RESEARCH REPORT SERIES
(Statistics #2011-06)

On the Computation of Autocovariances for Generalized Gegenbauer Processes

Tucker S. McElroy
Scott H. Holan[1]

[1] University of Missouri

Center for Statistical Research & Methodology
Research and Methodology Directorate
U.S. Census Bureau
Washington, D.C. 20233

Report Issued: April 20, 2011

Disclaimer: This report is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

On the Computation of Autocovariances for Generalized Gegenbauer Processes

Tucker S. McElroy[1] and Scott H. Holan[2]

[1] Statistical Research Division, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, tucker.s.mcelroy@census.gov

[2] Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO, 65211-6100, holans@missouri.edu

Abstract

Gegenbauer processes and their generalizations represent an extremely general way of modeling long memory and seasonal long memory; they include ARFIMA, seasonal ARFIMA, and GARMA processes as special cases. Models from this class of processes have been used extensively in economics, finance and in the physical sciences and are thus of widespread interest. Nonetheless, one obstacle to using this class of models is finding explicit formulas for the autocovariances that are valid for all lags. We provide a computationally efficient method of computing these autocovariances by determining the moving average representation of these processes, and also give an asymptotic formula for the determinant of the covariance matrix. The techniques are then illustrated using maximum likelihood estimation to model atmospheric CO_2 data.

Keywords. ARFIMA; Exponential model; FEXP model; GARMA; k-factor GARMA; k-factor GEXP; Long memory; Maximum likelihood; SARFIMA; Seasonality; Spectral density.

Disclaimer This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

1 Introduction

Gegenbauer processes (Gray, Zhang, and Woodward, 1989) provide a flexible way of modeling time series data that exhibit long memory and seasonal long memory. These processes and their generalizations (k-factor GARMA, discussed in Woodward, Cheng, and Gray (1998)) are easily seen to include standard ARFIMA (Hosking, 1981) and seasonal ARFIMA (Porter-Hudak, 1990) processes, as demonstrated below. If such a model is fitted to data using maximum likelihood estimation, it is essential that a fast, convenient method for computing the autocovariances is available. The same need is also present in Bayesian approaches to long memory (see Holan, McElroy, and Chakraborty (2009) and the references therein). This article provides a general method for computing autocovariances that is quite computationally efficient when there are multiple Gegenbauer polynomial terms in the model spectral density (i.e., multiple fractional memory parameters).

Over the past two decades there has been substantial interest in applying models that capture seasonal long-range dependance. In particular, since the initial introduction of the k-factor GARMA (Woodward, Cheng, and Gray, 1998), these models have become pervasive in economics (e.g., Bisaglia et al., 2003; Soares and Souza, 2006), finance (e.g., Ferrara and Guégan, 2000) and in the physical and natural sciences (e.g., Gil-Alana, 2008; Talamantes et al., 2007). Since successful fitting of these models is computationally expensive, our approach is critical for practical implementation.

We briefly mention related recent literature. In special cases, such as a pure ARFIMA(0,d,0) process, the exact autocovariances are known and can be convolved with any "short memory" portions of the model (i.e., those factors of the spectral density that have no poles). See Brockwell and Davis (1991) for the ARFIMA example. An efficient approach to approximating this convolution is achieved through the so-called "splitting method" outlined in Bertelli and Caporin (2002) and Hurvich (2002). This technique can also be extended to GARMA processes (Gray, Zhang, and Woodward, 1989), since there is only one seasonal long memory parameter.

Although there is an extensive literature on long memory time series (see Palma (2007), Beran (2010), and the references therein), this literature does not discuss the computation of autocovariances when multiple fractional memory parameters are present. Instead, the computation of autocovariances associated with single fractional memory parameter models (e.g., ARFIMA models) has been typically addressed; see Doornik and Ooms (2003) and the references therein. One notable exception is given by Bisognin and Lopes (2009), wherein the authors provide results for autocovariances for the SARFIMA model. Specifically, the authors explicitly derive the autocovariance generating function for the case of the SARFIMA$(P,\, D,\, Q)_s$. In the case of a SARFIMA$(p,\, d,\, q) \times (P,\, D,\, Q)_s$ the authors mention that this can be achieved through additional convolution. In principle, the case of multiple seasonal factors can be handled by adding on another convolution for each extra term, i.e., multiple splitting. Unfortunately, this is exceedingly expensive computationally; also since each convolution must be truncated, there is an additional loss of accuracy involved. The method we propose avoids splitting by instead computing the coefficients of the infinite order moving average representation of the Gegenbauer process explicitly through a recursive formula. The autocovariances are computed from these coefficients in the usual way, although this involves an infinite summation that must be truncated. We show how this truncation error can be accurately estimated using an asymptotic formula for the moving average coefficients, which is also fast to compute.

Furthermore, the Gaussian likelihood function requires computation of the determinant of the covariance matrix. This matrix is likely to be ill-conditioned due to multiple poles in the spectral density; instead, we can compute an approximation to this quantity by generalizing the approach of Chen, Hurvich, and Lu (2006) to this problem. Together, our results allow for computationally efficient maximum likelihood and/or Bayesian estimation of a broad class of processes exhibiting seasonal (and trend) long range dependence.

The remainder of this paper proceeds as follows. Section 2 presents the so-called k-factor GEXP model and draws connections to the ARFIMA and SARFIMA models. Our methodology is detailed in Section 3. In particular, this section provides theoretical justification for our computationally efficient approach to calculating the autocovariances. In addition, this section develops an asymptotic formula for calculating the determinant of the covariance matrix. An empirical study is presented in Section 4, illustrating the effectiveness and accuracy of our approach. Section 5 presents an application of our methodology to modeling atmospheric CO_2 data using maximum likelihood. Finally, Section 6 provides concluding discussion, and all proofs are left to the Appendix.

2 The k-GEXP Model

We consider the following k-factor Generalized Exponential model, or k-GEXP, whose spectral density can be written as

$$f(\lambda) = |1-e^{-i\lambda}|^{-2a}\,|1+e^{-i\lambda}|^{-2b}\prod_{l=1}^{K}|1-e^{-i\omega_l}e^{-i\lambda}|^{-2c_l}|1-e^{i\omega_l}e^{-i\lambda}|^{-2c_l}g(\lambda), \tag{1}$$

where the parameters a, b, $c_1,\cdots,c_K$ are each bounded in $(-1/2,1/2)$ in order to guarantee stationarity. The frequencies ω_l are distinct from one another, and not equal to zero or π. When a parameter a, b, or c_l is positive, there is a corresponding pole in the spectral density at frequency zero, π, or ω_l respectively – this is the case of long memory. On the other hand, negative parameters correspond to a zero in the spectrum, and correspond to intermediate memory (or negative memory, also called anti-persistence by some authors; see Beran, (2010) and the references therein). The function g is bounded, and represents the short memory portion of the spectrum; in particular, it corresponds to an EXP(q) model (Bloomfield, 1973) so that

$$g(\lambda) = \exp\left\{\sum_{j=1}^{q} g_j\cos(\lambda j)\right\}\sigma^2 = \exp\left\{g_0 + \frac{1}{2}\sum_{0<|j|\le q} g_j e^{-i\lambda j}\right\}, \tag{2}$$

where $g_{-j}=g_j$. So the innovation variance σ^2 of the model is equal to $\exp(g_0)$. It will be convenient to define $\kappa(z)=\exp\{\frac{1}{2}\sum_{j=1}^{q} g_j z^j\}$ and

$$\beta(z) = (1-z)^{-a}(1+z)^{-b}\prod_{l=1}^{K}(1-e^{-i\omega_l}z)^{-c_l}(1-e^{i\omega_l}z)^{-c_l}\kappa(z), \tag{3}$$

which can be written as $\sum_{k=0}^{\infty} \beta_k z^k$; then $g(\lambda) = |\kappa(e^{-i\lambda})|^2 \sigma^2$ and $f(\lambda) = |\beta(e^{-i\lambda})|^2 \sigma^2$. A model related to (1) was described in less generality, and using a different formulation, in Hsu and Tsai (2009). However, in that context, the authors conduct estimation using log periodogram regression. Importantly, except for the 1-GEXP, the model they propose can not be estimated using a likelihood or Bayesian approach, due to the lack of specification of the autocovariance sequence and the absence of methodology for estimating the determinant.

We claim that (1) is extremely general, including generalized Gegenbauer processes, as well as ARFIMA and seasonal ARFIMA. Recall (Gray, Zhang, and Woodward, 1989) that a Gegenbauer process has spectrum proportional to $|1 - 2ue^{-i\lambda} + e^{-i2\lambda}|^{-2d}$, with $|u| \leq 1$ and $|d| < 1/2$. The polynomial $1 - 2uB + B^2$ has either two real roots or a pair of complex conjugate roots; in both cases, the spectrum takes on the form of (1). The spectrum of a generalized Gegenbauer process contains many such factors multiplied together, which of course is also of the form (1). The seasonal ARFIMA (of which the ARFIMA is a special case) takes the form

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} |1 - e^{-is\lambda}|^{-2D} g(\lambda), \tag{4}$$

where g is allowed to have infinite order (i.e., $q = \infty$) corresponding to an ARMA process, d, D are restricted so as to guarantee stationarity, and s is the seasonal period. Noting that $1 - B^s = (1 - B)U(B)$ with $U(B)$ equal to the product of factors corresponding to the s roots of unity (excepting the one at frequency zero), the pole of f at frequency zero has exponent $d + D$, whereas the other roots of unity generate poles with exponent D. Therefore this corresponds to (1) with $a = d + D$, $b = D$, $c_l = D$, and $\omega_l = \pm 2\pi l/s$, $1 \leq l \leq s/2 - 1$.

3 Computation of the Model

We now turn to the practical issue of computing the autocovariances of (1). The idea is to first compute the Fourier coefficients of the log spectrum, and second to re-express the coefficients of the infinite moving average (MA) representation in terms of these, from which the autocovariances are determined in standard fashion. The reason why this approach is computationally efficient is that each extra multiplicative factor in (1) only adds an extra summand to the Fourier coefficients of the log spectrum – whereas a direct approach at obtaining the MA coefficients would involve an extra discrete convolution for each additional term. Now, defining

$$\theta_j = \frac{1}{\pi} \int_{-\pi}^{\pi} \log f(\lambda) \cos(\lambda j)\, d\lambda \tag{5}$$

for $j \geq 1$, we have $f(\lambda) = \exp\{\sum_{j \geq 1} \theta_j \cos(\lambda j)\} \sigma^2$. Here the innovation variance satisfies $\log \sigma^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda)\, d\lambda$ as usual. The following result tells us how to compute these coefficients (5) at once from the parameters in (1).

