Generalised Linear Cepstral Models for the Spectrum of a Time Series

Tommaso Proietti  
University of Rome Tor Vergata and CREATEs

Alessandra Luati  
University of Bologna

Abstract

The exponential model for the spectrum of a time series is based on the Fourier series expansion of the logarithm of the spectral density. The coefficients of the expansion are the cepstral coefficients and their collection is the cepstrum of the time series. Approximate likelihood inference based on the periodogram leads to a generalised linear model for asymptotically independent exponential data with logarithmic link. The paper introduces the class of generalised linear cepstral models with Box-Cox link, which is based on the truncated Fourier series expansion of the Box-Cox transformation of the spectral density; the coefficients of the expansions can be termed generalised cepstral coefficients and are related to the generalised autocovariances of the series. The link function depends on a power transformation parameter, and encompasses the exponential model. Other important special cases are the inverse link (which leads to modelling the inverse spectrum), and the identity link. One of the merits of this model class is the possibility of nesting alternative spectral estimation methods (autoregressive, exponential, etc.) under the same likelihood-based framework.

Key words and phrases: Frequency Domain Methods, Generalised linear models, Iteratively reweighted least squares.

JEL codes: C22, C52.

*Tommaso Proietti acknowledges support from CREATEs - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation. The authors wish to thank Victor Solo, Scott Holan and Tucker McElroy for their comments, suggestions and bibliographic references that led to several improvements in the paper. Preliminary versions of this paper circulated under the title The Exponential Model for the Spectrum of a Time Series: Extensions and Applications.
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 prediction and interpolation. Inferences on the spectrum are based on the periodogram, which possesses a well established large sample distribution theory that leads to convenient likelihood based estimation and testing methods.

This paper is concerned with a class of generalised linear models formulated for the logarithm of the spectral density of a time series, known as the exponential model, which 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 Fourier transform of the logarithm of the spectral density function plays an important role in the analysis of stationary stochastic processes. It is the key element of the spectral factorization at the basis of prediction theory, leading to the Kolmogorov-Szegö formula for the prediction error variance (see Doob, 1953, theorem 6.2, Grenander and Rosenblatt, 1957, section 2.2, and Pourahmadi, 2001, Theorem VII). Bogert, Healy and Tuckey (1963) advocated its use for the analysis of series that are contaminated by echoes, namely seismological data, whose spectral densities typically factorize as the product of two components, one of which is the echo component. We refer the reader to Oppenheim and Schafer (2010, ch. 13), Brillinger (2001) and Childers, Skinner and Kemerialt (1977) for historical reviews on the cepstrum and its applications in signal processing. Solo (1986) extended the cepstral approach for modelling bivariate random fields.

Bloomfield (1973) introduced the exponential (EXP) model and discussed its maximum likelihood estimation, relying on the distributional properties of the periodogram of a short memory process, based on Walker (1964) and Whittle (1953). As illustrated by Cameron and Turner (1987), maximum likelihood estimation is computationally very attractive, being carried out by iteratively reweighted least squares.

Local likelihood methods with logarithmic link for spectral estimation have been considered by Fan and Kreutzberger (1998). Also, the exponential model has played an important role in regularized estimation of the spectrum (Wahba, 1980; Pawitan and O’Sullivan, 1994), where smoothness priors are enforced by shrinking higher order cepstral coefficients toward zero, and has been recently considered in the estimation of time-varying spectra (Rosen, Stoffer and Wood, 2009, and Rosen, Wood and Stoffer, 2012). Among other uses of the EXP model we mention discrimination and clustering of time series, as in Fokianos and Savvides (2008).
The model was then generalised to processes featuring long range dependence by Robinson (1991) and Beran (1993), originating the fractional EXP model (FEXP), whereas Janacek (1982) proposed a method of moments estimator of the long memory parameter based on the sample cepstral coefficients estimated nonparametrically using the log-periodogram. Maximum likelihood estimation of the FEXP model has been dealt with recently by Narukawa and Matsuda (2011).

The paper contributes to the current literature by introducing the class of generalised linear cepstral 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 the identity link. The coefficients of the model are related to the generalised autocovariances, see Proietti and Luati (2012), and are termed generalised cepstral coefficients. To enforce the constraints needed to guarantee the positivity of the spectral density, we consider a reparameterization of the generalised cepstral coefficients and we show that this framework is able to nest alternative spectral estimation methods, in addition to the exponential approach, namely autoregressive spectral estimation (inverse link) and moving average estimation (identity link), so that the appropriate method can be selected in a likelihood based framework. We also discuss testing for white noise in this framework.

The paper is structured as follows: section 2 provides a review of cepstral analysis and introduces the exponential model and its long memory extension, the fractional exponential model. Section 3 considers frequency domain estimation of the cepstral coefficients based on the periodogram or sample spectrum. The class of generalised linear cepstral model is the subject of section 4. Illustrations are provided in section 5. Finally, in section 6 we trace some conclusions and directions for further research.

2 The Exponential Model and Cepstral Analysis

Let \( \{y_t\}_{t \in T} \) be a stationary zero-mean stochastic process indexed by a discrete time set \( T \), with covariance function \( \gamma_k = \int_{-\pi}^{\pi} e^{i\omega k} dF(\omega) \), where \( F(\omega) \) is the spectral distribution function of the process and \( i \) is the imaginary unit. We assume that the spectral density function of the process exists, \( F(\omega) = \int_{-\pi}^{\pi} f(\lambda) d\lambda \), and that the process is regular (Doob, 1953, p. 564), i.e. \( \int_{-\pi}^{\pi} \ln f(\omega) d\omega > -\infty \). We shall start by assuming that \( f(\omega) \) is bounded in \( (-\pi, \pi) \), i.e. the process \( \{y_t\}_{t \in T} \) is a short memory process.

