z) \neq 0 \iff |z| \leq 1$, the process is fractionally integrated of order $d = -\lambda^{-1}$. When $b_\lambda(z) = 1 - 2z + z^2$ and $\lambda = -2d^{-1}$, the process is a Gegenbauer process; see Hosking (1981) and Gray, Zhang and Woodward (1989).

2.2 The mutual information between past and future

The following theorem expresses the strong Szegö theorem (see Bingham, 2012, Theorem 6, and the references therein) in terms of generalised partial inverse autocorrelations introduced in (8).

Theorem 1 *Let $\{y_t\}_{t\in T}$ be a purely non deterministic Gaussian process with cepstral coefficients c_j, $j = 1, \ldots, \infty$, and generalised cepstral coefficients $c_{\lambda k}$, $k = 1, \ldots, K$. Then, for $\lambda \neq 0$,*

$$\sum_{j=1}^{\infty} j c_j^2 = \frac{1}{\lambda^2}\sum_{k=1}^{K} k \ln(1 - \varsigma_{\lambda k}^2) \tag{9}$$

where $\varsigma_{\lambda k}$ are the generalised partial inverse autocorrelations of the process and are related to the generalised cepstral coefficients by equations (7) and (8).

The term $\sum jc_j^2$ has several important uses. The mutual information between past and future of a Gaussian process is defined as

$$I_{p-f} = \frac{1}{2}\sum_{j=1}^{\infty} j c_j^2. \tag{10}$$

I_{p-f} measures the reduction in the uncertainty of about the future $\mathcal{F}_t$, the sigma-algebra generated by $\{y_{t+j}, j = 1, 2, \ldots,\}$, when the past $\mathcal{P}_t$, the sigma-algebra generated by $\{y_{t-j}, j = 1, 2, \ldots,\}$, is known (Ibragimov and Rozanov, 1978, Li, 2005). $I_{p-f} = 0$ for a Gaussian white noise process and $I_{p-f} < \infty$ for an absolutely regular process (Ibragimov and Rozanov, 1978, chapter IV). Theorem 1 also shows that I_{p-f} is infinite if $|\varsigma_{\lambda k}| = 1$, for some k, which occurs in the case of fractionally integrated processes (see Remark 4). Finally, according to the strong Szegö limit theorem (see Bingham, p. 305), $\sum jc_j^2$ is the limit of $\left[\ln|\Gamma_n| - n\ln\sigma^2\right]$ as $n \to \infty$, that can be used to approximate the log-determinantal term of the Gaussian likelihood of the sample time series $\{y_t, t = 1, \ldots, n\}$. The evaluation of $\sum jc_j^2$ is not trivial, except for the class of generalised cepstral models, for which Theorem 1 states that it can be computed from the K inverse partial autocorrelation coefficients as in (9).

3 Whittle likelihood estimation

The main tool for estimating the spectral density function and its functionals is the periodogram. Let $\{y_t, t = 1, 2, \dots, n\}$ denote a time series, which is assumed to be a sample realisation from a stationary short memory Gaussian process, characterised by an autocovariance sequence satisfying $\sum_{k=1}^{\infty} k\gamma_k^2 < \infty$, and let $\omega_j = \frac{2\pi j}{n}$, $j = 1, \dots, [n/2]$, denote the Fourier frequencies, where $[\cdot]$ denotes the integer part of the argument. The periodogram, or sample spectrum, is defined as

$$I(\omega_j) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n}(y_t - \bar{y})e^{-\imath\omega_j t}\right|^2,$$

where $\bar{y} = n^{-1}\sum_{t=1}^{n} y_t$. In large samples (Brockwell and Davis, 1991, ch. 10)

$$\frac{I(\omega_j)}{f(\omega_j)} \sim \text{IID}\ \frac{1}{2}\chi_2^2, \qquad \omega_j = \frac{2\pi j}{n},\ j = 1, \dots, [(n-1)/2], \tag{11}$$

whereas $\frac{I(\omega_j)}{f(\omega_j)} \sim \chi_1^2, \omega_j = 0, \pi$, where χ_m^2 denotes a chi-square random variable with m degrees of freedom, and, as a particular case, $\frac{1}{2}\chi_2^2$ is an exponential random variable with unit mean.

The above distributional results are the basis for approximate or Whittle maximum likelihood inference for the generalised cepstral model: writing $f(\omega)$ as in (3), and denoting by $\theta_\lambda = [c_{\lambda 0}, c_{\lambda 1}, \dots, c_{\lambda K}]'$ the vector containing the generalised cepstral coefficients, where $\theta_\lambda \in \Theta \subset \mathbb{R}^{K+1}$, the log-likelihood of $\{I(\omega_j), j = 1, \dots, N = [(n-1)/2]\}$, is:

$$\ell(\lambda, \theta_\lambda) = -\sum_{j=1}^{N}\left[\ln f(\omega_j) + \frac{I(\omega_j)}{f(\omega_j)}\right]. \tag{12}$$