Proposition 1 *Given the model (1), the coefficients in (5) are given by*

$$\theta_j = \frac{2}{j} \left\{ a + b(-1)^j + 2 \sum_{l=1}^{K} c_l \cos(\omega_l j) \right\} + g_j. \tag{6}$$

For Proposition 1 to be useful, one must know $\{g_j\}$; either these are known from the presumptive EXP(q) model for the short memory portion of the spectrum, or these coefficients are directly computed from known ARMA parameters. Suppose that we have an invertible ARMA model for the short memory portion, such that $\kappa(z) = \prod_k (1 - \zeta_k z)^{p_k}$ for (possibly complex) reciprocal roots ζ_k of the moving average and autoregressive polynomials; here p_k equals one if k corresponds to a moving average root, but equals negative one if k corresponds to an autoregressive root. Then g_j is given by

$$g_j = 2 \sum_k p_k \zeta_k^j / j. \tag{7}$$

The derivation of (7) can be found in McElroy and Holan (2009) and Hsu and Tsai (2009, Appendix B). Note that if ζ_k is complex, there exists a conjugate factor $\overline{\zeta}_k$ such that their sum is $|\zeta_k|^j 2 \cos(j\omega_k)$, where ω_k is the angular portion. Thus g_j is always real, and readily computed; due to the exponential decay in (7), it is always safe to truncate the sum using a relatively small number of terms (e.g., 100 terms).

Next, recall that we can write (1) as $|\beta(e^{-i\lambda})|^2 \sigma^2$. Then applying the Proposition of Pourahmadi (1984) – also see Hurvich (2002) for treatment of the FEXP case – we obtain

$$\beta_j = \frac{1}{2j} \sum_{k=1}^{j} k \theta_k \beta_{j-k} \tag{8}$$

for $j \geq 1$. This is a convolution, which is expensive to compute; moreover it must be carefully monitored for decay – see Theorem 1 below. We can expect these coefficients to decay slowly, e.g., in the case of an ARFIMA we have $b = 0 = c_l$ and $\beta_j \sim j^{-a-1}$ (Brockwell and Davis, 1991, p. 522).

Once we compute σ^2, the autocovariances are given by the usual formula:

$$\gamma_h = \sigma^2 \sum_{j \geq 0} \beta_j \beta_{j+h} \tag{9}$$

for $h \geq 0$ (and $\gamma_{-h} = \gamma_h$). So the algorithm amounts to the application of (6), (8), and a truncated version of (9). Now for processes where the long memory is pronounced (say a, b, or c_l is close to .5), the decay of the coefficients β_j is extremely slow, and thus computing γ_h via taking a truncation in (9) induces a substantial amount of error. One can take the truncation point farther out, but for fairly pronounced long memory the number of β_js needed becomes computationally prohibitive. (As an example, we found that with $a = .49$, even taking greater than $100,000$ β_js yielded autocovariance values that severely under-estimated the target.) Therefore, in practice, it is necessary to compute the truncation error via asymptotic formulas. We approach this problem by first presenting asymptotic formulas for the β_j sequence, which is a new result. Then we show how to use these approximations to compute an asymptotic form for the truncation error. If at least one of the memory parameters (a, b, $c_1, \cdots, c_K$) are positive, then we can use formula (11) below. However, if they are all negative, the process is intermediate memory and the β_j coefficients will decay rapidly.

Consider $\beta(z)$ in (3), and collect all of the factors associated with the minimal memory exponent α, i.e., $\alpha = \min\{-a, -b, -c_1, \cdots, -c_K\}$, and denote this (possibly non-monomial) polynomial by $\Theta(z)$; the remaining factors are gathered into $\Phi(z)$ (and it necessarily includes $\kappa(z)$ as a factor). Further, let $\Theta(z)$ consist of m factors $\Theta_k(z)$ of the form $(1-\zeta_k^{-1}z)^\alpha$, and denote $\beta(z)/\Theta_k(z)$ by $\beta_{-k}(z)$ (which has its pole at ζ_k removed).

Theorem 1 *Let α be the minimal memory exponent in (1) and (3), and suppose $\alpha < 0$. Then the overall rate of decay of β_j is governed by $\pi_j(\alpha) = \frac{\Gamma(j-\alpha)}{\Gamma(j+1)\Gamma(-\alpha)}$, and*

$$\pi_j(\alpha) = \frac{j^{-\alpha-1}}{\Gamma(-\alpha)}\{1 + O(j^{-1})\}. \tag{10}$$

Letting $E(j) = \sum_{k=1}^m \zeta_k^{-j}\beta_{-k}(\zeta_k)$, the exact asymptotics for β_j are

$$\beta_j = \pi_j(\alpha)E(j)\{1 + o(1)\}. \tag{11}$$

Remark 1 This is a general result, and applies to any $\beta(z)$ given as a product of factors of the form $(1-\zeta^{-1}z)^\gamma$ for unit roots ζ and non-zero $\gamma \in (-1/2, 1/2)$ (times any bounded function). For our purposes (3) is sufficient. It is easy to see that this result generalizes that of Chung (1996); in terms of our notation, that paper's equation (9) yields

$$\beta_j \sim \pi_j(\alpha) \cdot \cos\{(j-\alpha)\omega + \alpha\pi/2\}\, 2^{1+\alpha}\, \sin^\alpha(\omega),$$

where $\beta(z) = (1-e^{-i\omega}z)^\alpha(1-e^{i\omega}z)^\alpha$. The function multiplying $\pi_j(\alpha)$ on the right hand side can be rewritten as $e^{ij\omega}(1-e^{-i2\omega})^\alpha + e^{-ij\omega}(1-e^{i2\omega})^\alpha$, which agrees with (11).

Now, for some cutoff J, we can express (9) as

$$\gamma_h = \sigma^2 \sum_{j=0}^{J-1} \beta_j\beta_{j+h} + \sigma^2 \sum_{j=J}^{\infty} \beta_j\beta_{j+h} = B_J(h) + R_J(h). \tag{12}$$

The first term is computed using the exact β_js via (8), while the second term will use the approximate β_js given in (11). This remainder term – denoted $R_J(h)$ – can be written

$$R_J(h) = \sigma^2 \sum_{j=J}^{\infty} \beta_j\beta_{j+h} \sim \frac{\sigma^2}{\Gamma^2(-\alpha)} \sum_{k,l=1}^{m} \beta_{-k}(\zeta_k)\beta_{-l}(\zeta_l) \sum_{j=J}^{\infty} j^{-\alpha-1}(j+h)^{-\alpha-1}\zeta_k^{-j}\zeta_l^{-j-h}$$

as $J \to \infty$. This remainder term can be approximated by an easily computed expression, which is given in the proposition below. A notation that we shall use repeatedly is

$$A_h = \sum_{k=1}^{m} |\beta_{-k}(\zeta_k)|^2 \zeta_k^{-h}. \tag{13}$$

This expression, (13), is used in both Proposition 2 and Theorem 2.

Proposition 2 *The remainder term $R_J(h)$ for fixed h and $J \to \infty$ is given by*

$$R_J(h) = \left\{ J^{-1-2\alpha} \frac{\sigma^2 F(1+\alpha; 1+2\alpha; 2+2\alpha; -h/J) A_h}{\Gamma^2(-\alpha)(1+2\alpha)} \right\} \{1 + o(1)\}, \tag{14}$$

where $F(1+\alpha; 1+2\alpha; 2+2\alpha; z)$ is the hypergeometric function evaluated at z.

This result is proved using methods discussed in the proof of Theorem 2, and hence its proof is included in the Appendix at the end of that theorem's proof. We proceed with some specific examples.

1-GEXP The 1-GEXP model has only one memory parameter a, which can be associated with frequency $0, \pi$ or $\omega \in (0, \pi)$. In the latter case we can generalize the calculation from Remark 1, using $\beta(z) = (1 - e^{-i\omega} z)^{-c} (1 - e^{i\omega} z)^{-c} \kappa(z)$, and obtain

$$A_h = 2^{1-2c} \sin^{-2c}(\omega) \exp\left\{ \sum_{k=1}^{q} g_k \cos(\omega k) \right\} \cos(h\omega).$$

The frequency zero case is known as the FEXP, and $\beta(z) = (1 - z)^{-a} \kappa(z)$; then $A_h = \kappa^2(1)$. Similarly, the frequency π case has $\beta(z) = (1 + z)^{-b} \kappa(z)$ and $A_h = \kappa^2(-1)(-1)^h$.

2-GEXP The 2-GEXP model has two memory parameters. There are four main cases, depending on the locations of the poles: poles at 0 and π; poles at 0 and $\omega \in (0, \pi)$; poles at π and $\omega \in (0, \pi)$; poles at $\omega_1 \neq \omega_2 \in (0, \pi)$. In the first case $\beta(z) = (1 - z)^{-a} (1 + z)^{-b} \kappa(z)$. If $a > b$ then $A_h = 2^{-2b} \kappa^2(1)$, but if $b > a$ then $A_h = 2^{-2a} \kappa^2(-1)(-1)^h$. If $a = b$ we just sum (this is true for all the cases below): $A_h = 2^{-2b} \kappa^2(1) + 2^{-2a} \kappa^2(-1)(-1)^h$. In the second case $\beta(z) = (1 - z)^{-a} (1 - e^{-i\omega} z)^{-c} (1 - e^{i\omega} z)^{-c} \kappa(z)$. If $a > c$ then $A_h = |1 - e^{i\omega}|^{-4c} \kappa^2(1)$, but when $c > a$ we instead obtain

$$A_h = 2 |1 - e^{i\omega}|^{-2a} |1 - e^{i2\omega}|^{-2c} |\kappa(e^{i\omega})|^2 \cos(h\omega).$$

Thirdly, suppose $\beta(z) = (1 + z)^{-b} (1 - e^{-i\omega} z)^{-c} (1 - e^{i\omega} z)^{-c} \kappa(z)$. If $b > c$ we have

$$A_h = |1 + e^{i\omega}|^{-4c} \kappa^2(-1)(-1)^h,$$

and if $c > b$ we obtain

$$A_h = 2 |1 + e^{i\omega}|^{-2b} |1 - e^{i2\omega}|^{-2c} |\kappa(e^{i\omega})|^2 \cos(h\omega).$$

The final case is the most complicated. Here

$$\beta(z) = (1 - e^{-i\omega_1} z)^{-c_1} (1 - e^{i\omega_1} z)^{-c_1} (1 - e^{-i\omega_2} z)^{-c_2} (1 - e^{i\omega_2} z)^{-c_2} \kappa(z).$$

If $c_1 > c_2$ then we obtain

$$A_h = 2|1 - e^{i2\omega_1}|^{-2c_1} |1 - e^{i(\omega_1 - \omega_2)}|^{-2c_2} |1 - e^{i(\omega_1 + \omega_2)}|^{-2c_2} |\kappa(e^{i\omega_1})|^2 \cos(h\omega_1).$$

Clearly if $c_2 > c_1$, we just interchange ω_1 and ω_2 and c_1 and c_2 in the above formula.