As \( f(\omega) \) is a positive, smooth, even and periodic function of the frequency, its logarithm can be expanded in a Fourier series as follows,

\[
\ln[2\pi f(\omega)] = c_0 + 2 \sum_{k=1}^{\infty} c_k \cos(k\omega),
\]

(1)
where the coefficients \( c_k \), \( k = 0, 1, \ldots \), are obtained by the (inverse) Fourier transform of \( \ln[2\pi f(\omega)] \),

\[
c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln[2\pi f(\omega)] \exp(\imath \omega k) \, d\omega.
\]

The coefficients \( c_k \) are known as the cepstral coefficients and the sequence \( \{c_k\}_{k=0,1,\ldots} \) is known as the cepstrum (Bogert, Healy and Tukey, 1963). The interpretation of the cepstral coefficients as pseudo-autocovariances is also discussed in Bogert, Healy and Tukey (1963) and essentially follows from the analogy with the Fourier pair \( 2\pi f(\omega) = \gamma_0 + 2\sum_{k=1}^{\infty} \gamma_k \cos(k\omega) \) and \( \gamma_k = \int_{-\pi}^{\pi} f(\omega) \exp(\imath \omega k) \, d\omega \).

Important characteristics of the underlying process can be obtained from the cepstral coefficients. The intercept is related to the one-step ahead prediction error variance (p.e.v.), \( \sigma^2 = \text{Var}(y_t|\mathcal{F}_{t-1}) \), where \( \mathcal{F}_t \) is the information set up to time \( t \): by the Szegő-Kolmogorov formula,

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

we get immediately that \( c_0 = \ln \sigma^2 \). Moreover, the long run variance is obtained as

\[
2\pi f(0) = \exp\left( c_0 + 2 \sum_{k=1}^{\infty} c_k \right).
\]

Also, if we let \( y_t = \varphi(B)\xi_t \) denote the invertible Wold representation of the process, with \( \varphi(B) = 1 + \varphi_1 B + \varphi_2 B^2 + \ldots, \sum_j \varphi_j^2 < \infty, \xi_t \sim \text{WN}(0,\sigma^2) \), where \( B \) is the lag operator, \( B^j y_t = y_{t-j} \), then the moving average coefficients of the Wold representation are obtained recursively from the formula

\[
\varphi_j = j^{-1} \sum_{r=1}^{j} r c_r \varphi_{j-r}, \quad j = 1, 2, \ldots, \tag{2}
\]

with \( \varphi_0 = 1 \). The derivation, see Janacek (1982), Pourahmadi (1983) and Hurvich (2002), is based on the spectral factorization \( 2\pi f(\omega) = \sigma^2 \varphi(\imath \omega) \varphi(\imath \omega) \); setting \( \varphi(z) = \exp\left( \sum_{k=1}^{\infty} c_k z^k \right) \), and equating the derivatives of both sides with respect to \( z \) at the origin, enables to express the Wold coefficients in terms of the cepstral coefficients, giving (2). The autoregressive representation \( \pi(B)y_t = \xi_t \), where \( \pi(B) = \sum_{j=0}^{\infty} \pi_j B^j = \varphi(B)^{-1} \), is easily determined from the relationship \( \ln \pi(z) = -\ln \varphi(z) \), and it is such that \( \pi_0 = 1 \) and \( \pi_j = -j^{-1} \sum_{r=1}^{j} r c_r \pi_{j-r}, \quad j = 1, 2, \ldots. \)

The mutual information between the past and the future of a Gaussian time series is defined in terms of the cepstral coefficients by Li (2005). \( I_{p-f} = \frac{1}{2} \sum_{k=1}^{\infty} k c_k^2 \), provided that \( \sum_{k=\infty}^{-\infty} k c_k^2 < \infty \), and the following relation hold between cepstral coefficients and the partial autocorrelation coefficients, \( \{\phi_{kk}\}_{k=1,2,\ldots} \), the so called reflectrum identity, \( \sum_{k=1}^{\infty} k c_k^2 = - \sum_{k=1}^{\infty} k \ln(1 - \phi_{kk}^2) \) and \( c_0 = \ln \gamma_0 + \sum_{k=1}^{\infty} \ln(1 - \phi_{kk}^2) \), the latter being a consequence of the Kolmogorov-Szegő formula.

We also note that the Fourier expansion (2) is equivalent to express the logarithm of the spectral density function as \( \ln[2\pi f(\omega)] = c_0 + s(\omega) \) where \( s(\omega) \) is a linear spline function, \( s(\omega) = \int_{0}^{\omega} B(z) \, dz \), and
B(z) is a Wiener process, when the canonical representation for the spline basis functions is chosen, i.e. via the Demmler-Reinsch basis functions (Demmler and Reinsch, 1975, see also Eubank, 1999). This representation is applied in a Bayesian setting by Rosen, Wood and Stoffer (2012) and Rosen, Stoffer and Wood (2009).

2.1 Cepstral Analysis of ARMA Processes

If yt is a white noise (WN) process, c_{k} = 0, k > 0. Figure II displays the cepstrum of the AR(1) process
\[ y_t = \phi y_{t-1} + \xi_t, \xi_t \sim WN(0, \sigma^2) \]
with \( \phi = 0.9 \) and coefficients \( c_k = \phi^k/k \) (top left plot). The behaviour of the cepstrum is analogous to that of the autocovariance function, although it will dampen more quickly due to the presence of the factor \( k^{-1} \). The upper right plot is the cepstrum of the MA(1) process \( y_t = \xi_t + \theta \xi_{t-1} \), with \( \theta = 0.9 \); the general expression is \( c_k = -(-\theta)^k/k \). Notice that if \( c_k, k > 1 \), are the cepstral coefficients of an AR model, that of an MA model of the same order with parameters \( \theta_j = -\phi_j \) are \(-c_k\). Hence, for instance, the cepstral coefficients of an MA(1) process with coefficient \( \theta = -0.9 \) are obtained by reversing the first plot. The bottom plots concern the cepstra of two pseudo-cyclical processes: the AR(2) process \( y_t = 1.25y_{t-1} - 0.95y_{t-2} + \xi_t \) with complex roots, and the ARMA(2,1) process \( y_t = 1.75y_{t-1} - 0.95y_{t-2} + \xi_t + 0.5\xi_{t-1} \). The cepstra behave like a damped sinusoidal, and again the damping is more pronounced than it shows in the autocovariance function. Notice also that even for finite \( p \) and \( q \) we need infinite coefficients \( c_k \) to represent an ARMA model.