Letting $z(\omega) = [1, 2\cos(\omega), 2\cos(2\omega), \dots, 2\cos(K\omega)]'$, and writing $g_\lambda(\omega_j) = z(\omega_j)'\theta_\lambda$, (12) can be expressed as $\ell(\lambda, \theta_\lambda) = -\sum_{j=1}^{N}\ell_j(\lambda, \theta_\lambda)$, where, for $1 + \lambda z(\omega_j)'\theta_\lambda > 0$,

$$\ell_j(\lambda, \theta_\lambda) = \begin{cases} \frac{1}{\lambda}\ln\left(1 + \lambda z(\omega_j)'\theta_\lambda\right) + \frac{2\pi I(\omega_j)}{(1+\lambda z(\omega_j)'\theta_\lambda)^{\frac{1}{\lambda}}}, & \lambda \neq 0, \\ z(\omega_j)'\theta_0 + \frac{2\pi I(\omega_j)}{\exp(z(\omega_j)'\theta_0)}, & \lambda = 0. \end{cases}$$

Notice that we have excluded the frequencies $\omega = 0, \pi$ from the analysis; the latter may be included with little effort, and their effect on the inferences is negligible in large samples.

An alternative derivation of the Whittle likelihood is based on the following argument. Let $y = (y_1, \dots y_n)' \sim \mathrm{N}(0, \Gamma_n)$, an asymptotic approximation to the true Gaussian log-likelihood,

$$\ell^*(\lambda, \theta_\lambda) = -\frac{n}{2}\ln 2\pi - \frac{1}{2}\ln|\Gamma_n| - \frac{1}{2}y'\Gamma_n^{-1}y, \tag{13}$$

is, apart from a constant:

$$\ell^\dagger(\lambda, \theta_\lambda) = -\frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\ln f(\omega) + \frac{I(\omega)}{f(\omega)}\right]\mathrm{d}\omega, \tag{14}$$

so that $\ell(\lambda, \theta_\lambda)/N$, where $\ell(\lambda, \theta_\lambda)$ is given by (12), converges in probability to $\ell^*(\lambda, \theta_\lambda)$.

The profile likelihood of the model as λ varies can be used to select the spectral model for y_t. A similar idea has been used by Koenker and Yoon (2009) for the selection of the appropriate link function for binomial data. Let $\ell_{max}(\lambda)$ denote the partially maximised, or profile, Whittle likelihood $\ell_{max}(\lambda) = \max_{\theta_\lambda \in \Theta} \ell(\lambda, \theta_\lambda)$, or equivalently $\ell_{max}(\lambda) = \ell(\lambda, \tilde{\theta}_\lambda)$, where $\tilde{\theta}_\lambda = \text{argmax}_{\theta_\lambda \in \Theta} \ell(\lambda, \theta_\lambda)$. The maximum likelihood estimate of the transformation parameter is obtained as the value of λ which maximises the profile Whittle likelihood.

The truncation parameter, K, can be chosen as the value minimizing an information criterion, such as the Akaike Information Criterion (AIC) or the Bayesian Information Criterion (BIC), given, respectively, by:

$$AIC(K, \lambda) = -2\ell(\lambda, \tilde{\theta}_\lambda) + 2K, \quad BIC(K, \lambda) = -2\ell(\lambda, \tilde{\theta}_\lambda) + \ln(N)K. \tag{15}$$

Remark 5 In the long memory case, the above distributional results cannot be invoked, as the large sample distribution of the periodogram (normalized by dividing for the spectral density) is no longer IID exponential in the vicinity of the zero frequency, see Künsch (1986), Hurvich and Beltrao (1993) and Robinson (1995). However, the Whittle estimator of the unknown parameters, i.e. the maximiser of (12) can be shown to be consistent and asymptotically normal also in the long memory case. See Dahlhaus (1989) for the general theory.

3.1 Reparameterization

The main difficulty with maximum likelihood estimation of the generalised linear cepstral model in the case $\lambda \neq 0$ is enforcing the condition $1 + \lambda z(\omega_j)'\theta_\lambda > 0$. This problem is well known in the literature concerning generalised linear models for gamma distributed observations, for which the canonical link is the inverse link (McCullagh and Nelder, 1989).

The most appropriate solution that ensures the positive definiteness and the regularity of the spectral density is to reparameterise the generalised cepstral coefficients as in (7). For $k = 1, \dots, K$, the parameters $b_{\lambda k}$ are in turn expressed as a function of the the generalised inverse partial autocorrelations $|\varsigma_{\lambda k}| < 1$. In practice, we estimate K unconstrained real parameters $\vartheta_{\lambda k}$, that are mapped into the interval $(-1, 1)$ by the Fisher inverse transformations $\varsigma_{\lambda k} = \frac{\exp(2\vartheta_{\lambda k}) - 1}{\exp(2\vartheta_{\lambda k}) + 1}$ for $j = 1, \dots, K$. Also, we set $\vartheta_{\lambda 0} = \ln(\sigma^2_\lambda)$.