SFEXP We refer to a process following (4), with short memory EXP dynamics (i.e., $g(\lambda) \sim$ EXP(q)), by a Seasonal Fractional EXP model, or SFEXP for short. There are s poles at various seasonal frequencies, but we have $b = c_k$ for $k \leq s/2 - 1$. It is then straightforward to apply (6) and (8) to obtain β_j for $j < J$ (the small index case). For the large index case, clearly we need to compute A_h; also recall that α is the negative of the larger of a and b, the two memory parameters. Now $\beta(z) = (1 - z)^{-a} U^{-b}(z) \kappa(z)$, and note that $U(z)/(1 - \zeta^{-1} z)|_{z=\zeta} = s/(1 - \zeta)$ for any ζ unit root of $U(z)$. So if $a > b$, then

$$\beta_j \sim \pi_j(-a)\, s^{-b}\, \kappa(1)$$

follows from (11), and $A_h = s^{-2b} \kappa^2(1)$. Alternatively, if $b > a$ we obtain $\beta_j \sim \pi_j(-b)$ times the sum over $s - 1$ unit roots ζ for $U(z)$ of $U^{-b}(z)(1 - z)^{-a} \kappa(z)(1 - \zeta^{-1} z)^b \zeta^{-j}$ evaluated at $z = \zeta$, i.e., $(1 - \zeta)^{b-a} \kappa(\zeta) s^{-b} \zeta^{-j}$. Focusing on the case $s = 12$, we can simplify to

$$A_h = 12^{-2b} \left\{ 2^{2(b-a)} \kappa^2(-1)(-1)^h + 2 \sum_{k=1}^{5} |1 - e^{i\pi k/6}|^{2(b-a)} |\kappa(e^{i\pi k/6})|^2 \cos(\pi k h/6) \right\}.$$

Finally, in this case $a \neq b$ unless either d or D equals zero.

It is also of interest to know the asymptotic behavior of the autocovariances γ_h as $h \to \infty$. This can be useful in computations where many lags are involved, and a quicker formula is needed.

Theorem 2 *Let α be the minimal memory exponent in (1) and (3), and suppose $\alpha < 0$. Then the autocovariances are asymptotically given by*

$$\gamma_h = h^{-2\alpha - 1} \frac{\sigma^2 \Gamma(1 + 2\alpha)}{\Gamma(-\alpha)\Gamma(1 + \alpha)} A_h \{1 + o(1)\} \tag{15}$$

as $h \to \infty$, where A_h is defined in (13).

Remark 2 In the formula (15), specific information about the long memory poles at non-zero frequencies is located in the A_h term – see (13) – which distinguishes this asymptotic result from the more basic FEXP and ARFIMA processes. Of course, estimation of GARMA processes can be handled using the same result, since (13) holds for GARMA (only with the form of $\beta(z)$ being somewhat different through the specification of $\kappa(z)$).

Remark 3 In a sense (15) is a special case of (14) when $h \to \infty$ and $P = h/J \to \infty$. In the last line of the proof, we let $P \to 0$ in the integral and use $P = J/h$ to obtain $h^{-1-2\alpha}\Gamma(-\alpha)\Gamma(1+2\alpha)/\Gamma(1+\alpha)$, again via 3.194.3 of Gradshteyn and Ryzhik (1994), which produces (15).

We proceed by considering a few specific examples.

1-GEXP First if $\beta(z) = (1 - e^{-i\omega} z)^{-c}(1 - e^{i\omega} z)^{-c}\kappa(z)$ then $\gamma_h \sim h^{2c-1}\frac{\sigma^2\Gamma(1-2c)}{\Gamma(c)\Gamma(1-c)}2\{2\sin(\omega)\}^{-2c}\times$ $\cos(h\omega)|\kappa(e^{i\omega})|^2$. Using $\Gamma(c)\Gamma(1-c) = \pi/\sin(\pi c)$ and $\Gamma(h+2c)/\Gamma(h+1) \sim h^{2c-1}$, we see this agrees exactly with the formula of Chung (1996). If the pole is at frequency zero or π, we obtain $\kappa^2(1)$ and $\kappa^2(-1)(-1)^h$ respectively, multiplied each by $\sigma^2\Gamma(1-2c)h^{2c-1}/\{\Gamma(c)\Gamma(1-c)\}$.

2-GEXP Let $\xi(\alpha) = \sigma^2\Gamma(1+2\alpha)/\{\Gamma(-\alpha)\Gamma(1+\alpha)\}$, so that $\gamma_h = h^{-2\alpha-1}\xi(\alpha)A_h$. The values of A_h have been previously computed for the various cases of the 2-GEXP.

SFEXP The expression for A_h given after Proposition 2 can be used when $s = 12$, and γ_h is easily obtained.

In this way the autocovariances can be efficiently computed. To summarize, the procedure is the following:

Step 1: determine all parameters of (1);

Step 2: compute a sufficient number of θ_j via (6);

Step 3: compute a sufficient number of β_j via (8);

Step 4: compute $B_J(h)$ for h small via (12);

Step 5: compute A_h for all desired h via (13);

Step 6: compute $R_J(h)$ for h small via (14);

Step 7: compute γ_h for h large via (15).

This necessarily involves some choices, such as the cutoff between small and large h, and the truncation level J. In practice it makes little difference how these are chosen, but it is recommended that J be taken large relative to the lag cutoff, e.g., $J \geq 2000$ and the h cutoff at 100. Additionally, in many cases, the number of desired lags is sufficiently small enough that γ_h can be calculated for all h without appealing to (15) (i.e., omitting Step 7 of the above procedure).

For maximum likelihood (and Bayesian) estimation, it is also necessary to compute the determinant of the covariance matrix $\Sigma_n(f)$. In general, this can be computationally expensive. One path to computing this determinant is to use the Durbin-Levinson algorithm applied to the autocovariance function (see Brockwell and Davis, 1991). However, this can be computationally infeasible for large sample sizes (Chen et al., 2006). Alternatively, an approximate formula can be obtained from Theorem 5.47 of Böttcher and Silbermann (1999):

Proposition 3 *Considering the model (1) we have:*

$$|\Sigma_n(f)| \sim \sigma^{2n}\, n^{a^2+b^2+2\sum_{l=1}^{K} c_l^2}\, E(f),$$

where the constant $E(f)$ is equal to

$$E(f) = \exp\left[\sum_{j=1}^{q} j g_j^2/4 + \sum_{j=1}^{q} g_j \left\{a + b(-1)^j + 2\sum_{l=1}^{K} c_l \cos(\omega_l j)\right\}\right]$$

$$\cdot\, 2^{-2ab} \prod_{l=1}^{K} \{2 - 2\cos(\omega_l)\}^{-2ac_l} \{2 + 2\cos(\omega_l)\}^{-2bc_l} \{2 - 2\cos(2\omega_l)\}^{-c_l^2}$$

$$\prod_{m>l} \{2 - 2\cos(\omega_m - \omega_l)\}^{-2c_m c_l} \{2 - 2\cos(\omega_m + \omega_l)\}^{-2c_m c_l}$$

$$\cdot\, \frac{G^2(1-a)}{G(1-2a)} \frac{G^2(1-b)}{G(1-2b)} \prod_{l=1}^{K} \frac{G^4(1-c_l)}{G^2(1-2c_l)}.$$

Here G is the Barnes G function (Böttcher and Silbermann, 1999) given by

$$G(z+1) = (2\pi)^{z/2} \exp[-\{z(z+1) + \gamma z^2\}/2] \prod_{n\geq 1} \left\{(1 + z/n)^n e^{-z+z^2/(2n)}\right\}$$

with γ Euler's constant $\approx .57721$.

The proof is omitted, as the result directly follows from Böttcher and Silbermann (1999). Altogether, this provides a method for the fast and accurate evaluation of the Gaussian likelihood function for (1).

In summary, the preceding algorithm can be used to construct the authocovariance matrix needed to compute the Gaussian likelihood. In some cases it is possible to invert the autocovariance matrix and to find its' determinant directly using standard software packages (e.g. R or Matlab - R Development Core Team, (2010); Mathworks, Inc. (2010)). When this is not possible, inversion can be handled using a conjugate gradient (or pre-conditioned conjugate gradient) approach (Golub and Van Loan, 1996) and the log determinant can be approximated using Proposition 3.

4 Empirical Results

This section discusses the accuracy of the autocovariance approximations and asymptotic formula for the determinant (Proposition 3). Exact quantification of the error necessarily depends on the ability able to compute the true autocovariance sequence and true determinant. Unfortunately this is not always possible.

In some cases, such as fractional Gaussian noise (i.e. FGN or ARFIMA$(0, d, 0)$), the autocovariances can be computed exactly and the determinant calculated numerically. In this case we computed the log determinant numerically using the autocovariance obtained from (12), the log determinant obtained from Proposition 3, and the log determinant obtained numerically from the exact autocovariances. In all cases, even for $d = .45$ (and J=5000), the difference between all three methods is less than 10^{-5}. (A comprehensive breakdown of these results are available upon request.)