For a general ARMA process \( y_t \sim ARMA(p, q) \), \( \phi(B) y_t = \theta(B) \xi_t \), with AR and MA polynomials factorized in terms of their roots, as in Brockwell and Davis (1991, section 4.4),
\[ \phi(B) = \prod_{j=1}^{p}(1 - a_j^{-1}B), \quad \theta(B) = \prod_{j=1}^{q}(1 - b_j^{-1}B), |a_j| > 1, |b_j| > 1, \]
we have the following general result
\[
\ln[2\pi f(\omega)] = c_0 + 2 \sum_{k=1}^{\infty} \left( \sum_{j=1}^{q} c_{jk}^{(b)} - \sum_{j=1}^{p} c_{jk}^{(a)} \right) \cos(k\omega),
\]
where
\[
c_{jk}^{(a)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln|1 - a_j^{-1}e^{-i\omega}|^2 \cos(\omega) d\omega, \quad c_{jk}^{(b)} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \ln|1 - b_j^{-1}e^{-i\omega}|^2 \cos(\omega) d\omega.
\]
This is the sum of elementary cepstral processes corresponding to polynomial factors. When \( a_j \) and \( b_j \) are real \( c_{jk}^{(a)} = -a_j^{-k}/k \) and \( c_{jk}^{(b)} = -b_j^{-k}/k \) (see Gradshteyn and Ryzhik, 1994, 4.397.6).

When there are two complex conjugate roots, \( a_j = r^{-1}e^{i\varpi}, \bar{a}_j = r^{-1}e^{-i\varpi} \), with modulus \( 1/r \) and phase \( \varpi \), their contribution to the cepstrum is via the coefficients \( r^k \cos(\varpi k)/k \). Hence, the cepstral coefficients of the stationary cyclical process \( (1 - 2r \cos \varpi B + r^2 B^2)y_t = \xi_t \) are \( c_k = r^k \cos(\varpi k)/k, k = 1, 2, \ldots \);
Figure 1: Cepstral coefficients $c_k, k = 1, \ldots, 20$ for selected ARMA models
see the bottom left plot of figure \( \Box \) for which \( r = 0.97 \) and \( \varpi = 0.87 \): the cepstral coefficient have a period of about 7 units.

### 2.2 Truncated Cepstral Processes

The class of stochastic processes characterised by an exponential spectrum was proposed by Bloomfield (1973), who suggested truncating the Fourier representation of \( \ln 2\pi f(\omega) \) to the \( K \) term (EXP(\( K \)) process), so as to obtain:

\[
\ln[2\pi f(\omega)] = c_0 + 2\sum_{k=1}^{K} c_k \cos(\omega k). \tag{3}
\]

The EXP(1) process, characterized by the spectral density \( f(\omega) = (2\pi)^{−1}\exp(c_0 + 2c_1 \cos \omega) \), has autocovariance function

\[
\gamma_k = \sigma^2 I_k(2c_1) = \sigma^2 \sum_{j=0}^{\infty} \frac{c_{1+j}^2}{(k+j)!j!},
\]

where \( I_k(2c_1) \) is the modified Bessel function of the first kind of order \( k \), see Abramowitz and Stegun, (1972), 9.6.10 and 9.6.19, and \( \sigma^2 = \exp(c_0) \). Notice that there is an interesting analogy with the Von Mises distribution on the circle, \( f(\omega) = \sigma^2 I_0(c_1)g(\omega; 0, 2c_1) \), where \( g(.) \) is the density of the Von Mises distribution, see Mardia and Jupp (2000). For integer \( k \), \( \gamma_k \) is real and symmetric. The coefficients of the Wold represntation are obtained from (1): \( \varphi_j = (j!)^{-1}c_1^j \), \( j > 0 \), which highlights the differences with the impulse response of autoregressive (AR) process of order 1 (it converges to zero at a faster rate than geometric).

The truncated cepstral process of order \( K \), with \( f(\omega) = \frac{1}{2\pi} \exp(c_0 + 2\sum_{k=1}^{K} c_k \cos(\omega k)) \), is such that the spectral density can be factorized as

\[
f(\omega) = \frac{\sigma^2}{2\pi} \prod_{k=1}^{K} I_0(2c_k) + 2\sum_{j=1}^{\infty} I_j(c_k) \cos(\omega k j).
\]

This result comes from the Fourier expansion of the factors \( \exp(2c_k \cos(\omega k)) \).

### 2.3 Fractional Exponential (FEXP) processes

We now move to long memory processes, whose spectral density is unbounded at the origin. The process \( \{y_t\}_{t \in T} \) is generated according to the equation \( y_t = (1 − B)^{-d} \xi_t \), where \( \xi_t \) is a short memory process and \( d \) is the long memory parameter, \( 0 < d < 0.5 \). The spectral density can be written as \( 2\sin(\omega/2)^{−2d} f_\xi(\omega) \), where the first factor is the power transfer function of the filter \( (1 − B)^{-d} \), i.e. \( |1 − e^{-i\omega}|^{−2d} = |2\sin(\omega/2)|^{−2d} \), and \( f_\xi(\omega) \) is the spectral density function of a short memory process, whose logarithm admits a Fourier series expansion.
The logarithm of the spectral generating function of \( y_t \) is thus linear in \( d \) and in the cepstral coefficients of the Fourier expansion of \( \ln[2\pi f(\omega)] \), denoted \( c_k, k = 1, 2, \ldots \):

\[
\ln[2\pi f(\omega)] = -2d \ln \left| 2 \sin \frac{\omega}{2} \right| + c_0 + 2 \sum_{k=1}^{\infty} c_k \cos(\omega k). \tag{4}
\]

Here, \( c_0 \) retains its link to the p.e.v., \( \sigma^2 = \exp(c_0) \), as \( \int_{-\pi}^{\pi} \ln \left| 2 \sin \frac{\omega}{2} \right| d\omega = 0 \). In view of the result

\[- \ln \left| 2 \sin \frac{\omega}{2} \right| = \sum_{k=1}^{\infty} \frac{\cos(\omega k)}{k}\]

(see also Gradshteyn and Ryzhik, 1994, formula 1.441.2), which tends to infinity when \( \omega \to 0 \), we rewrite (4) as

\[
\ln[2\pi f(\omega)] = c_0 + 2 \sum_{k=1}^{\infty} (c_k^* + c_k) \cos(\omega k),
\]

with

\[
c_k^* = -\frac{1}{2\pi} \int_{-\pi}^{\pi} 2d \ln \left| 2 \sin(\omega/2) \right| \cos(k\omega) d\omega = \frac{d}{k}, k, k > 0.
\]

Hence, for a fractional noise (FN) process, for which \( y_t \sim \text{WN}(0, \sigma^2) \), the cepstral coefficients show a hyperbolic decline (\( c_k = d/k, k, k > 0 \)).

