Statistica Sinica **22** (2012), 1661-1687
doi:http://dx.doi.org/10.5705/ss.2010.186

ON THE COMPUTATION OF AUTOCOVARIANCES FOR GENERALIZED GEGENBAUER PROCESSES

Tucker S. McElroy and Scott H. Holan

U.S. Census Bureau and University of Missouri

Abstract: Gegenbauer processes and their generalizations represent a 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. An obstacle to using this class of models is in 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. This allows feasible calculation of the exact Gaussian likelihood, while also making simulation, forecasting, and signal extraction practicable. The techniques are illustrated using maximum likelihood estimation to model atmospheric CO_2 data.

Key words and phrases: 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, Bordignon, and Lisi (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, Behseta and Zender (2007). Because 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. 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). In principal, using the results of Lapsa (1997), 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. Two exceptions to this are Giraitis and Leipus (1995), which provides asymptotic formulas for autocovariances at high lags, and Bisognin and Lopes (2009), which provides results for autocovariances of the SARFIMA model. Specifically, Bisognin and Lopes (2009) explicitly derive the autocovariance generating function (at all lags) 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 theory, the case of multiple seasonal factors can be handled by adding another convolution for each extra term, i.e., iterated 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.

The Gaussian likelihood function requires computation of the determinant of the covariance matrix as well. 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.

Although our primary focus is on rapid accurate computation of the autocovariances for generalized Gegenbauer processes, the results provided here also facilitate solutions across a broad range of applications. For example, our results can be used to efficiently compute the exact Gaussian likelihood and, thus, used to obtain maximum likelihood estimates. Other references for maximum likelihood estimation for long range dependent processes include Chan and Palma (2005), Reisen, Rodriguez, and Palma (2006), Palma (2007), and the references therein. Combined with the asymptotic properties of the Gaussian maximum likelihood estimates for seasonal long range dependent processes presented in Palma and Chan (2005), our results increase the utility of these models in practice (e.g., see Section 5). In addition, having efficient methods to compute the exact likelihood associated with generalized Gegenbauer processes also facilitates Bayesian estimation. Previous research on Bayesian long memory has primarily been restricted to ARFIMA models (e.g., see Pai and Ravishanker (1998) and Ko and Vannucci (2006)) or FEXP models (Holan, McElroy, and Chakraborty (2009)); however, utilizing the approach described here Bayesian estimation for seasonal long memory process becomes relatively straightforward. Finally, given an efficient method for computing the autocovariances, a host of other problems can be readily addressed including simulation, signal extraction, and imputation, among others.

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. Asymptotic results for moving average coefficients and autocovariances are presented in Section 3, providing theoretical justification for our methodology. In addition, Section 3 develops an asymptotic formula for calculating the determinant of the covariance matrix. Section 4 provides relevant details regarding model computation, and also discusses several possible applications. An empirical study is presented in Section 5, illustrating the effectiveness and accuracy of our approach. Section 6 presents an application of our methodology

to modeling atmospheric CO_2 data using maximum likelihood. Finally, Section 7 provides discussion; all proofs are left to the Appendix.

2. The k-GEXP Model

Let $\{X_t\}$ be a mean zero covariance stationary time series. We suppose that $\{X_t\}$ follows a k-factor Generalized Exponential model, or k-GEXP, whose spectral density can be written as

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2a} |1 + e^{-i\lambda}|^{-2b} \prod_{\ell=1}^{k} |1 - e^{-i\omega_\ell} e^{-i\lambda}|^{-2c_\ell} |1 - e^{i\omega_\ell} e^{-i\lambda}|^{-2c_\ell} g(\lambda), \tag{2.1}$$

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

$$g(\lambda) = \exp\Big\{ \sum_{j=1}^{q} g_j \cos(\lambda j) \Big\} \sigma^2 = \exp\Big\{ g_0 + \frac{1}{2} \sum_{0<|j|\leq q} g_j e^{-i\lambda j} \Big\}, \tag{2.2}$$