In the case of the FEXP(q) model, the autocovariance sequence can not be computed exactly. Nevertheless, the autocovariance sequence can be computed up to any degree of accuracy using the so-called splitting method (cf. Section 1 and Bertelli and Caporin (2002)). Similar to the case of FGN, we computed the log determinant numerically using the autocovariances obtained from (12), the log determinant obtained from Proposition 3, and the log determinant obtained numerically from the autocovariances obtained using the splitting method. In all cases, for various short memory specifications g and even for $d = .45$ (J=5000), the difference between all three methods was less than 10^{-4}. (A comprehensive breakdown of these results are available upon request.)

In principal, the autocovariances for 1-GEXP case can be computed to any degree of accuracy. However, this would require use of the splitting method and the autocovariance function associated with the Gegenbauer process (i.e., one seasonal long memory factor). Alternatively, a formula for approximating the autocovariance of the Gegenbauer process has been derived by Chung (1996) and has a complicated form involving Legendre functions that entail recursive calculations. In order to avoid this complex form, Chung (1996) also provides an approximate asymptotic formula that is seen to be a special case of Theorem 1 (cf. Remark 1).

In the case of the 1-GEXP, 2-GEXP and SFEXP the autocovariances sequence can not be computed exactly. Therefore, in practice, the true autocovariance from these models are unknown. Additionally, use of the splitting method for the 1-GEXP case would require an explicit formula for the autocovariances associated with the Gegenbauer process. Specifically, this would require the autocovariance function of a GARMA(0,0) model, which has a complicated form (Chung, 1996). Therefore in order to assess the accuracy of our method, in the case of these models, we computed the log determinant numerically using the autocovariance obtained from (12) and the log determinant obtained from Proposition 3. Although, in principal, the autocovariances for the multiple memory parameter case could be obtained by numerically calculating the inverse Fourier transform of the spectrum, this numerical integration can be prohibitively slow and becomes unstable, even for "moderate" size memory parameters. As a consequence, in practice, this estimate is subject to numerical error and often not obtainable; thus, it is not included in our comparison.

As demonstrated in Tables 1 and 2 the approximations for the 1-GEXP closely agree (i.e. for

$J \geq 10,000$ and $c \leq .35$ the difference is on the order of 10^{-4}). In fact, by increasing J, even for $c = .45$ the difference in the approximations can be made arbitrarily small. In the case of the 2-GEXP, Tables 3 and 4 provide a sense of the accuracy of the approximations. Specifically, in this case, unless both memory parameters are pronounced (i.e. both c_1 and c_2 are greater than or equal to .4), the difference in the approximations can be made on the order of 10^{-2} (for moderate size J).

The SFEXP exhibits similar behavior to the 2-GEXP model. In this case, the trend long-memory parameter equals $d + D$ whereas the seasonal long-memory parameter is given by D. As demonstrated in Tables 5 and 6, when both the trend $(d + D)$ and seasonal (D) long-memory parameters are larger than .3 approximation accuracy is diminished. However, for moderate size seasonal long-memory parameter and substantial trend long memory parameter there is exceptional agreement between the independent computations (i.e., a difference on the order of 10^{-2} for $J \geq 5000$), indicating excellent accuracy.

As expected, when models have multiple memory parameters approaching the nonstationary region of .5 there is decreased accuracy in the approximation of the autocovariance sequence. Nonetheless, it is important to note that no exact formulas exist for the autocovariances associated with seasonal long memory models having multiple memory parameters. Further, the cases where our method suffers from loss of accuracy are exactly the cases that are intractable using standard numerical approaches (i.e., numerical calculation of the inverse Fourier transform of the spectrum). In this context, our approach allows calculation of autocovariance sequences, with minimal loss of accuracy, that otherwise would not be possible. Additionally, the comparison we have conducted is "cumulative", in the sense that the error in the autocovariance sequence is determined through the log determinant rather than by comparison of each individual autocovariance (i.e. for each lag), the latter being impossible to calculate directly in many situations. Finally, by coupling (12) with (15) computational efficiency can be increased for large sample sizes and/or substantial memory parameters.

5 Mauna Loa Data

To illustrate the utility of our approach we model 382 monthly atmospheric CO_2 measurements collected at the summit of Mauna Loa in Hawaii beginning in March 1958 (Keeling et al., 1989). This data was previously analyzed by Woodward et al. (1998) using a 2-factor GARMA model for the second differences of the atmospheric CO_2 data. Figures 1 and 2 display the original data, its sample autocorrelations, the log periodogram of the second differenced data and the sample autocorrelation of the second differenced data. Looking at the log periodogram, it is immediately apparent that the spectrum possesses multiple peaks. To accommodate this behavior we fit a 2-GEXP model using maximum likelihood estimation.

The goal of our analysis is to demonstrate the effectiveness of embedding our autocovariance computations into a maximum likelihood analysis. As such, no efforts have been made in terms of formal model selection. To this end, we fit a 2-GEXP(4) model with unknown peak frequencies. Figure 3 displays our estimated model spectrum with the log periodogram plus γ superimposed. The value $\gamma = .57211$ is the Euler constant and is added to the log periodogram, as this forms an unbiased estimate of the log spectrum (see Percival and Walden (2000) for a comprehensive discussion).

Several salient features of our analysis are important to note. First, the estimated peak frequencies are $\boldsymbol{\omega} = (\omega_1, \omega_2) = (.5239,\ 1.048)$. These correspond to 12- and 6-month cycles and corroborates the analysis of Woodward et al., (1998). Additionally, the corresponding memory parameters $\mathbf{c} = (c_1,\ c_2) = (.4972,\ .4970)$ and also coincide with the analysis of Woodward et al., (1998). The associated standard errors can be obtained from the estimated inverse Hessian. For the peak frequencies the standard errors are 1.42×10^{-7} and 6.66×10^{-7} for ω_1 and ω_2 respectively. The standard errors for the memory parameters are 5.33×10^{-6} and 4.29×10^{-6} for c_1 and c_2 respectively. Finally, the short memory portion of the model is given by $(g_0, g_1, g_2, g_3, g_4) = (-2.162, -1.188, -1.175, 1.150)$ with standard error $(1.54 \times 10^{-2},\ 4.99 \times 10^{-2},\ 3.17 \times 10^{-2},\ 1.70 \times 10^{-2},\ 1.22 \times 10^{-2})$ and the estimated mean (of the twice differenced data) is equal to .198, with standard error 3.98×10^{-4}.

6 Conclusion

Flexible modeling of seasonal long-range dependent processes has been severely hampered due to the lack of computationally efficient methods for calculating the associated model autocovariances. Additionally, for seasonal long-memory models, approaches to approximating the determinant of the autocovariance matrix needed for evaluating the exact Gaussian likelihood have been lacking. As a result, in general, generalized Gegenbauer processes have experienced limited use. Further, in their limited usage, these models have been necessarily estimated using log periodogram regression or Whittle approximations to the Gaussian likelihood.

The approach presented here allows for fast accurate computation of the autocovariances for seasonal long-memory models having multiple memory parameters. As a consequence, flexible models for long-range dependent data can be estimated using exact likelihood or Bayesian methods. To assess the accuracy of our method we presented the results of an empirical study that compares independent estimates of the log determinant. The comparisons are achieved using the estimated autocovariance sequence directly and through the proposed asymptotic approximation to the log determinant. In general, we find that the estimated autocovariance sequences produce log determinants that agree with high accuracy to their asymptotic approximation counterpart.

Additionally, we illustrate the utility of the autocovariance computation approach by embedding our approximations in a maximum likelihood analysis for the Mauna Loa data. The results obtained from this analysis are seen to provide sensible estimates that corroborate the analysis of Woodward et al. (1998). This analysis, along with the results of the empirical study, demonstrate the effectiveness of our computational formulas, allowing for the estimation of a general class of models using exact maximum likelihood and Bayesian methodology.

Appendix - Proofs of Results

Proof of Proposition 1. Taking the logarithm of (1) yields

$$\log f(\lambda) = -a\log(2-2\cos\lambda) - b\log(2+2\cos\lambda) - \sum_{l=1}^{K} c_l \log\left[\{2-2\cos(\omega_l-\lambda)\}\{2-2\cos(\omega_l+\lambda)\}\right] + \log g(\lambda).$$

Using integration by parts we obtain (for $j \geq 1$)

$$\frac{1}{\pi}\int_{-\pi}^{\pi} \log(2-2\cos\lambda)\,\cos(\lambda j)\,d\lambda = -\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin(\lambda j)\,\sin\lambda}{j(1-\cos\lambda)}\,d\lambda.$$

For rigor, the integral should be broken into two integrals over $[-\pi, 0)$ and $(0, \pi]$ and re-assembled, which is used to show that the boundary terms in the integration by parts amount to zero. Next, letting Ω denote the unit circle in the complex plane we have

$$-\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin(\lambda j)\,\sin\lambda}{j(1-\cos\lambda)}\,d\lambda = -\frac{1}{2\pi i j}\int_{\Omega} z^{-(j+1)}(z+1)\sum_{k=0}^{2j-1} z^k\,dz = -\frac{1}{jj!}\frac{\partial^j}{\partial z^j}\left\{(z+1)\sum_{k=0}^{2j-1} z^k\right\}|_{z=0} = -\frac{2}{j},$$

since there is a pole of order $j+1$ at zero, making use of the residue formula (see Henrici (1974)). Similarly,

$$\begin{aligned}
\frac{1}{\pi}\int_{-\pi}^{\pi} \log(2+2\cos\lambda)\,\cos(\lambda j)\,d\lambda &= \frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin(\lambda j)\,\sin\lambda}{j(1+\cos\lambda)}\,d\lambda \\
&= \frac{1}{2\pi i j}\int_{\Omega} z^{-(j+1)}(z-1)\sum_{k=0}^{2j-1} (-z)^k\,dz \\
&= \frac{1}{jj!}\frac{\partial^j}{\partial z^j}\left\{(z-1)\sum_{k=0}^{2j-1} (-z)^k\right\}|_{z=0} \\
&= -\frac{2(-1)^j}{j}.
\end{aligned}$$

Finally, suppose that $\omega \neq 0, \pi$.