When \( \ln 2\pi f(\omega) \) is approximated by an \( \text{EXP}(K) \) process, i.e. the last summand of (4) is truncated at \( K \), a fractional exponential process of order \( K \), \( \text{FEXP}(K) \), arises. The fractional noise case corresponds to the \( \text{FEXP}(0) \) process.

### 3 The Periodogram and the Whittle likelihood

The main tool for estimating the spectral density function and its functionals is the periodogram. Due to its sampling properties, a generalised linear model for exponential random variables with logarithmic link can be formulated for the spectral analysis of a time series in the short memory case. The strength of the approach lies in the linearity of the log-spectrum in the cepstral coefficients.

Let \( \{y_t, t = 1, 2, \ldots, n\} \) denote a time series, which is assumed to be a sample realisation from a stationary short memory Gaussian process, and let \( \omega_j = \frac{2\pi j}{n}, j = 1, \ldots, [n/2] \), denote the Fourier frequencies, where \([\cdot]\) denotes the integer part of the argument. The periodogram, or sample spectrum, is defined as

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

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

\[
\frac{I(\omega_j)}{f(\omega_j)} \sim \text{IID} \frac{1}{2\lambda^2}, \quad \omega_j = \frac{2\pi j}{n}, j = 1, \ldots, [(n-1)/2],
\tag{5}
\]

8
whereas \( I(\omega_j) / f(\omega_j) \sim \chi^2_m \), \( \omega_j = 0, \pi \), where \( \chi^2_m \) denotes a chi-square random variable with \( m \) degrees of freedom, or, equivalently, a Gamma\((m/2,2)\) random variable. As a particular case, \( \frac{1}{2} \chi^2_2 \) is an exponential random variable with unit mean.

The above distributional result is the basis for approximate or Whittle maximum likelihood inference on the EXP\((K)\) model for the spectrum of a time series: denoting by \( \theta \) the vector of cepstral coefficients, \( \theta' = (c_0, c_1, c_2, \ldots, c_K) \), and writing \( f(\omega) = (2\pi)^{-1} \exp(c_0 + 2 \sum_{k=1}^{K} c_k \cos(\omega k)) \), the log-likelihood of \( \{I(\omega_j), j = 1, \ldots, N = [(n - 1)/2]\} \), is:

\[
\ell(\theta) = - \sum_{j=1}^{N} \left[ \ln f(\omega_j) + \frac{I(\omega_j)}{f(\omega_j)} \right].
\]

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. Estimation by maximum likelihood (ML) of the EXP\((K)\) model was proposed by Bloomfield (1973); later Cameron and Turner (1987) showed that ML estimation is carried out by iteratively reweighted least squares (IRLS).

In some cases, tapering the series before computing the periodogram may improve the estimation, in that tapering reduces the finite sample bias due to the leakage in spectral estimation that arises from neighbouring regions characterised by high spectral density. A taper is a data window taking the form of a sequence of positive weights \( h_t, t = 1, \ldots, n \) that leaves unaltered the series in the middle of the sample and down-weights the observations at the extremes. In other words, tapering amounts to smoothing the observed sample transition from zero to the observed values when estimating convolutions of data sequences such as the periodogram. Tapering is particularly important for spectral density with large dynamic range due to the presence of cyclical components. On the converse, data tapering may induce some correlation among the observations and may cause a loss in efficiency. Brillinger (1981, Theorem 5.2.7) shows that for the tapered periodogram the same distributional result as (6) holds. In our applications we shall use a taper formed for zeroth-order discrete prolate spheroidal sequences (dpss), for which we refer to Percival and Walden (1981, sec. 3.9 and ch. 7).

4 Generalized Linear Cepstral Models with Power Link

The EXP model is a generalised linear model (GLM, McCullagh and Nelder, 1989) for exponentially distributed data, the periodogram ordinates at the Fourier frequencies, with mean function given by the spectral density and a logarithmic link function.

The generalization that we propose in this section\(^1\) refers to the short memory case and is based on the

---

\(^1\) After the preparation of this paper, Scott H. Holan (University of Missouri) pointed out a very insightful paper by Parzen (1992), in which he had considered the development of power and cepstral correlation analysis as one of the new directions of time
observation that any continuous monotonic transform of the spectral density function can be expanded as a Fourier series. We focus, in particular, on a parametric class of link functions, the Box-Cox link (Box and Cox, 1964), depending on a power transformation parameter, that encompasses the EXP model (logarithmic link), as well as the the identity and the inverse links; the latter is also the canonical link for exponentially distributed observations.

Let us thus consider the Box-Cox transform of the spectral density function $2\pi f(\omega)$,

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

Its Fourier series expansion, truncated at $K$, is

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

and the coefficients

$$c_{\lambda,k} = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_\lambda(\omega) \cos(\omega k) d\omega$$

will be named generalised cepstral coefficients (GCC).

Hence, a linear model is formulated for $g_\lambda(\omega)$. The spectral model with Box-Cox link and mean function

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

will be referred to as a generalised cepstral model (GCM) with parameter $\lambda$ and order $K$, GCM($\lambda,K$), in short. The EXP model thus corresponds the the case when the power parameter $\lambda$ is equal to zero, and $c_{0,k} = c_k$, are the usual cepstral coefficients.