3.2 Asymptotic properties

We hereby prove the consistency and the asymptotic normality of the Whittle MLE of the vector θ_λ, which we denoted as $\tilde{\theta}_\lambda$, under the hypothesis that λ is known and equal to its true value. In practice, the transformation parameter is estimated by maximising the profile likelihood $\ell_{max}(\lambda)$. In the sequel we focus separately on the cases $\lambda \neq 0$ and $\lambda = 0$.

When $\lambda \neq 0$, $[2\pi f(\omega)]^\lambda = 1 + \lambda g_\lambda(\omega)$ and $g_\lambda(\omega) = z(\omega)'\theta_\lambda$. The asymptotic theory for the Whittle MLE of θ_λ is based on the following assumptions.

A1. $\{y_t\}_{t\in\mathcal{T}}$ is a zero mean stationary Gaussian process and $\exists\, m > 0$ such that $1 + \lambda g_\lambda(\omega) \geq m$.

A2. $\theta_{\lambda 0} \in \Theta \subset \mathbb{R}^{K+1}$ where $\theta_{\lambda 0}$ is the true parameter value and Θ is a closed subset of the $K+1$ Euclidean space.

A3. The Fourier series expansion $\ln(1 + \lambda g_\lambda(\omega)) = \kappa_0 + 2\sum_{j=1}^{\infty}\kappa_j \cos(\omega j)$ has coefficients satisfying the condition $\sum_{j=1}^{\infty} j\kappa_j^2 < \infty$, or, equivalently, the cepstral coefficients in $\ln[2\pi f(\omega)] = \sum_j c_j e^{-\imath\omega j}$ satisfy $\sum_{j=1}^{\infty} j c_j^2 < \infty$.

Theorem 2 *Under conditions A1-A3,*

$$\tilde{\theta}_\lambda \rightarrow_p \theta_{\lambda 0},$$

$$\sqrt{n}(\tilde{\theta_\lambda} - \theta_{\lambda 0}) \rightarrow_d N(0, V_\lambda),$$

with

$$V_\lambda^{-1} = \frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{1}{[2\pi f(\omega)]^{2\lambda}} z(\omega)z(\omega)'\mathrm{d}\omega.$$

In the exponential case, when $\lambda = 0$, the finitess of $\sum_{j=1}^{K} j c_j^2$ implies that $\sum_{j=1}^{\infty} j\gamma_j^2$ and the remaining conditions of Theorems II.2.1 and II.2.2 in Dzhaparidze (1986) are fulfilled. Hence, under assumption A2 the Whittle estimates of the cepstral coefficients are consistent and $\sqrt{n}(\tilde{\theta_\lambda} - \theta_{\lambda 0}) \rightarrow_d \mathrm{N}(0, V_\lambda)$, with $V_\lambda^{-1} = \frac{1}{4\pi}\int_{-\pi}^{\pi} z(\omega)z(\omega)'\mathrm{d}\omega$. See Bloomfield (1973).

Remark 6 The function that maps the partial autocorrelation coefficients to the model parameters is one to one and smooth (see Barndorff-Nielsen and Schou, 1973, Theorem 2), so that the asymptotic properties of the Whittle estimator continue to hold.

4 Illustrations

The proposed generalizations will now be applied to time series that have been analyzed extensively in the literature and that provide a useful testbed for the class of generalised linear spectral models. In the applications we will use the tapered periodogram $I(\omega_j) = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} h_t(y_t - \bar{y})e^{-\imath\omega_j t}\right|^2$, where $h_t \geq 0, t = 1, \ldots, n$, and $\sum_{t=1}^{n} h_t^2 = 1$. We shall use a taper formed for zeroth-order discrete prolate spheroidal sequences (DPSS), obtained as the eigenvector corresponding to the largest eigenvalue of the matrix $A(\nu)$, with elements

$$a_{ij}(\nu) = \begin{cases} \frac{\sin(2\pi\nu(i-j))}{\pi(i-j)}, & \text{for } i \neq j, \\ 2\nu, & \text{for } i = j, \end{cases}$$

$i, j = 1, \ldots, n$. The matrix depends on the bandwidth parameter ν, which is often set equal to $\nu = 2/n$. We refer to Percival and Walden (1993, sec. 3.9 and ch. 7) for further details. Brillinger (1981, Theorem 5.2.7) shows that for the tapered periodogram the same distributional result as (11) holds.

In the sequel we refer to the generalised linear cepstral model with transformation paramater λ and order K as GLCM(λ, K).

4.1 Southern Oscillation Index

The Southern Oscillation Index (SOI) measures the difference in surface air pressure between Tahiti and Darwin and it is an important indicator of the strength of El Niño and La Niña events, with values below -8 indicating an El Niño event while positive values above 8 indicate a La Niña event. The index reflects the cyclic warming (negative SOI) and cooling (positive SOI) of the eastern and central Pacific, which affects the sea level pressure at the two locations. The monthly series from January 1876 to December 2013 is plotted in figure 1 along with the autocorrelation function. The series has a periodic behaviour: often the El Niño and La Niña episodes alternate and this confers the SOI a cyclical feature, with an irregular period of about 3-7 years (see e.g. http://earthobservatory.nasa.gov).