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

$$\beta(z) = (1 - z)^{-a}(1 + z)^{-b} \prod_{\ell=1}^{k} (1 - e^{-i\omega_\ell} z)^{-c_\ell} (1 - e^{i\omega_\ell} z)^{-c_\ell} \kappa(z), \tag{2.3}$$

which can be written as $\sum_{\ell=0}^{\infty} \beta_\ell z^\ell$; then $g(\lambda) = |\kappa(e^{-i\lambda})|^2 \sigma^2$ and $f(\lambda) = |\beta(e^{-i\lambda})|^2 \sigma^2$. Note, by setting $b = c_1 = \cdots, c_k \equiv 0$, (2.1) reduces to the fractionally differenced exponential model (FEXP). A model related to (2.1) 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 cannot 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 (2.1) 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 (2.1). The spectrum of a generalized Gegenbauer process contains many such factors multiplied together, which of course is also of the form (2.1). The seasonal ARFIMA (of which the ARFIMA is a special case) takes the form

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} |1 - e^{-is\lambda}|^{-2D} g(\lambda), \tag{2.4}$$

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 (2.1) with $a = d + D$, $b = D$, $c_\ell = D$, and $\omega_\ell = \pm 2\pi\ell/s$, $1 \leq \ell \leq s/2 - 1$.

3. Asymptotic Properties of the Model

We begin with some asymptotic properties of the moving average coefficients and autocovariances. Although an initial treatment of this topic was given in Giraitis and Leipus (1995), our asymptotics are written slightly differently, in a form more convenient for the calculations presented in Section 4. Moreover, our formulation explicitly provides an asymptotic expression for the truncation error and, in practice, accurate computation critically depends on accounting for this error. Another distinction, albeit minor in theory, is that our short memory dynamics are given by an EXP(q) spectral representation, whereas Giraitis and Leipus (1995) work with an ARMA model. This distinction becomes increasingly important in applications where parameter estimation is required. In this context, the generality of the EXP(q) representation is advantageous because no constraints need to be imposed on the short memory parameters to ensure the resulting autocovariance matrix remains positive definite.

Consider $\beta(z)$ in (2.3), 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 $\Phi(z)$ (necessarily including $\kappa(z)$ as a factor). Further, let $\Theta(z)$ consist of m factors $\Theta_\ell(z)$ of the form $(1 - \zeta_\ell^{-1} z)^\alpha$, and denote $\beta(z)/\Theta_\ell(z)$ by $\beta_{-\ell}(z)$ (which has its pole at ζ_ℓ removed).

Theorem 1. *Let α be the minimal memory exponent in* (2.1) *and* (2.3)*, and suppose $\alpha < 0$. Then the overall rate of decay of β_j is governed by $\pi_j(\alpha) = \Gamma(j-\alpha)/\{\Gamma(j+1)\Gamma(-\alpha)\}$, and*

$$\pi_j(\alpha) = \frac{j^{-\alpha-1}}{\Gamma(-\alpha)}\{1 + O(j^{-1})\}. \tag{3.1}$$

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

$$\beta_j = \pi_j(\alpha)E(j)\{1 + o(1)\}. \tag{3.2}$$

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). We are principally interested in the case that (2.3) holds. It is easy to see that this result generalizes that of Chung (1996); in terms of our notation, that paper's equation (9) yields

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

where $\beta(z) = (1-e^{-i\omega}z)^\alpha(1-e^{i\omega}z)^\alpha$. One can rewrite $\cos\{(j-\alpha)\omega+\alpha\pi/2\}s^{1+\alpha}$ $\sin^\alpha(\omega)$ as $e^{ij\omega}(1-e^{-i2\omega})^\alpha + e^{-ij\omega}(1-e^{i2\omega})^\alpha$, which agrees with (3.2).