For $\lambda \neq 0$, the GCC’s are related to the generalised autocovariance function, introduced by Proietti and Luati (2012),

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

by the following relationships:

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

In turn, the generalised autocovariances are interpreted as the traditional autocovariance function of the power-transformed process:

$$u_\lambda = \left[ \sigma_\varphi \left( B^{S(\lambda)} \right) \right]^{\lambda} \xi_t,$$

series analysis. We are grateful to Scott Holan for providing this reference.
where $\xi_t^* = \sigma^{-1}\xi_t$, $s(\lambda)$ is the sign of $\lambda$, and $[\varphi(B^{\lambda})]$ is a series in the lag operator whose coefficients can be derived in a recursive manner based on the Wold coefficients by Gould (1974). For $\lambda = 1$, $c_{1k} = \gamma_k$, the autocovariance function of the process is obtained. In the case $\lambda = -1$ and $k \neq 0$, $c_{-1,k} = -\gamma_k$, where $\gamma_k$ is the inverse autocovariance of $y_t$ (see Cleveland, 1972). The intercept $c_{00}$ 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, $\exp(c_{00}) = \sigma^2$, the prediction error variance, and $c_{1,0} + 1 = \gamma_0$ is the unconditional variance of $y_t$. Also, for $\lambda \to 0$, $c_{\lambda k} \to c_k$, i.e. the cepstrum is the limit of the GCC as $\lambda$ goes to zero.

For a fractional noise process the GCCs are zero for $\lambda = -d^{-1}$ and $k > 1$. This is so since $2\pi f(\omega)^\lambda = \sigma^2 |2\sin(\omega/2)|^{-2d\lambda} = \sigma^2 |1 - e^{-i\omega}|^2$ for $\lambda = -d^{-1}$, which is the spectrum of a non-invertible moving average process of order 1, whose autocovariances are $\gamma_{\lambda k} = 0$ for $k > 1$.

### 4.1 Whittle Likelihood Estimation

Let $g_\lambda(\omega_j) = z_j^{' \theta_\lambda}$, with

$$z_j = [1, 2 \cos \omega_j, 2 \cos(2\omega_j), \ldots, 2 \cos(K\omega_j)]'$$

and

$$\theta_\lambda = [c_{\lambda 0}, c_{\lambda 1}, \ldots, c_{\lambda K}]'.$$

Then, the approximate Whittle likelihood is

$$\ell(\theta_\lambda) = N \ln 2\pi - \sum_{j=1}^N \ell_j(\theta_\lambda),$$

where, for $1 + \lambda z_j^{' \theta_\lambda} > 0$,

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

and the approximate Whittle estimator of $\theta_\lambda$ is $\tilde{\theta}_\lambda$ such that

$$\ell(\tilde{\theta}_\lambda) = \max_{\theta_\lambda \in \Theta} \ell(\theta_\lambda),$$

where $\Theta \subseteq \mathbb{R}^{K+1}$.

The score vector and the Hessian matrix, when $\lambda \neq 0$, are respectively

$$s(\theta_\lambda) = \frac{\partial \ell(\theta_\lambda)}{\partial \theta_\lambda} = -\sum_j z_j^* u_j, \quad z_j^* = \frac{z_j}{1 + \lambda z_j^{' \theta_\lambda}}, \quad u_j = \begin{cases} 
1 - \frac{2\pi I(\omega_j)}{(1 + \lambda z_j^{' \theta_\lambda})^{\frac{1}{\lambda}}}, & \lambda \neq 0, \\
1 - \frac{2\pi I(\omega_j)}{\exp(z_j^{' \theta_\lambda})}, & \lambda = 0
\end{cases}$$
\[ H(\theta_\lambda) = \frac{\partial^2 \ell(\theta_\lambda)}{\partial \theta_\lambda \partial \theta'_\lambda} = -\sum_j W_j^* z_j^* z'_j, \]

with

\[ W_j^* = \begin{cases} \frac{2\pi I(\omega_j)}{(1+\lambda z_j^* \theta_\lambda)^2} - \lambda u_j, & \lambda \neq 0, \\ \frac{2\pi I(\omega_j)}{\exp(z_j^* \theta_\lambda)}, & \lambda = 0, \end{cases} \]

so that the expected Fisher information is \( \mathcal{I}(\theta) = -E[H(\theta)] = \sum_j z_j^* z'_j. \)

Maximum likelihood estimation is carried out numerically by the Newton-Raphson algorithm, i.e. iterating until convergence

\[ \hat{\theta}_{i+1} = \hat{\theta}_i - [H(\hat{\theta}_i)]^{-1} s(\hat{\theta}_i) \]

or by the method of scoring:

\[ \hat{\theta}_{i+1} = \hat{\theta}_i + [I(\hat{\theta}_i)]^{-1} s(\hat{\theta}_i) \]

with fixed initial conditions, e.g. the white noise case.

The asymptotic theory for \( \hat{\theta}_\lambda \) is based on Dzhaparidze (1986, ch. 2). For a Gaussian process with positive spectral density function, under the usual regularity conditions (the true parameter \( \theta_0 \) is an inner point of \( \Theta \), the model is identifiable and the derivatives \( \frac{\partial f^{-1}(\omega)}{\partial \theta(\lambda)} \) exist and are continuous for all \( l \)), then \( \hat{\theta} \to_p \theta \). If, additionally, \( \frac{h'z_j}{[2\pi f_0(\omega)]^\lambda} \neq 0 \) for all \( \omega \) and \( h = (h_0, \ldots, h_K) \neq 0 \) then, setting \( z(\omega) = [1, \cos(\omega), \cos(2\omega), \ldots, \cos(K\omega)]' \),

\[ \sqrt{n}(\hat{\theta}_\lambda - \theta_\lambda) \to_d N(0, V_\lambda), \quad V_\lambda^{-1} = \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{1}{[2\pi f(\omega)]^{2\lambda}} z(\omega) z(\omega)' d\omega. \]

In the exponential case, when \( \lambda = 0 \), the derivation is the same and so are the asymptotic results, provided that \( z_j^* \) is replaced by \( z_j \).

In the case of misspecification, it is possible to derive the asymptotic covariance matrix of \( \hat{\theta}_\lambda \) by following Bloomfield (1972, 1973). However, in the present case, when we are faced with the further assumption of positivity of the inverse Box-Cox transform, we suggest to reparametrize the model as in section 4.3.