We investigate what GLCM(λ, K) representation provides the best fit to the sample spectrum of the time series. This depends on two crucial parameters, the truncation parameter K and the power parameter λ, which can be selected according to the information criteria given in (15). Estimating the GLCM(λ, K) on a grid of values for λ in the range [-2.50, 1.00] with step 0.01 and for K ranging from 0 to 10, and computing the AIC and BIC criteria, leads to selecting $K = 7$ and $\tilde{\lambda} = -2.28$. Figure 2 displays the prediction error variance and the profile Whittle likelihood of GLCM($\lambda, 7$) models, as a function of λ, which also shows that the optimal value of the power transformation parameter is $\tilde{\lambda} = -2.28$. The third plot displays the corresponding estimates of I_{p-f}, which peaks at around $\lambda = -2$. The plot illustrates that the pair (λ, K) that minimises the AIC, does not necessarily maximises I_{p-f}.

The estimated spectrum is $\tilde{f}(\omega) = \frac{1}{2\pi}\left[\tilde{\sigma}^2_\lambda \tilde{b}_\lambda(e^{-\imath\omega})\tilde{b}_\lambda(e^{\imath\omega})\right]^{-1/2.278}$, with $\tilde{\sigma}^2_\lambda = 2161.74$, and

$$\tilde{b}_\lambda(e^{-\imath\omega}) = 1 - 1.02e^{-\imath\omega} - .03e^{-\imath 2\omega} - .05e^{-\imath 3\omega} - .08e^{-\imath 4\omega} + .04e^{-\imath 5\omega} - .02e^{-\imath 6\omega} + .23e^{-\imath 7\omega}$$

From the second panel of figure 2 it is evident that the likelihood ratio test of $\lambda = -2$ for a GLCM(λ, K) model with $K = 7$ is not significant, so that the spectrum that is fitted by maximum likelihood does not differ from that arising from fitting an autoregressive model of order 14 such that the autoregressive polynomial

Figure 1: Southern Oscillation Index. Time series and sample autocorrelation function. In the first plot the horizontal lines are drawn at ± 8,

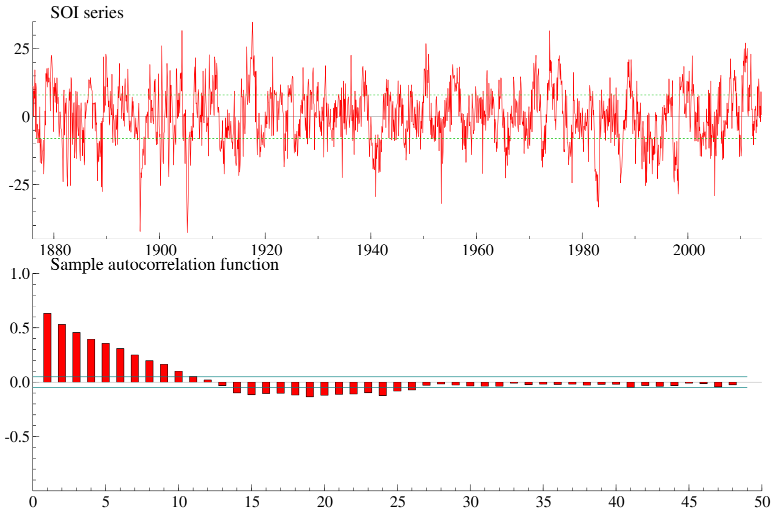


Figure 2: Southern Oscillation Index. Whittle likelihood, Prediction error variance and Mutual Information as a function of λ for GLMS(λ, K) models with $K = 7$.

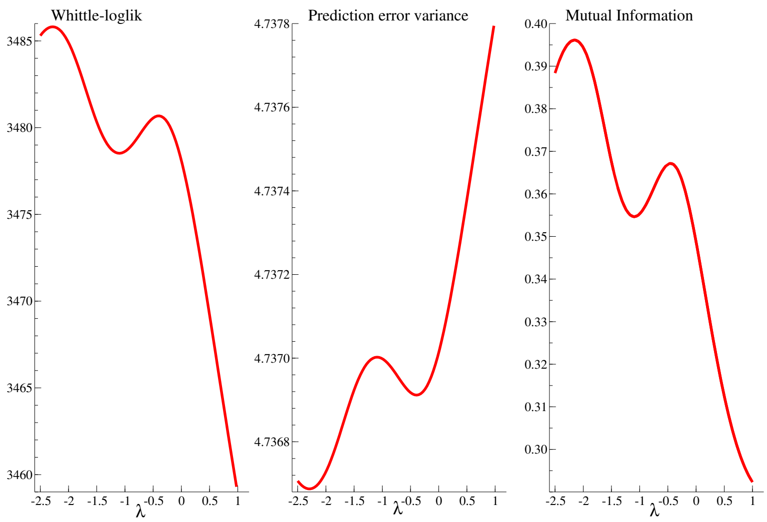
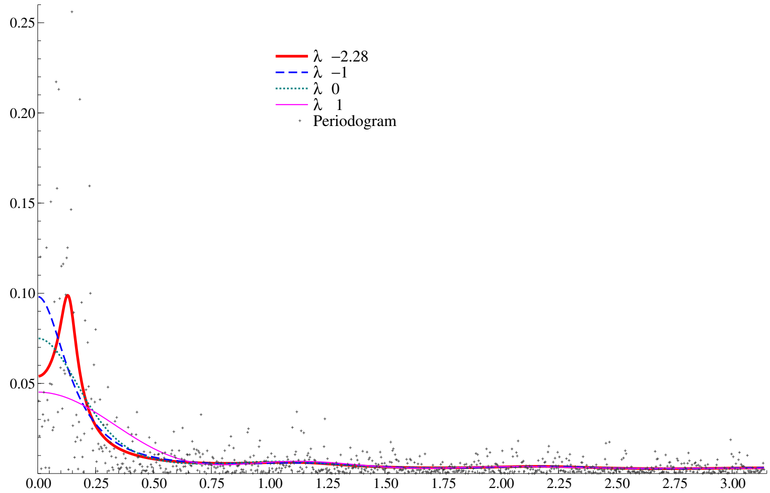