$$\begin{aligned}
\frac{1}{\pi}\int_{-\pi}^{\pi} \log\{2-2\cos(\lambda+\omega)\}\,\cos(\lambda j)\,d\lambda &= -\frac{1}{\pi}\int_{-\pi}^{\pi} \frac{\sin(\lambda j)\,\sin(\lambda+\omega)}{j\{1-\cos(\lambda+\omega)\}}\,d\lambda \\
&= -\frac{1}{2\pi i j}\int_{\Omega} \frac{(ze^{i\omega}+1)z^{-(j+1)}(z^{2j}-1)}{ze^{i\omega}-1}\,dz.
\end{aligned}$$

The integrand has a simple pole at $e^{-i\omega}$, unless $e^{i\omega 2j} = 1$ (in which case there is a cancelation). The residue is $2(e^{-i\omega j} - e^{i\omega j})$, which gets halved because it lies on Ω. The pole at zero is of order

$j+1$, and its residue comes out to be $2e^{i\omega j}$. As a final result we obtain $-2\cos(\omega j)/j$ if $e^{i\omega 2j} \neq 1$, and $-2/je^{i\omega j}$ otherwise. However, in this latter case we have $\cos(\omega j) = e^{i\omega j}$, so that $-2\cos(\omega j)/j$ is a valid formula for both cases. Finally, note that these roots $\omega \in (0,\pi)$ found in (1) always occur in pairs, which accounts for the doubling of these terms in (6). This concludes the derivation. □

Proof of Theorem 1. As a first step, consider the case that $\Theta(z)$ consists of a single factor, i.e., $\Theta(z) = (1-\zeta^{-1}z)^{\alpha}$. Expanding in z yields (see Theorem 13.2.1 of Brockwell and Davis (1991)) $\Theta(z) = \sum_{j\geq 0} \pi_j(\alpha)\zeta^{-j}z^j$, with the asymptotic form of $\pi_j(\alpha)$ given by (10). Now $\Phi(z)$ consists of $\kappa(z)$, whose coefficients decay at exponential rate (see Hurvich, 2002), multiplied by various factors $(1-\xi^{-1}z)^{\gamma}$ where $\gamma > \alpha$. We proceed by induction on the number of these factors; first suppose that $\Phi = \kappa$. Then

$$\beta_j = \sum_{k=0}^{j} \pi_k(\alpha)\zeta^{-k}\phi_{j-k} \tag{A.1}$$

follows from $\beta(z) = \Theta(z)\Phi(z)$, with ϕ_j the coefficients of Φ. We split the sum (A.1) into four parts using a sequence $\tau(j)$ such that $1/\tau(j) + \tau(j)/j \to 0$: sum over $0 \leq k < \tau(j) - 1$ (part I); sum over $\tau(j) \leq k < j/2$ (part II); sum over $j/2 \leq k < j - \tau(j)$ (part III); sum over $j - \tau(j) \leq k \leq j$ (part IV). For part I, we have ϕ_{j-k} is asymptotic to ϕ_j times a function of exponential decay, so that $\phi_{j-k}/\phi_j \sim r^{-k}$ (for some $r < 1$) times other bounded functions. Then the sum in part I is asymptotic to $\phi_j \sum_{k=0}^{\tau(j)} \pi_k(\alpha)\zeta^{-k}r^{-k}$ (leaving out the bounded functions, since they don't affect the argument), which tends to zero since ϕ_j is dominant. For part II, we note that $\pi_k(\alpha)$'s asymptotic form (10) can be substituted since $k \geq \tau(j) \to \infty$. The resulting sum converges (because ϕ_{j-k} has exponential decay), but since it is a tail sum starting at $\tau(j)$ part II converges to zero. For parts III and IV we first make a change of variable so that the summands look like $\pi_{j-k}(\alpha)\zeta^{k-j}\phi_k$. Since $\pi_{j-k}(\alpha)/\pi_j(\alpha) \to 1$ for $k \leq \tau(j)$, part III is asymptotic to $\pi_j(\alpha)\zeta^{-j}$ times the tail sum of a convergent series, and hence is $o(\pi_j(\alpha))$. Only part IV is left, which is asymptotic to $\pi_j(\alpha)\zeta^{-j}\sum_{k=0}^{\tau(j)} \zeta^k\phi_k$, and this sum tends to $\Phi(\zeta)$. But this is (11).

Now suppose that $\Phi(z)$ consists of $\kappa(z)$ times n factors of the form $(1-\xi^{-1}z)^{\gamma}$, and consider $\beta(z) = (1-\zeta^{-1}z)^{\alpha}\Phi(z)(1-\xi^{-1}z)^{\gamma}$. Group the first term with $\Phi(z)$, and call this $\Delta(z)$; then by the induction hypothesis, we know that its coefficients δ_j satisfy (11), i.e., $\delta_j \sim \pi_j(\alpha)\zeta^{-j}\Phi(\zeta)$. The second term $(1-\xi^{-1}z)^{\gamma}$ has coefficients ρ_j, and we know that $\rho_j = \pi_j(\gamma)\xi^{-j}$. Now similar to (A.1), $\beta_j = \sum_{k=0}^{j} \delta_k\rho_{j-k}$ and we can break the sum into four parts as in the base case. Although ρ_j does not have exponential decay, we can use the fact that $\rho_j/\pi_j(\alpha) \to 0$ – as well as the summability of $\rho_k\zeta^{-k}\xi^{-k}$, as ξ and ζ are unit roots – to conclude that the first three parts are $o(\pi_j(\alpha))$. Part IV is asymptotic to $\delta_j \sum_{k=0}^{\tau(j)} \xi^k\rho_k \sim \pi_j(\alpha)\zeta^{-j}\Phi(\zeta)(1-\xi^{-1}\zeta)^{\gamma}$, as desired.

Finally, we must induct on the number of factors in $\Theta(z)$. We have established the base case already (one factor), so suppose that (11) holds for m factors, i.e., $\beta(z) = \Theta(z)(1-\zeta_{m+1}^{-1}z)^{\alpha}$, and θ_j satisfies (11). Since the coefficients of this latter factor are exactly $\pi_j(\alpha)\zeta_{m+1}^{-j}$, (A.1) yields

$\beta_j = \sum_{k=0}^{j} \theta_k \pi_{j-k}(\alpha)\zeta_{m+1}^{k-j}$. This can be decomposed into four portions as before, but now only parts II and III are negligible. This is because for both these summations we can asymptotically pull out a $\pi_j(\alpha)$ term, leaving the tail sum of a convergent series (again, because the presence of the unit roots ζ_l bring about convergence), making them both $o(\pi_j(\alpha))$. For part I we obtain the asymptotic $\zeta_{m+1}^{-j}\pi_j(\alpha)\sum_{k=0}^{\tau(j)} \theta_k \zeta_{m+1}^{k} \sim \pi_j(\alpha)\zeta_{m+1}^{-j}\Theta(\zeta_{m+1})$. For part IV we first observe that

$$\frac{\theta_{j-k}}{\theta_j} \sim \frac{\sum_{l=1}^{m} \zeta_l^{-j+k}\Theta_{-l}(\zeta_l)}{\sum_{l=1}^{m} \zeta_l^{-j}\Theta_{-l}(\zeta_l)},$$

so that the asymptotic is $\theta_j \sum_{k=0}^{\tau(j)} \sum_{l=1}^{m} \zeta_l^{-j+k}\Theta_{-l}(\zeta_l)\pi_k(\alpha)\zeta_{m+1}^{-k} / \sum_{l=1}^{m} \zeta_l^{-j}\Theta_{-l}(\zeta_l)$, which is $\pi_j(\alpha)\sum_{l=1}^{m} \zeta_l^{k}\Theta_{-l}(\zeta_l)(1-\zeta_l/\zeta_{m+1})^{\alpha}$. Putting parts I and IV together yields

$$\frac{\beta_j}{\pi_j(\alpha)} \sim \zeta_{m+1}^{-j}\Theta(\zeta_{m+1}) + \sum_{l=1}^{m} \zeta_l^{k}\Theta_{-l}(\zeta_l)(1-\zeta_l/\zeta_{m+1})^{\alpha} = \sum_{l=1}^{m+1} \zeta_l^{k}\beta_{-l}(\zeta_l).$$

We make some further comments on the order of approximation. By examining the Gamma function and using results of Gradshteyn and Ryzhik (1994), it is possible to show that $\pi_j(\alpha) = j^{-(1+\alpha)}\Gamma^{-1}(-\alpha)\{1 + O(j^{-1})\}$. In the analysis of β_j, there are error terms that are $o(\pi_j(\alpha))$ and cannot be improved in general; thus $\beta_j = \pi_j(\alpha)E(j)\{1 + o(1)\}$. This concludes the proof. $\square$

Proof of Theorem 2. Without loss of generality set $\sigma^2 = 1$. Let $\tau(h)$ play the same role as in the proof of Theorem 1, and break the sum in γ_h into two parts: $j \leq \tau(h)$ and $j > \tau(h)$. The first portion is

$$\sum_{j=0}^{\tau(h)} \beta_j\beta_{j+h} = \sum_{j=0}^{\tau(h)} \beta_j (j+h)^{-(1+\alpha)}\Gamma^{-1}(-\alpha)E(j+h)\{1+o(1)\}$$
$$= h^{-(1+\alpha)}\Gamma^{-1}(-\alpha)\sum_{j=0}^{\tau(h)} \beta_j (1+j/h)^{-(1+\alpha)}E(j+h)\{1+o(1)\}.$$

The $o(1)$ term is as $h \to \infty$. Because of the asymptotic rate of decay of β_j, a bound on the divergence of the sum is $\tau(h)^{-\alpha} = o(h^{-\alpha})$. Hence the overall bound on the first term is $O(h^{-(1+2\alpha)})$. Turning to the second term, we have

$$\sum_{j>\tau(h)} \beta_j\beta_{j+h} = \Gamma^{-2}(-\alpha)\sum_{j>\tau(h)} j^{-(1+\alpha)}(j+h)^{-(1+\alpha)}E(j)E(j+h)\{1+o(1)\}. \tag{A.2}$$