4.2 White Noise and Goodness of Fit Test

We consider the problem of testing the white noise hypothesis \( H_0 : c_{\lambda 1} = c_{\lambda 2} = \cdots = c_{\lambda K} = 0 \) in the GCM framework, with \( \lambda \) and \( K \) given. Interestingly, the score test statistic is invariant with \( \lambda \) and is asymptotically equivalent to the Box-Pierce (1970) test statistic. This is immediate to see for the traditional EXP(\( K \)) case, that is when \( \lambda = 0 \). Under the assumption \( y_t \sim WN(0, \sigma^2) \), \( 2\pi f(\omega) = \exp(c_0) \) and the Whittle estimator of \( c_0 \) is the logarithm of the sample periodogram mean, \( 2\pi I = \frac{1}{N} \sum_j 2\pi I(\omega_j) \) (which is also an estimate of the variance - for a WN process the p.e.v. equals the variance): \( \tilde{c}_0 = \log(2\pi I). \)
The score test of the null $H_0 : c_1 = c_2 = \cdots = c_K = 0$ in an EXP($K$) model is

$$S_{WN}(K) = n^{-1} \left( \sum z_j^T \tilde{u}_j \right)^T \left( \sum z_j \tilde{u}_j \right) \approx n \sum_{k=1}^{K} \rho_k^2,$$

where $z_j = [1, 2 \cos \omega_j, 2 \cos(2 \omega_j), \ldots, 2 \cos(K \omega_j)]$, $\tilde{u}_j = 1 - I(\omega_j)/\bar{I}$, and we rely on the large sample approximations: $2N \approx n$, $\frac{1}{N} \sum_{j=1}^{N} I(\omega_j) \cos(\omega_j k) \approx \tilde{\rho}_k$, the lag $k$ sample autocorrelation (see Brockwell and Davis, prop. 10.1.2). Hence, $S_{WN}(K)$ is the same as the Box-Pierce (1970) portmanteau test statistic. The same holds in the case when $\lambda \neq 0$, where $g_{\lambda}(\omega) = c_0$, $\tilde{c}_0 = (\frac{2\pi I}{\lambda} - 1)$ and $1 + \lambda \tilde{c}_0 = (2\pi \bar{I})^\lambda$.

On the contrary, the likelihood ratio test can be shown to be equal to

$$LR = 2N \left( \ln \bar{I} - \frac{1}{N} \sum_{j=1}^{N} \ln \tilde{f}(\omega_j) \right) = 2N \ln \frac{V}{p.e.v.},$$

where $V = 2\pi \bar{I}$ is the unconditional variance of the series, estimated by the averaged periodogram, and the prediction error variance in the denominator is estimated by the geometric average of the estimated spectrum under the alternative. The former is also the p.e.v. implied by the null model; the latter depends on $\lambda$. Interestingly, the LR test can be considered a parametric version of the test proposed by Davis and Jones (1968), based on the comparison of the unconditional and the prediction error variance.

### 4.3 Reparameterization

The main difficulty with maximum likelihood estimation of the GCM($\lambda, K$) model for $\lambda \neq 0$ is enforcing the condition $1 + \lambda z_j^T \theta_\lambda > 0$. This problem is well known in the literature concerning generalised linear models with the inverse link for gamma distributed observations, for which the canonical link is the inverse link (McCullagh and Nelder, 1989). Several strategies may help overcoming this problem, such as periodogram pooling (Bloomfield, 1973, Moulines and Soulier, 1999, Faï, Moulines and Soulier, 2002), which reduces the influence of the periodogram ordinates close to zero, and weighting the periodogram, so as to exclude in the estimation those frequencies for which the positivity constraint is violated.

The most appropriate solution is to reparameterize the generalised autocovariances and the cepstral coefficients as follows:

$$[2\pi f(\omega)]^\lambda = \sigma_\lambda^2 b_\lambda(e^{-i\omega}) b_\lambda(e^{i\omega}), \quad b_\lambda(e^{-i\omega}) = 1 + b_1 e^{-i\omega} + \cdots + b_K e^{-i\omega K},$$

where the $b_k$ coefficients are such that the roots of the polynomial $1 + b_1 z + \cdots + b_K z^K$ lie outside the unit circle, so that, for $\lambda \neq 0$, the GCC’s are obtained as

$$c_{\lambda 0} = \frac{1}{\lambda} \left[ \sigma_\lambda^2 (1 + b_1^2 + \cdots + b_K^2) - 1 \right], \quad c_{\lambda k} = \frac{1}{\lambda} \sigma_\lambda^2 \sum_{j=k}^{K} b_j b_{j-k}.$$
To ensure the positive definiteness and the regularity of the spectral density we adopt a reparameterization due to Barndorff-Nielsen and Schou (1973) and Monahan (1984): given $K$ coefficients $\varsigma_{\lambda k}, k = 1, \ldots, K$, such that $|\varsigma_{\lambda k}| < 1$, the coefficients of the polynomial $b_{\lambda}(z)$ are obtained from the last iteration of the Durbin-Levinson recursion

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

for $k = 1, \ldots, K$ and $j = 1, \ldots, k - 1$. The coefficients $\varsigma_{\lambda j}$ are in turn obtained as the Fisher inverse transformations of unconstrained real parameters $\vartheta_j, j = 1, \ldots, K$, i.e. $\varsigma_{\lambda j} = \frac{\exp(2\vartheta_j) - 1}{\exp(2\vartheta_j) + 1}$ for $j = 1, \ldots, K$, which are estimated unrestrictedly. Also, we set $\vartheta_0 = \ln(\sigma^2_{\chi})$.

By this reparameterisation, alternative spectral estimation methods are nested within the GCM($\lambda, K$) model. In particular, along with the EXP model ($\lambda = 0$), autoregressive estimation of the spectrum arises in the case $\lambda = -1$, whereas $\lambda = 1$ (identity link) amounts to fitting the spectrum of a moving average model of order $K$ to the series. 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. The profile likelihood of the GCM($\lambda, K$) as $\lambda$ 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; another possibility is to test for the adequacy of a maintained link (e.g. the logarithmic one) using the goodness of link test proposed by Pregibon (1980).