Figure 3: Southern Oscillation Index. Comparison of the spectral density estimates arising from different GLCM(λ, K) models with $K = 7$.



is the square of a polynomial of order 7. This polynomial has three pairs of complex conjugate roots and a real root.

Figure 3 plots the periodogram of the SOI series and overimposes the spectral densities fitted by the GLCM(λ, K) model with $K = 7$ and λ set equal to $1, 0, -1$ and $\tilde{\lambda} = -2.28$. The case when λ is set equal to 1 corresponds to fitting an MA(7) model to the series, whereas the case $\lambda = 0$ corresponds to fitting Bloomfield's exponential model of order $K = 7$; $\lambda = -1$ corresponds to fitting an AR(7). It should be noticed that in none of these cases a spectral peak arises at a frequency other than zero. The spectrum fitted by maximum likelihood, on the contrary has a clear mode at a frequency corresponding to a period of about four years.

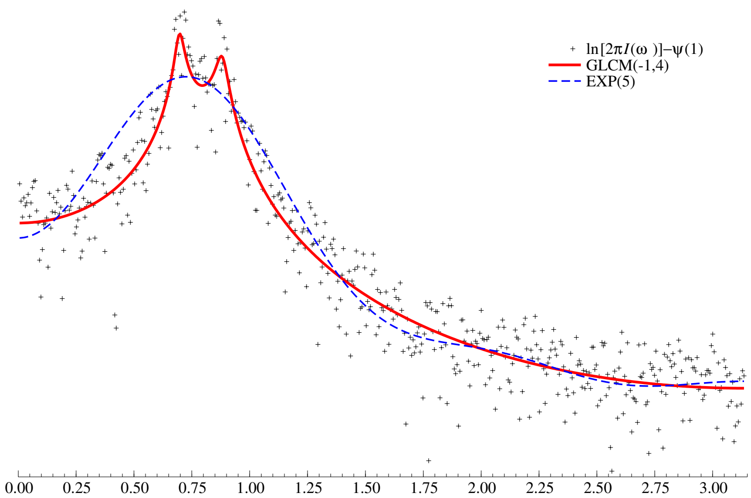
4.2 Simulated AR(4) Process

This illustration concerns the estimation of the power spectrum of the AR(4) stochastic process

$$y_t = 2.7607 y_{t-1} - 3.8106 y_{t-2} + 2.6535 y_{t-3} - 0.9238 y_{t-4} + \xi_t, \xi_t \sim \text{NID}(0, 1). \quad (16)$$

The interest in this application lies in the nature of the spectral density to be estimated, which is bimodal, featuring two peaks located very closely in the frequency range. In fact, the AR polynomial features two pairs of complex conjugate roots with modulus 1.01 and 1.02 and phases 0.69 and 0.88, respectively. The closeness of the two modes renders the estimation of the spectrum rather problematic, and thus this process constitutes a test case for spectral estimation methods; see Percival and Walden (1993).

Figure 4: Periodogram and log-spectra estimated by the GLCM$(-1, 4)$, selected by BIC, and the exponential model with $K = 5$, EXP(5).



A sample time series of length $n = 1024$ from the above AR(4) process is obtained from the book by Percival and Walden (1993) and its detailed analysis is presented here for illustrative purposes.

The specifications of the class GLCM(λ, K) selected by AIC and BIC differ slightly. While the latter selects the true generating model, that is $\lambda = -1$ and $K = 4$, AIC selects $\lambda = -1$ and $K = 6$. However, the likelihood ratio test of the null that $K = 4$ is a mere 4.8.