Here there is error that is $o(1)$ as $j \to \infty$ and as $j+h \to \infty$, which amounts to just $o(1)$ as $h \to \infty$, since $j > \tau(h)$. Now writing out $E(j)$ and $E(j+h)$, we must compute $\sum_{j>\tau(h)} j^{-(1+\alpha)}(j+h)^{-(1+\alpha)}\zeta_l^{-j}\zeta_k^{-j}$ for unit roots ζ_l, ζ_k. If $\zeta_l\zeta_k \neq 1$, they make the sum oscillatory fostering convergence, such that we obtain a bound of $O(h^{-(1+\alpha)})$; else if $\zeta_l\zeta_k = 1$ the sum decays at the slower rate of $h^{-(1+2\alpha)}$, as shown below. Letting $k_0 = \tau(h)$ and $k_p = hp$ for $p \geq 1$, we can rewrite as

$$h^{-2(1+\alpha)}\left\{\sum_{p\geq 1}\sum_{j=k_{p-1}+1}^{k_p} (j/h)^{-(1+\alpha)}(1+j/h)^{-(1+\alpha)}\right\} \sim h^{-(1+2\alpha)}\sum_{p\geq 1}\int_{p-1}^{p} x^{-(1+\alpha)}(1+x)^{-(1+\alpha)}\,dx,$$

since it is a Riemann sum. We also use the fact that $\tau(h)/h \to 0$. Now $\int_0^\infty x^{-(1+\alpha)}(1+x)^{-(1+\alpha)}\,dx = \Gamma(-\alpha)\Gamma(1+2\alpha)/\Gamma(1+\alpha)$ by 3.194.3 of Gradshteyn and Ryzhik (1994). Next, we must determine how many unit root pairs satisfy $\zeta_l\zeta_k = 1$, i.e., are conjugate to one another. If the root is ± 1, it is self-conjugate. Otherwise, there is always a conjugate root present in $\beta(z)$ since these factors come in pairs. In any event, the only terms in $E_j E_{j+h}$ that need be considered are those such that $\zeta_l\zeta_k = 1$ for all $1 \leq l, k \leq m$, which is

$$\sum_{l,k=1}^{m} 1_{\{\zeta_l\zeta_k=1\}}\beta_{-l}(\zeta_l)\beta_{-k}(\zeta_k)\zeta_k^{-h} = \sum_{k=1}^{m} |\beta_{-k}(\zeta_k)|^2\zeta_k^{-h} = A_h.$$

Putting this together with $h^{-(1+2\alpha)}\Gamma(-\alpha)\Gamma(1+2\alpha)/\Gamma(1+\alpha)$ yields the stated result (15).

Here we also prove Proposition 2, using the same techniques. Let $J = hP$, and suppose h is fixed but J and P are tending to infinity. Then $R_J(h)$ can be expressed by (A.2), but with $\tau(h)$ replaced by J. Similar arguments (but now taking a J asymptotic instead of an h asymptotic) allow us to focus on unit roots such that $\zeta_k\zeta_l = 1$, and the summation for such k, l is

$$\begin{aligned}\sum_{j=J}^{\infty} j^{-\alpha-1}(j+h)^{-\alpha-1} &= J^{-2-2\alpha}\sum_{r\geq 1}\sum_{j=rJ}^{(r+1)J}(j/J)^{-1-\alpha}(j/J+1/P)^{-1-\alpha} \\ &\approx J^{-1-2\alpha}\int_1^\infty x^{-1-\alpha}(x+1/P)^{-1-\alpha}\,dx.\end{aligned}$$

The approximation used here is the Riemann integration approximation for bounded integrands, so the error is $O(J^{-2-2\alpha})$. However, the approximation error from discounting the $\zeta_k\zeta_l \neq 1$ terms is lower order, namely $o(J^{-1-2\alpha})$. Using 3.194.2 of Gradshteyn and Ryzhik (1994), the integral is equal to

$$P^{1+2\alpha}\int_P^\infty y^{-1-\alpha}(1+y)^{-1-\alpha}\,dx = \frac{F(1+\alpha;1+2\alpha;2+2\alpha;-1/P)}{1+2\alpha},$$

using the change of variable $y = Px$. This completes the proof. $\square$

References

[1] Beran, J. (2010) Long-range dependence. *WIREs Computational Statistics* **2**, 26–35.

[2] Bertelli, S. and Caporin, M. (2002) A note on calculating autocovariances of long-memory processes. *Journal of Time Series Analysis* **23**, 503–508.

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

[4] Bisognin, C. and Lopes, S.R.C. (2009). Properties of seasonal long memory processes. *Mathematical and Computer Modelling.* **49**, 1837–1851.

[5] Bloomfield, P. (1973). An exponential model for the spectrum of a scalar time series. *Biometrika* **60**, 217–226.

[6] Böttcher, A. and Silbermann, B. (1999) *Introduction to Large Truncated Toeplitz Matrices.* New York: Springer-Verlag.

[7] Brockwell, P. and Davis, R. (1991) *Time Series: Theory and Methods.* New York: Springer.

[8] Chen, W., Hurvich, C., and Lu, Y. (2006) On the correlation matrix of the discrete Fourier transform and the fast solution of large Toeplitz systems for long-memory time series. *Journal of the American Statistical Association* **101**, 812–822.

[9] Chung, C. (1996) Estimating a generalized long memory process. *Journal of Econometrics*, **73**, 237–259.

[10] Doornik, J. and Ooms, M. (2003) Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Computational Statistics and Data Analysis*, **42**, 333–348.

[11] Ferrara, L. and D. Guégan. (2000) Forecasting financial time series with generalized long memory processes, in *Advances in Quantitative Asset Management*, C.L. Dunis eds., 319-342, Kluwer Academic Publishers.

[12] Gil-Alana, L.A. (2008) Cyclical long-range dependence and the warming effect in a long temperature time series. *International Journal of Climatology.* **28** 1435–1443.

[13] Gradshteyn, I. and Ryzhik, I. (1994) *Table of Integrals, Series, and Products.* Orlando: Academic Press.

[14] Gray, H., Zhang, N., and Woodward, W. (1989) On generalized fractional processes. *Journal of Time Series Analysis* **10**, 233-257.

[15] Golub, G.H. and Van Loan, C.F. (1996) *Matrix Computations* (3rd ed.), Baltimore: Johns Hopkins University Press.

[16] Holan, S., McElroy, T., and Chakraborty, S. (2009) A Bayesian Approach to Estimating the Long Memory Parameter. *Bayesian Analysis* **4**, 159–190.

[17] Hosking, J. (1981) Fractional differencing. *Biometrika* **68**, 165–176.

[18] Hsu, N. and Tsai, H. (2009) Semiparametric estimation for seasonal long-memory time series using generalized exponential models. *Journal of Statistical Planning and Inference*, **139**, 1992-2009.

[19] Hurvich, C. (2002) Multistep forecasting of long memory series using fractional exponential models. *International Journal of Forecasting* **18**, 167–179.

[20] Hurvich, C., Moulines, E., and Soulier, P. (2002) The FEXP estimator for potentially non-stationary linear time series. *Stochastic Processes and Their Applications* **97**, 307–340.

[21] Keeling, C.D., Bacastow, R.B., Carter, A.F., Piper, S.C., and Whore, T.P. (1989) *Am. Geophys. Union, Geophys. Monogr.* **55**, 165–236.

[22] Mathworks, Inc. (2010) Matlab, www.mathworks.com.

[23] McElroy, T. and Holan, S. (2009) Spectral Domain Diagnostics for Testing Model Proximity and Disparity in Time Series Data. *Statistical Methodology* **6**, 1–20.

[24] Palma, W. (2007) *Long Memory Time Series: Theory and Methods.* New York: Wiley.

[25] Percival, D. and Walden, A. (2000) Wavelet Methods for Time Series Analysis. Cambridge: University Press.

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

[27] Pourahmadi, M. (1984) Taylor expansion of $\exp(\sum_{k=0}^{\infty} a_k z^k)$ and some applications. *The American Mathematical Monthly* **91**, 303–307.

[28] R Development Core Team, (2010) R: A Language and Environment for Statistical Computing, http://www.R-project.org.

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

[30] Talamantes, J., Behseta, S. and Zender, C. (2007) Statistical modeling of valley fever data in Kern County, California, *International Journal of Biometeorology*, **51**, 307–313.

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

1-GEXP: $g = (0, .75)$**;** $\omega = .56$**;** $n = 500$					
$c = c_0$ $(\log \lvert\Sigma_{\text{prop}}\rvert)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = .1$ (.4299165)	.4299322 (1.568073e-5)	.4299343 (1.779021e-5)	.4299339 (1.741245e-5)	.429934 (1.748945e-5)	.429934 (1.746436e-5)
$c = .25$ (1.582358)	1.582165 (-1.930936e-4)	1.582420 (6.205987e-5)	1.582377 (1.854933e-5)	1.582400 (4.148341e-5)	1.582391 (3.288235e-5)
$c = .35$ (3.058414)	3.055264 (-.003149895)	3.058549 (1.351969e-4)	3.058041 (-3.730171e-4)	3.058518 (1.033682e-4)	3.058331 (-8.28445e-5)
$c = .45$ (5.973976)	5.939837 (-.03413864)	5.972566 (-.001410034)	5.968087 (-.005889043)	5.97576 (-.001783619)	5.972568 (-.001407956)

Table 1: Log determinant of the autocovariance sequence, for a 1-GEXP model, obtained from Equation (12) for a given J. The number in parenthesis below denotes the difference between the estimate from Equation (12) and $\log |\Sigma_{\text{prop}}|$ obtained from Proposition 3. Recall that $g_0 = 0$ implies unit innovation variance.

1-GEXP: $g = (0, .75)$**;** $\omega = .56$**;** $n = 1000$					
$c = c_0$ $(\log \lvert\Sigma_{\text{prop}}\rvert)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = .1$ (.4437794)	.4437847 (5.240159e-6)	.4437889 (9.415287e-6)	.4437881 (8.660819e-6)	.4437833 (8.817557e-6)	.4437882 (8.76521e-6)
$c = .25$ (1.669002)	1.668565 (-4.361900e-4)	1.669074 (7.217425e-5)	1.668987 (-1.433973e-5)	1.669033 (3.161854e-5)	1.669016 (1.435415e-5)
$c = .35$ (3.228235)	3.221981 (-.006254165)	3.228549 (3.133181e-4)	3.227534 (-7.010782e-4)	3.228488 (2.526479e-4)	3.228115 (-1.201864e-4)
$c = .45$ (6.2547)	6.1866 (-.06804101)	6.252189 (-.002511493)	6.243225 (-.011475681)	6.25859 (.003889263)	6.252199 (-.002501926)

Table 2: Log determinant of the autocovariance sequence, for a 1-GEXP model, obtained from Equation (12) for a given J. The number in parenthesis below denotes the difference between the estimate from Equation (12) and $\log |\Sigma_{\text{prop}}|$ obtained from Proposition 3. Recall that $g_0 = 0$ implies unit innovation variance.