In conclusion, the GCM framework enables the selection of a spectral estimation method in a likelihood based framework. Another possible application of the GCM($\lambda, K$) is the direct estimation of the inverse spectrum and inverse autocorrelations up to the lag $K$, which arises for $\lambda = -1$ (this corresponds to the inverse link) and of the optimal interpolator (Battaglia, 1983), which is obtained in our case from the corresponding $b_k$ coefficients as $\sum_{k=1}^{K} \rho_{-1,k} y_{t+k}$ with

$$\rho_{-1,k} = \frac{\sum_{j=k}^{K} b_{-1,j} b_{-1,j-k}}{\sum_{j=0}^{K} b_{-1,j}^2}.$$

which represents the inverse autocorrelation at lag $k$ of $y_t$.

## 5 Illustrations

The proposed generalizations will now be applied to three well known time series that have been analyzed extensively in the literature and that provide a useful testbed for the class of generalised linear cepstral models.
5.1 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 2 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 GCM(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 by an information criterion, such as the Akaike Information Criterion (AIC):

\[ AIC(K, \lambda) = -2\ell(\hat{\theta}_\lambda) + 2K, \]

where \( \ell(\hat{\theta}_\lambda) \) is the Whittle likelihood of the GCM(λ, K) model, evaluated at the maximum likelihood estimate of the parameters \( \theta_\lambda \); the latter is obtained using the reparameterisation considered in section 4.3.

The AIC always leads to \( K = 7 \) for all values of \( \lambda \). Figure 3 displays the prediction error variance and the profile Whittle likelihood of GCM(λ, 7) models, as a function of \( \lambda \). This suggests that the optimal value of the power transformation parameter is \( \hat{\lambda} = -2.28 \).

The estimated spectrum is \( \hat{f}(\omega) = \frac{1}{2\pi} \left[ \tilde{\sigma}_\lambda^2 \tilde{b}_\lambda(e^{-i\omega})\tilde{b}_\lambda(e^{i\omega}) \right]^{-1/2.278} \), with \( \tilde{\sigma}_\lambda^2 = 2161.74 \), and \( \tilde{b}_\lambda(e^{-i\omega}) = 1 - 1.02e^{-i\omega} - .03e^{-i2\omega} - .05e^{-i3\omega} - .08e^{-i4\omega} + .04e^{-i5\omega} - .02e^{-i6\omega} + .23e^{-i7\omega} \).

From the second panel of figure 3 it is evident that the likelihood ratio test of \( \lambda = -2 \) for a GCM(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 is the square of a polynomial of order 7. This polynomial has three pairs of complex conjugate roots and a real root.

Figure 4 plots the periodogram of the SOI series and overimposes the spectral densities fitted by the GCM(λ, K) model with \( K = 7 \) and \( \lambda \) set equal to 1, 0, -1 and \( \hat{\lambda} = -2.28 \). The case when \( \lambda \) is set equal to 1 corresponds to fitting an MA(7) model to the series, whereas the case \( \lambda = 0 \) corresponds to fitting the EXP(7) model; \( \lambda = -1 \) 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.
Figure 2: Southern Oscillation Index. Time series and sample autocorrelation function. In the first plot the horizontal lines are drawn at ±8.

Figure 3: Southern Oscillation Index. Prediction error variance and Whittle likelihood as a function of $\lambda$ for GCM($\lambda, K$) models with $K = 7$. 
Figure 4: Southern Oscillation Index. Comparison of the spectral density estimates arising from different GCM(λ, K) models with K = 7.

5.2 Box and Jenkins Series A

Our second empirical illustration deals with a time series popularised by Box and Jenkins (1970), concerning a sequence of n = 200 readings of a chemical process concentration, known as Series A. The series, plotted in figure 5, was investigated in the original paper by Bloomfield (1973), with the intent of comparing the EXP model with ARMA models. Bloomfield fitted a model with K chosen so as to match the number of ARMA parameters. Box and Jenkins (1970) had fitted an ARMA(1,1) model to the levels and an AR(1) to the differences. The estimated p.e.v. resulted 0.097 and 0.101, respectively. Thus, Bloomfield fitted the EXP(2) model to the levels and an EXP(1) to the 1st differences by maximum likelihood, using a modification which entails concentrating σ² out of the likelihood function. The estimated p.e.v. resulted 0.146 and 0.164, respectively. He found this rather disappointing and concluded that ARMA models are more flexible.

Actually, there is no reason for constraining K to the number of parameters of the ARMA model. If model selection is carried out and estimation by MLE is performed by IRLS, AIC selects an EXP(7) for the levels and an EXP(5) for the 1st differences. The estimated p.e.v. is 0.099 and 0.097, respectively. BIC selects an EXP(3) for both series. The p.e.v. estimates are 0.103 and 0.103.

Also, the FEXP(0) provides an excellent fit: the d parameter is estimated equal to 0.437 (with standard
Figure 5: Box and Jenkins (1970) Series A. Chemical process concentration readings.

error 0.058), and the p.e.v. is 0.100.

Figure 6 presents the periodogram and the fitted spectra for the two EXP specifications and the FEXP model (left plot). The right plot displays the centered log-periodogram \( \ln [2\pi I(\omega_j)] - \psi(1) \) and compares the fitted log-spectra. It could be argued that EXP(7) is prone to overfitting and that the FEXP(0) model provides a very good fit, the first periodogram ordinate \( I(\omega_1) \) being very influential in determining the fit.

The FEXP(0) model estimated on the first differences yields an estimate of the memory parameter \( d \) equal to -0.564 (s.e. 0.056), and the p.e.v. is 0.098. These results are consistent with the FEXP model applied to the levels, as a negative \( d \) is estimated.

This example illustrates that the exponential model provides a fit that is comparable to that of an ARMA model, in terms of the prediction error variance. There is a possibility that the series has long memory, which agrees with the finding in Beran (1995) and Velasco and Robinson (2000).