The estimated coefficients of the GLCM$(-1, 4)$ model and their estimation standard errors are

$b_{\lambda k}$	std. err.	true value
-2.7490	0.0007	-2.7607
3.7901	0.0016	3.8106
-2.6353	0.0007	-2.6535
0.9201	0.0025	0.9238

The comparison with the true autoregressive coefficients (reported in the last column) stresses that the coefficient estimates are remarkably accurate. Figure 4 displays the centered periodogram and compares the log-spectra fitted by the selected GLCM$(-1, 4)$ model and the exponential model with $K = 5$, which emerges if Box-Cox transformation parameter is set equal to zero. The latter fit is clearly suboptimal, as it fails to capture the two spectral modes.

In order to evaluate whether the above result are generalisable, we have carried out a Monte Carlo experiment by which 5,000 replications of length $n = 1024$ are generated independently according to the

AR(4) Gaussian process (16). For each replication we estimate a sequence of GLCM(λ, K) models with λ taking values on a grid from -1.75 to 1 with step 0.05 and for K ranging from 0 to 8; model selection is carried out according to AIC and BIC and the maximum likelihood estimates of the parameters, $\tilde{\sigma}^2_\lambda$ and $\tilde{b}_k$, $k = 1, \ldots, K$, are recorded, as well as the estimated spectral density.

The main results are presented in table 1, which reports some summary statistics of the distribution of the maximum likelihood estimates of the GLCM(λ, K) models, separately for the two cases when model selection is carried out according to the AIC and the BIC, namely the median, the mean, the standard deviation (St. dev.) and the root mean square estimation error (RMSE). The main evidence can be summarised as follows. Model selection by BIC yields more accurate estimates of the parameters and the order of the model; the RMSE is systematically smaller and both the variance and the bias of the estimator are smaller. The distribution of the selection frequency for the order of the GLCM(λ, K) is the following

K	0	1	2	3	4	5	6	7	8
AIC	0.00	0.00	0.00	0.00	0.36	0.17	0.14	0.16	0.17
BIC	0.00	0.00	0.00	0.00	0.78	0.16	0.04	0.02	0.01

Finally, the standard error in the estimation of the log-spectrum, computed by averaging across the simulations,

$$\left[\frac{1}{N}\sum_{j=1}^{N}\Big(\ln \tilde{f}(\omega_j) - \ln f(\omega_j)\Big)^2\right]^{1/2},$$

where $\tilde{f}(\omega_j)$ is the fitted spectral density, equals 0.1788 and 0.1520, respectively when the model is selected by AIC and BIC; this entails that AIC results in an efficiency loss of about 18% in estimating the log-spectrum of the series.

The overall conclusion is that, despite the differences due to the adoption of a selection criterion, the GLCM(λ, K) proves a very effective spectral estimation method, yielding an autoregressive spectral estimate ($\lambda = -1$) or a neighbouring estimate in all cases.

5 Conclusions

We have proposed a general frequency domain estimation framework which nests the exponential model for the spectrum as a special case and allows for any power transformation of the spectrum to be modelled, so that alternative spectral fits can be encompassed. As a direction for future research we think that the generalised linear cepstral model can have successful applications for modelling the time-varying spectrum of a locally stationary processes (Dahlhaus, 2012), by allowing the cepstral coefficients to vary over time, e.g. with autoregressive dynamics. Finally, a multivariate extension, the matrix-logarithmic spectral model

Table 1: Summary statistics of the Monte Carlo distribution of the maximum likelihood estimates of the parameters of the GLCM(λ, K) model, based on 5,000 replications from the AR(4) model (16).

	AIC				BIC				True
Parameter	Median	Mean	St. dev.	RMSE	Median	Mean	St. dev.	RMSE	values
$\tilde{\lambda}$	-1.0000	-1.0242	0.0903	0.0242	-1.0000	-1.0016	0.0305	0.0016	-1
$\tilde{K}$	5.0000	5.5968	1.5177	1.5968	4.0000	4.3072	0.6752	0.3072	4
$\tilde{b}_1$	-2.7581	-2.8169	0.2783	0.0562	-2.7554	-2.7579	0.0888	0.0028	-2.7607
$\tilde{b}_2$	3.8026	4.0037	0.7305	0.1931	3.7974	3.8071	0.2496	0.0035	3.8106
$\tilde{b}_3$	-2.6486	-2.9514	1.0288	0.2979	-2.6388	-2.6576	0.3447	0.0041	-2.6535
$\tilde{b}_4$	0.9224	1.1586	0.7674	0.2348	0.9176	0.9337	0.2547	0.0099	0.9238
$\tilde{b}_5$	0.0000	-0.0847	0.3370	0.0847	0.0000	-0.0060	0.1228	0.0060	0
$\tilde{b}_6$	0.0000	-0.0015	0.1979	0.0015	0.0000	-0.0008	0.0770	0.0008	0
$\tilde{b}_7$	0.0000	0.0034	0.1089	0.0034	0.0000	0.0010	0.0361	0.0010	0
$\tilde{b}_8$	0.0000	0.0048	0.0391	0.0048	0.0000	0.0002	0.0089	0.0002	0
$\tilde{\sigma}_\lambda$	1.0161	1.5414	1.5363	0.5414	1.0044	1.0535	0.5061	0.0535	1

for the spectrum of a vector time series, could be envisaged, along the lines of the model formulated by Chiu, Leonard and Tsui (1996) for covariance structures.