2-GEXP: $g = (0, .75)$; $\omega = (\omega_1, \omega_2) = (.1, .56)$; $n = 500$					
$c = (c_1, c_2)$ ($\log \lvert\Sigma_{\text{prop}}\rvert$)	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = (.1, .2)$ (1.528072)	1.523229 (-.004843934)	1.52541 (-.002662391)	1.526801 (-.001271907)	1.527357 (-7.154246e-4)	1.527665 (-4.072432e-4)
$c = (.1, .3)$ (2.714834)	2.708252 (-.006581986)	2.711868 (-.002966315)	2.713324 (-.001509996)	2.71408 (-7.54218e-4)	2.714331 (-5.028845e-4)
$c = (.1, .45)$ (6.538299)	6.486128 (-.05217059)	6.519585 (-.01871311)	6.527094 (-.01120481)	6.538506 (2.078213e-4)	6.53538 (-2.918457e-4)
$c = (.45, .2)$ (8.134616)	7.668759 (-0.4658567)	7.90187 (-0.2327459)	8.190575 (0.05595938)	8.13965 (0.00503351)	8.115619 (-0.01899707)
$c = (.45, .3)$ (9.49205)	8.5508 (-0.9412503)	8.945401 (-0.5466491)	9.494429 (0.002378529)	9.444705 (-0.04734516)	9.411312 (-0.08073868)
$c = (.45, .4)$ (11.68251)	9.423213 (-2.259298)	9.978965 (-1.703546)	11.26194 (-0.420568)	11.27301 (-0.4094976)	11.22001 (-0.4625005)

Table 3: Log determinant of the autocovariance sequence, for a 2-GEXP model, obtained from Equation (12) for a given J. The number in parenthesis below denotes the difference between the estimate from Equation (12) and $\log \lvert\Sigma_{\text{prop}}\rvert$ obtained from Proposition 3. Recall that $g_0 = 0$ implies unit innovation variance.

2-GEXP: $g = (0, .75)$; $\omega = (\omega_1, \omega_2) = (.1, .56)$; $n = 1000$					
$c = (c_1, c_2)$ ($\log \lvert\Sigma_{\text{prop}}\rvert$)	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = (.1, .2)$ (1.597387)	1.588833 (-.008554216)	1.592713 (-.004673783)	1.595149 (-.002238315)	1.596123 (-.001263973)	1.596659 (-7.277024e-3)
$c = (.1, .3)$ (2.853464)	2.841542 (-.01192180)	2.848285 (-.005178798)	2.850855 (-.002608591)	2.852228 (-.001235417)	2.852651 (-8.126799e-4)
$c = (.1, .45)$ (6.832886)	6.730148 (-.1027383)	6.796645 (-.03624135)	6.811344 (-.02154171)	6.834062 (.001175967)	6.827721 (-.005165404)
$c = (.45, .2)$ (8.470792)	7.557704 (-0.9130878)	8.018156 (-0.4526367)	8.59122 (0.1204275)	8.486669 (0.01587655)	8.43687 (-0.0339227)
$c = (.45, .3)$ (9.897542)	8.142808 (-1.754733)	8.902406 (-0.9951356)	9.97242 (0.07487789)	9.856134 (-0.04140792)	9.776608 (-0.1209333)
$c = (.45, .4)$ (12.18504)	8.318515 (-3.866528)	9.338821 (-2.846222)	11.80701 (-0.3780357)	11.76417 (-0.4208718)	11.60253 (-0.5825138)

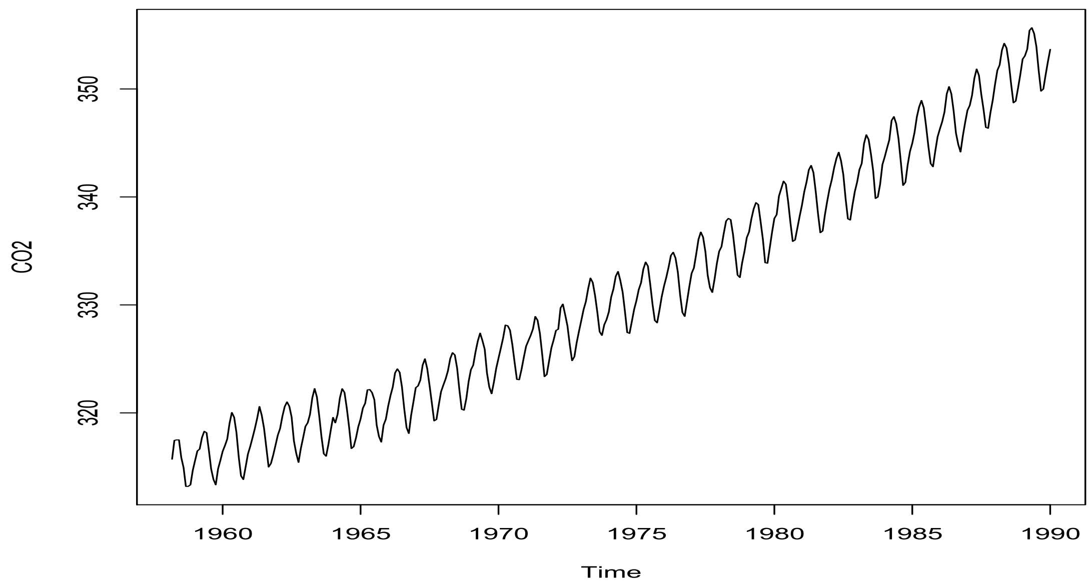
Table 4: Log determinant of the autocovariance sequence, for a 2-GEXP model, obtained from Equation (12) for a given J. The number in parenthesis below denotes the difference between the estimate from Equation (12) and $\log \lvert\Sigma_{\text{prop}}\rvert$ obtained from Proposition 3. Recall that $g_0 = 0$ implies unit innovation variance.

SFEXP: $g = (0, .75)$; $n = 500$					
$(d, D) = (d_0, D_0)$ ($\log\lvert\Sigma_{\text{prop}}\rvert$)	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$(d, D) = (.1, .2)$ (3.606141)	3.376883 (-.2292581)	3.455379 (-.1507615)	3.518562 (-.08757812)	3.547613 (-.05852793)	3.566701 (-.03943965)
$(d, D) = (.1, .3)$ (8.584982)	7.459553 (-1.125429)	7.747775 (-.837207)	8.01178 (-.5732017)	8.151985 (-.4329975)	8.256607 (-.3253756)
$(d, D) = (.1, .35)$ (12.71477)	10.422228 (-2.292494)	10.90555 (-1.809218)	11.37301 (-1.341755)	11.63730 (-1.077465)	11.84609 (-.8686839)
$(d, D) = (.2, .1)$ (1.648142)	1.622271 (-.02587158)	1.633137 (-.01500525)	1.640714 (-.007428354)	1.643690 (-.004452358)	1.645398 (-.002744174)
$(d, D) = (.3, .1)$ (2.599647)	2.573756 (-.02589099)	2.584518 (-.01512898)	2.592081 (-.007565892)	2.595056 (-.004590889)	2.596761 (-.002886382)
$(d, D) = (.35, .1)$ (3.472745)	3.446911 (-.02583382)	3.457536 (-.01520905)	3.465090 (-.007655405)	3.468067 (-.004678368)	3.469766 (-.002979346)

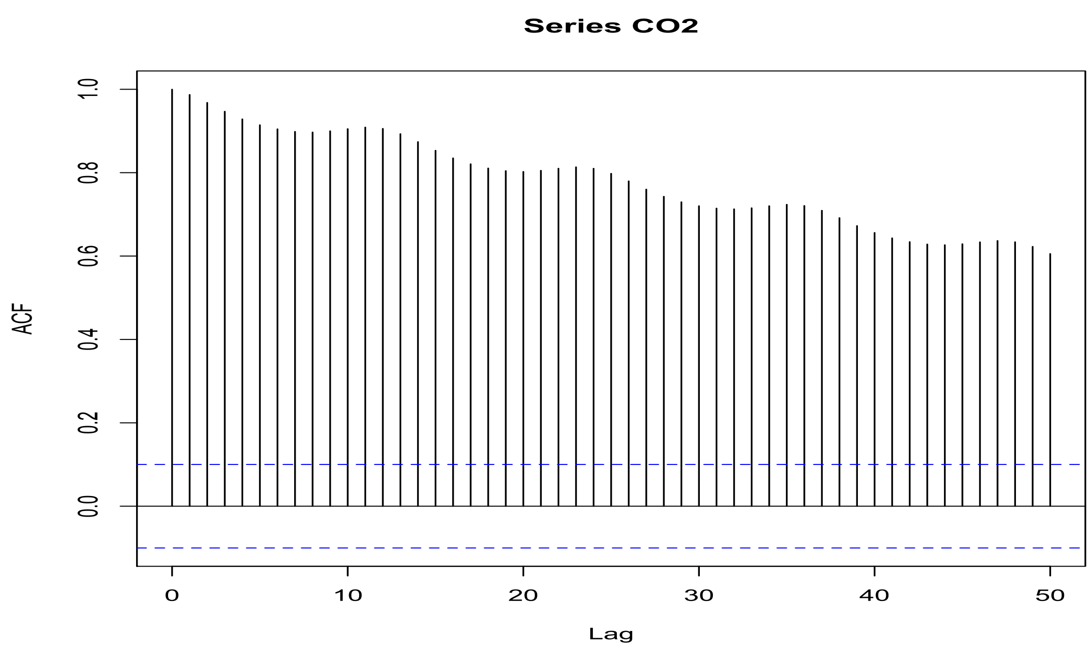
Table 5: Log determinant of the autocovariance sequence, for a SFEXP model, obtained from Equation (12) for a given J. The number in parenthesis below denotes the difference between the estimate from Equation (12) and $\log\lvert\Sigma_{\text{prop}}\rvert$ obtained from Proposition 3. Recall that $g_0 = 0$ implies unit innovation variance.