When we move to fitting the more general class of GMC(\( \lambda, K \)) models, both AIC and BIC select the model GMC(-2.29, 1); see table 1, which refers to the AIC. Notice that the EXP(5) and EXP(7) are characterised by a much higher AIC (see the row corresponding to \( \lambda = 0 \)). Table 2 displays the values of the estimated \( b_1 \) coefficient and the corresponding generalised cepstral coefficient \( c_{\lambda 1} \), as well as the value of the maximised likelihood and prediction error variance, for the first order model GMC(\( \lambda, 1 \)), as the transformation parameter varies. For the specification selected according to information criteria, the implied spectrum is

\[
2\pi \hat{f}(\omega) = \sigma^2_{\lambda} |1 + \tilde{b}_1 e^{-i\omega/2\lambda}|^{-2|2\sin(\omega/2)|^{-2\times 0.44}},
\]

which results from replacing \( \tilde{b}_1 = -1 \) and
Figure 6: BJ Series A. Spectrum and log-spectrum estimation by an exponential model with $K$ selected by AIC and a fractional exponential model.
Table 1: BJ Series A. Values of the Akaike Information Criterion for GCM(\(\lambda, K\)) models. The selected model is GCM(-2.29, 1).

<table>
<thead>
<tr>
<th>(\lambda)</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.50</td>
<td>-3.000</td>
<td>-3.083</td>
<td>-3.085</td>
<td>-3.079</td>
<td>-3.071</td>
<td>-3.070</td>
<td>-3.087</td>
</tr>
<tr>
<td>0.00</td>
<td>-2.955</td>
<td>-3.055</td>
<td>-3.072</td>
<td>-3.069</td>
<td>-3.061</td>
<td>-3.056</td>
<td>-3.077</td>
</tr>
<tr>
<td>0.50</td>
<td>-2.918</td>
<td>-3.024</td>
<td>-3.057</td>
<td>-3.062</td>
<td>-3.053</td>
<td>-3.045</td>
<td>-3.063</td>
</tr>
<tr>
<td>1.00</td>
<td>-2.888</td>
<td>-2.994</td>
<td>-3.037</td>
<td>-3.053</td>
<td>-3.048</td>
<td>-3.038</td>
<td>-3.049</td>
</tr>
</tbody>
</table>

\(\hat{\lambda} = -2.29\) and by application of the two-angle trigonometric formula. This is the spectral density of a fractional noise process with memory parameter \(d = 0.44\). It is indeed remarkable that likelihood inferences and model selection applied to the GCM(\(\lambda, K\)) model point to the same results obtained by the FEXP(0) model discussed above. We interpret these results as further confirming the long memory nature of the series.

5.3 Simulated AR(4) Process

As our third example we consider \(n = 1024\) observations simulated from the AR(4) stochastic process

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

The series is obtained from Percival and Walden (1993) and constitutes a test case for spectral estimation methods, as the data generating process features a bimodal spectrum, with the peaks located very closely. 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. As in Percival and Walden, the series is preprocessed by a dpss data taper (with bandwidth parameter \(W = 2/n\), see Percival and Walden, 1993, sections 6.4 and 6.18, for more details).

The specifications of the class GCM(\(\lambda, K\)) 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 GCM\((-1, 4)\) model and their estimation standard errors are
Table 2: GCM(\(\lambda, 1\)) models: Whittle likelihood estimates of \(b_1\), the GCC \(c_{\lambda_1}\); value of log-likelihood at the maximum, \(\ell(\hat{\theta})\), and prediction error variance (p.e.v.).

<table>
<thead>
<tr>
<th>(\lambda)</th>
<th>(b_1)</th>
<th>(c_{\lambda_1})</th>
<th>(\ell(\hat{\theta}))</th>
<th>p.e.v</th>
</tr>
</thead>
<tbody>
<tr>
<td>-2.50</td>
<td>-1.000</td>
<td>124.952</td>
<td>309.399</td>
<td>0.100</td>
</tr>
<tr>
<td>-2.29</td>
<td>-1.000</td>
<td>84.230</td>
<td>309.609</td>
<td>0.100</td>
</tr>
<tr>
<td>-2.25</td>
<td>-1.000</td>
<td>78.112</td>
<td>309.602</td>
<td>0.101</td>
</tr>
<tr>
<td>-2.00</td>
<td>-1.000</td>
<td>48.654</td>
<td>309.070</td>
<td>0.101</td>
</tr>
<tr>
<td>-1.50</td>
<td>-0.840</td>
<td>16.878</td>
<td>305.880</td>
<td>0.103</td>
</tr>
<tr>
<td>-1.00</td>
<td>-0.578</td>
<td>5.363</td>
<td>301.346</td>
<td>0.108</td>
</tr>
<tr>
<td>-0.50</td>
<td>-0.274</td>
<td>1.631</td>
<td>296.491</td>
<td>0.113</td>
</tr>
<tr>
<td>0.00</td>
<td>-</td>
<td>0.496</td>
<td>292.100</td>
<td>0.117</td>
</tr>
<tr>
<td>0.50</td>
<td>0.221</td>
<td>0.154</td>
<td>288.432</td>
<td>0.122</td>
</tr>
<tr>
<td>1.00</td>
<td>0.393</td>
<td>0.049</td>
<td>285.433</td>
<td>0.126</td>
</tr>
</tbody>
</table>

Figure 7: Periodogram and log-spectra estimated by the GCM(-1, 4), selected by BIC, and EXP(5) models.
The comparison with the true autoregressive coefficients (reported in the last column) stresses that they are remarkably accurate. Figure 7 displays the centered periodogram and compares the log-spectra fitted by the selected GCM(−1, 4) model and the EXP(5) model, 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.

6 Conclusions

Modelling the log-spectrum has a long tradition in the analysis of univariate time series and leads to computationally attractive likelihood based methods. We have devised a general frequency domain estimation framework within 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 exponential framework 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 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.

References


and long memory autoregressive integrated moving average models, *Journal of the Royal Statistical


2, 217–226.

strum, pseudo-autocovariance, cross-cepstrum, and saphe cracking, in Rosenblatt M. Ed. *Proceedings of


1509-1526.

Holden-Day.


ings of the IEEE*, 65, 1428-1443.


[17] Cleveland, W.S. (1972), The Inverse Autocorrelations of a Time Series and Their Applications, *Techn-
nometrics*, 14, 2, 277–293.