Appendix

A Proof of theorem 1

The coefficients $\{\varsigma_{\lambda k}, k = 1, \ldots, K\}$ are the partial autocorrelations of the process $b_\lambda(B)x_t = \nu_t, \nu_t \sim \text{WN}(0, \sigma^2_\nu)$ (see Bhansali, 1983), whose spectrum $f_x(\omega)$ is proportional to $[f(\omega)]^{-\lambda}$. The process x_t has a cepstrum $\{\kappa_j, j = 0, 1, \ldots\}$,

$$\kappa_j = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln[2\pi f_x(\omega)] \cos(\omega j) \mathrm{d}\omega,$$

so that, by the strong Szegö theorem, see Bingham (2012, Theorem 6),

$$\prod_{k=1}^{K}(1 - \varsigma^2_{\lambda k})^{-k} = \exp\left(\sum_{j=1}^{\infty} j\kappa_j^2\right).$$

For $j > 0$, the cepstral coefficients of x_t are related to those of y_t by $\kappa_j = -\lambda c_j$, so that

$$\frac{1}{\lambda^2}\sum_{k=1}^{K} k \ln(1 - \varsigma^2_{\lambda k}) = \sum_{j=1}^{\infty} j c_j^2.$$

B Proof of theorem 2

Under the stated assumptions Theorems II.2.1 (consistency) and II.2.2 of Dzhaparidze (1986, chapter II, pp. 105 and 109, respectively) hold.

Condition A1 ensures that $1 + \lambda g_\lambda(\omega) = \gamma_{\lambda 0} + 2\sum_{k=1}^{K} \gamma_{\lambda k}\cos(\omega k)$ is positive. As a matter of fact, $1+\lambda g_\lambda(\omega)$ is a proper spectral density and under A1 the generalised autocovariances $\{\gamma_{\lambda k}\}$ form a positive definite sequence. Also, as $g_\lambda = z(\omega)'\theta_\lambda$, $1 + \lambda g_\lambda(\omega)$ is a smooth and symmetric function of $\omega \in [-\pi, \pi]$. Assumption A2 states that the true parameter vector is interior to the parameter space. Notice also that our GLCM is identified, that is $\theta_{\lambda 1} \neq \theta_{\lambda 2} \to f_{\theta_{\lambda 1}} \neq f_{\theta_{\lambda 2}}$ for almost all ω and $\theta_{\lambda_1}, \theta_{\lambda_2} \in \Theta$, where f_{θ_λ} denotes $(1 + \lambda z(\omega)'\theta_\lambda)^{1/\lambda}$.

Assumption A3 implies that the autocovariances of y_t satisfy the condition $\sum_{k=1}^{\infty} k\gamma_k^2 < \infty$. The latter condition, along with $f(\omega) > 0$, by Corollary I.3.1 in Dzhaparidze (1986, p. 66) guarantees that the principal part of the Gaussian log-likelihood can be approximated by the Whittle likelihood, i.e. it is needed for approximating the Gaussian likelihood $\ell^*(\lambda, \theta_\lambda)$ in (13) with $\ell^\dagger(\lambda, \theta_\lambda)$ in (14). The condition $\sum_{k=1}^{\infty} k\gamma_k^2 < \infty$ is implied by the summability of the cepstral coefficients, $\sum_{j=1}^{\infty} jc_j^2 < \infty$, which is fullfilled as $\ln[2\pi f(\omega)] = \frac{1}{\lambda}\ln(1 + \lambda g_\lambda(\omega))$, so that $\sum_{j=1}^{\infty} j\kappa_j^2 < \infty \implies \sum_{j=1}^{\infty} jc_j^2 < \infty$. Now, $\sum_{j=1}^{\infty} jc_j^2 < \infty \iff \sum_{j=1}^{\infty} k\gamma_j^2 < \infty$ is implied by Theorem 1 in Li (2005), who shows the equivalence of the two conditions when $f(\omega)$ is positive and continuous.

The derivatives of the inverse of the spectral density,

$$\frac{\partial}{\partial\theta_\lambda} f(\omega)^{-1} = -2\pi[1 + \lambda z(\omega)'\theta_\lambda]^{-(1+\lambda)/\lambda} z(\omega),$$

exist and are continuous in θ_λ. Hence, by Theorem II.2.1 in Dzhaparidze (1986, p. 105), $\tilde{\theta}_\lambda \to_p \theta_{\lambda 0}$

Moreover, the spectral density $f(\omega)$ is a twice differentiable function of θ_λ and the second derivatives

$$\frac{\partial^2}{\partial\theta_\lambda \partial\theta_\lambda'} f(\omega) = \frac{\lambda}{2\pi}\left(\frac{1}{\lambda} - 1\right)\left[1 + \lambda z(\omega)'\theta_\lambda\right]^{\frac{1}{\lambda}-2} z(\omega)z(\omega)'$$

are continuous in ω. Thus, by Theorem II.2.2 in Dzhaparidze (1986, p. 109), $\tilde{\theta}_\lambda$ is asymptotically normal as stated in the theorem.