SFEXP: $g = (0, .75)$; $n = 1000$					
$(d, D) = (d_0, D_0)$ ($\log\lvert\Sigma_{\text{prop}}\rvert$)	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$(d, D) = (.1, .2)$ (3.973509)	3.622366 (-.3511428)	3.745153 (-.228356)	3.842621 (-.1308877)	3.88708 (-.0864286)	3.916191 (-.05731751)
$(d, D) = (.1, .3)$ (9.382102)	7.848874 (-1.53227)	8.257978 (-1.124124)	8.622112 (-.75999)	8.812097 (-.5700042)	8.952638 (-.4294634)
$(d, D) = (.1, .35)$ (13.78915)	10.83478 (-2.954365)	11.49619 (-2.292958)	12.11198 (-1.677172)	12.45169 (-1.337460)	12.71667 (-1.072474)
$(d, D) = (.2, .1)$ (1.786772)	1.741965 (-.04480619)	1.761085 (-.02568656)	1.774356 (-.01241565)	1.779555 (-.007216655)	1.782536 (-.004235872)
$(d, D) = (.3, .1)$ (2.786797)	2.742162 (-.04463461)	2.761073 (-.02572414)	2.774320 (-.01247751)	2.779518 (-.007279237)	2.782491 (-.004305617)
$(d, D) = (.35, .1)$ (3.689353)	3.64497 (-.04438323)	3.663606 (-.02574781)	3.676836 (-.01251715)	3.68204 (-.0073137)	3.685002 (-.004350972)

Table 6: Log determinant of the autocovariance sequence, for a SFEXP model, obtained from Equation (12) for a given J. The number in parenthesis below denotes the difference between the estimate from Equation (12) and $\log\lvert\Sigma_{\text{prop}}\rvert$ obtained from Proposition 3. Recall that $g_0 = 0$ implies unit innovation variance.

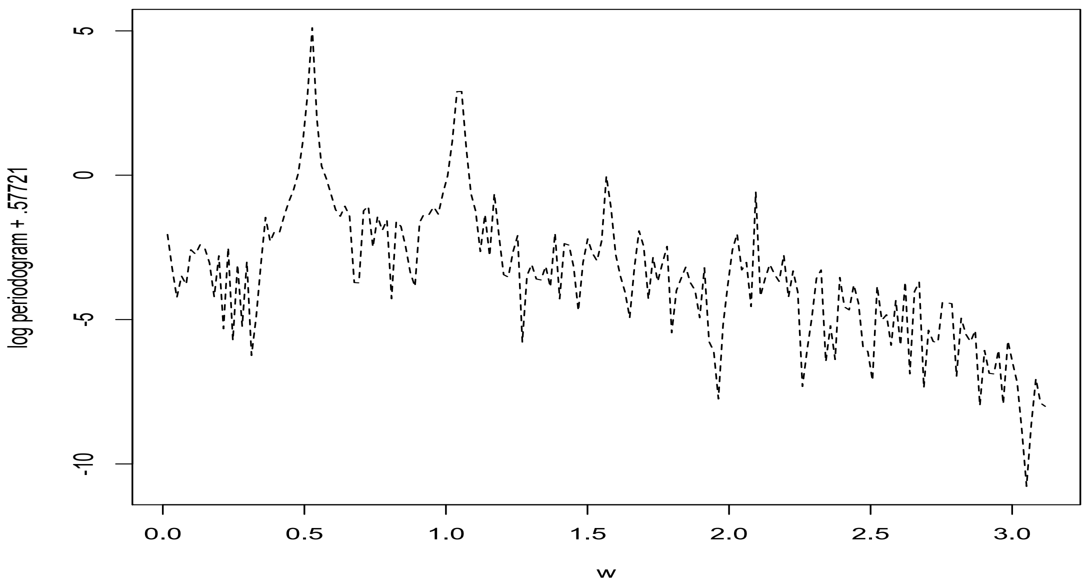


(a)

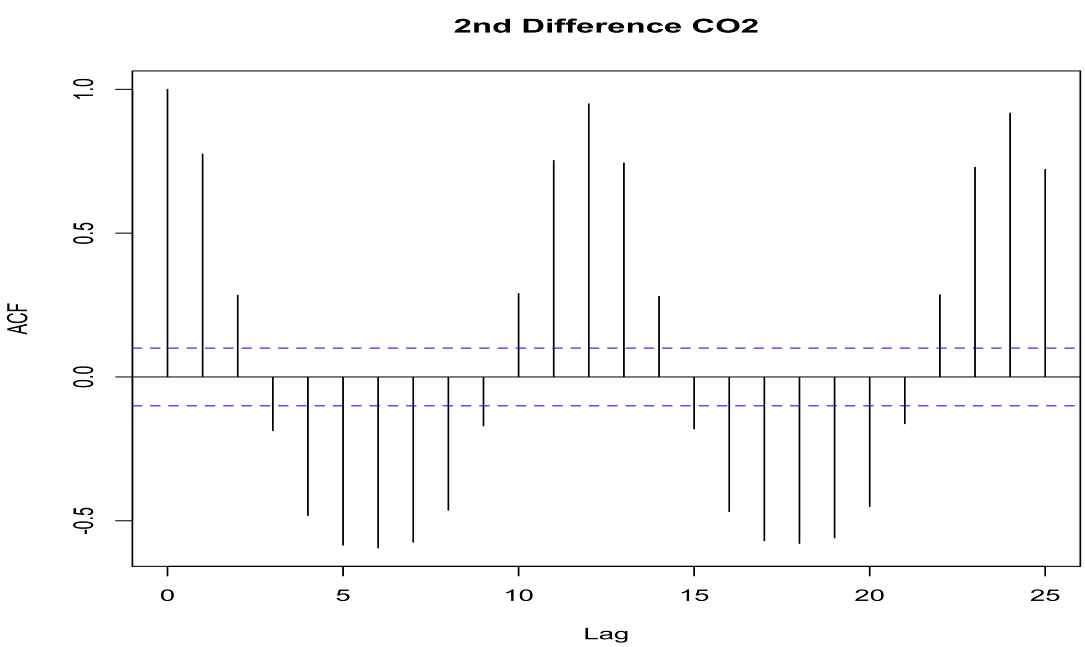


(b)

Figure 1: (a) Mauna Loa CO_2 data, 382 monthly observations beginning March 1958. (b) Autocorrelation function of the Mauna Loa CO_2 data.



(a)



(b)

Figure 2: (a) Log periodogram + .57721 of the twice differenced Mauna Loa CO_2 data. (b) Autocorrelation function of the twice differenced Mauna Loa CO_2 data.

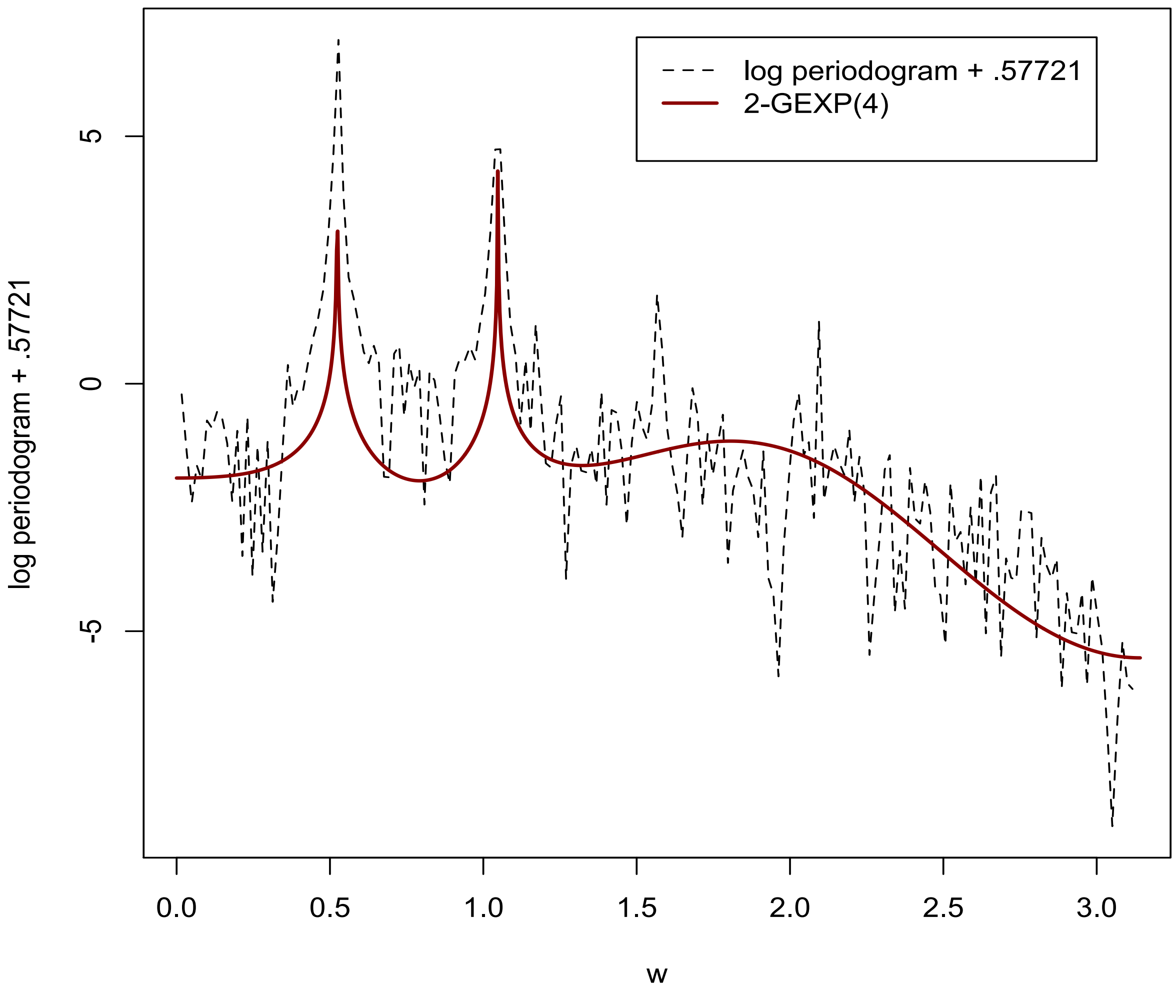


Figure 3: Log periodogram + .57721 of the twice differenced Mauna Loa CO_2 data with the estimated 2-GEXP(4) model superimposed.