References

Bhansali, R. J. (1983), The Inverse Partial Correlation Function of a Time Series and its Applications, *Journal of Multivariate Analysis*, 13, 310-327.

Barndorff-Nielsen, O.E., Schou, G. (1973), On the Parametrization of Autoregressive Models by Partial Autocorrelations, *Journal of Multivariate Analysis*, 3, 408–419.

Bingham, N. (2012), Szegö's theorem and its probabilistic descendants, *Probability Surveys*, 9, 287–324.

Bloomfield, P. (1973), An Exponential Model for the Spectrum of a Scalar Time Series, *Biometrika*, 60, 2, 217–226.

Bogert B.P., Healy M.J.R. and Tukey J.W. (1963), The quefrency alanysis of time series for echoes: cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe cracking, in Rosenblatt M. (ed.) *Proceedings of the Symposium on Time Series Analysis*, 209–243. Wiley, New York.

Box, G.E.P., and Cox, D.R. (1964), An analysis of transformations (with discussion), *Journal of the Royal Statistical Society, B*, 26, 211–246.

Brillinger, D.R. (2001), *Time Series. Data Analysis and Theory* , SIAM, Holden Day, San Francisco.

Brockwell, P.J. and Davis, R.A. (1991), *Time Series: Theory and Methods*, Springer-Verlag, New York.

Chiu, T.Y.M., Leonard, T. and Tsui, K-W. (1996), The Matrix-Logarithmic Covariance Model, *Journal of the American Statistical Association*, 91, 198-210.

Cleveland, W.S. (1972), The Inverse Autocorrelations of a Time Series and Their Applications, *Technometrics*, 14, 2, 277–293.

Dahlhaus, R. (1989), Efficient Parameter Estimation for Self Similar Processes. *The Annals of Statistics*, 17, 4, 1749-1766.

Dahlhaus, R. (2012). Locally Stationary Processes. In *Handbook of Statistics*, *Time Series Analysis: Methods and Applications*, Volume, 30, chapter 13, Elsevier, p. 351–408.

Doob, J.L. (1953), *Stochastic Processes*, John Wiley and Sons, New York.

Dzhaparidze, K. (1986), *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.*, Springer-Verlag, New York.

Giraitis, L., Koul, H.L. and Surgailis, D. (2012), *Large Sample Inference for Long Memory Processes*, Imperial College Press, London.

Gould, H.W. (1974), Coefficient Identities for Powers of Taylor and Dirichlet Series, *The American Mathematical Monthly*, 81, 1, 3–14.

Gray, H.L, Zhang, N.F. and Woodward, W.A. (1989), On Generalized Fractional Processes, *Journal of Time Series Analysis*, 10, 233–257.

Grenander, U. and Szegö, G. (1958), *Toeplitz Forms and Their Applications*, University of California Press, 1958.

Hosking, J.R.M. (1981), Fractional Differencing, *Biometrika*, 68, 165–176.

Hurvich, C.M., and Beltrao, K.I. (1993), Asymptotics for the low-frequency ordinates of the periodogram of a long-memory time series. *Journal of Time Series Analysis*, 14, 455-472.

Ibragimov, I.A. and Rozanov Y.A. (1978), *Gaussian random processes*, Springer-Verlag Berlin Heidelberg.

Kobayashi, T. and Imai, S. (1984), Spectral Analysis Using Generalized Cepstrum, *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-32, 1087–1089.

Koenker, R. and Yoon, J. (2009), Parametric Links for Binary Choice Models: A Fisherian-Bayesian Colloquy, *Journal of Econometrics*, 152, 120–130.

Künsch, H. R. (1986), Discrimination between monotonic trends and long-range dependence, *Journal of Applied Probability*, 23, 1025-1030.

Li, L.M. (2005), Some Notes on Mutual Information between Past and Future, *Journal of Time Series Analysis*, 27, 309–322.

McCullagh, T.S. and Nelder, J.A. (1989), *Generalized Linear Models*, Chapman & Hall Cambridge, UK.

Monahan, J.F. (1984), A Note Enforcing Stationarity in Autoregressive-Moving Average Models, *Biometrika*, 71, 2, 403–404.

Parzen, E. (1992), Time Series, Statistics, and Information, in *New Directions in Time Series Analysis*, (ed. E. Parzen et al) Springer Verlag, New York, 265-286.

Percival, D.B. and Walden, A.T. (1993), *Spectral Analysis for Physical Applications*, Cambridge University Press.

Pourahmadi M. (1983), Exact factorization of the spectral density and its application to forecasting and time series analysis, *Communications in Statistics*, 12, 18, 2085–2094.

Proietti, T., Luati, A. (2015), The Generalised Autocovariance function. Journal of Econometrics, **187**, 245–257.

Robinson, P. (1995). Log-Periodogram Regression of Time Series with Long Range Dependence, *Annals of Statistics*, 23, 1048–1072.

Whittle, P. (1953), Estimation and Information in Stationary Time Series, *Arkiv för Matematik*, 2, 423–434.