Our (3.2) is identical to (10) in Giraitis and Leipus (1995) after accounting for differences of notation; that paper's expansion involves a sum over various terms governed by the asymptotic behavior of $\pi_j(d_\ell)$, where d_ℓ ranges over the various memory exponents as ℓ increments. However, we have instead collapsed these terms into one dominant asymptotic expression, with $E(j)$ summarizing the contribution due to the residual dynamics; in fact, each summand in the formula for $E(j)$ corresponds to the various $D(\ell)$ expressions – multiplied by cosine sequences – of Theorem 1 of Giraitis and Leipus (1995). Writing the asymptotics of β_j in terms of the leading, or minimal, memory exponent is more convenient for deriving the truncation error expression for the autocovariances that is discussed in Section 4.

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. A notation that we use repeatedly is

$$A_h = \sum_{\ell=1}^{m} |\beta_{-\ell}(\zeta_\ell)|^2 \zeta_\ell^{-h}. \tag{3.3}$$

This is similar to $E(h)$ in Theorem 1, but is appropriate for gauging the growth of autocovariances rather than moving average coefficients.

Theorem 2. *Let α be the minimal memory exponent in* (2.1) *and* (2.3)*, and suppose $\alpha < 0$. Then the autocovariances are*

$$\gamma_h = h^{-2\alpha-1} \frac{\sigma^2 \Gamma(1+2\alpha)}{\Gamma(-\alpha)\Gamma(1+\alpha)} A_h \{1 + o(1)\} \tag{3.4}$$

as $h \to \infty$, where A_h is defined in (3.3)*.*

Remark 2. In (3.4), specific information about the long memory poles at non-zero frequencies is located in the A_h term – see (3.3) – 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 (3.3) holds for GARMA (only with the form of $\beta(z)$ being somewhat different through the specification of $\kappa(z)$).

This asymptotic form is similar to that of Theorem 1 of Giraitis and Leipus (1995), although we have expressed things in terms of a single dominant exponent. Formula (13) of Giraitis and Leipus (1995) sums various rates of growth over the various memory exponents; accounting for this difference of presentation, the asymptotics are identical, our A_h corresponding to their weighted sum of cosines. We proceed by considering a few 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

$$A_h = 2^{1-2c} \sin^{-2c}(\omega) \exp\Big\{ \sum_{\ell=1}^{q} g_\ell \cos(\omega\ell) \Big\} \cos(h\omega).$$

Also, as $h \to \infty$, it follows that $\gamma_h \sim h^{2c-1}[(\sigma^2\Gamma(1-2c))/(\Gamma(c)\Gamma(1-c))] \times 2 \times \{2\sin(\omega)\}^{-2c} \times \cos(h\omega)|\kappa(e^{i\omega})|^2$. 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). 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 $A_h = \kappa^2(-1)(-1)^h$. As for the autocovariances, 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 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. As $h \to \infty$, the autocovariances for any of these cases are given by $\gamma_h = h^{-2\alpha-1}\xi(\alpha)A_h$, with $\xi(\alpha) = \sigma^2\Gamma(1+2\alpha)/\{\Gamma(-\alpha)\Gamma(1+\alpha)\}$.

SFEXP We refer to a process following (2.4), with short memory EXP dynamics (i.e., $g(\lambda) \sim \text{EXP}(q)$), as 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 (4.2) and (4.4) 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 (3.2), and $A_h = s^{-2b}\kappa^2(1)$. Alternatively, if $b > a$ we obtain $\beta_j \sim \pi_j(-b)$ times the sum, over the $s-1$ unit roots ζ of $U(z)$, of the function $U^{-b}(z)(1-z)^{-a}\kappa(z)(1-\zeta^{-1}z)^b\zeta^{-j}$ 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

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

Finally, in this case $a \neq b$ unless either d or D equals zero. The above formula, once plugged into (3.4), yields the asymptotic autocovariances at once.

Finally, it is of interest to know the asymptotic form of the determinant of the dimension n autocovariance matrix $\Sigma_n(f)$, which is defined as having jkth entry γ_{j-k}, the autocovariance corresponding to spectral density f given by (2.1). An approximate formula is obtained from Theorem 5.47 of Böttcher and Silbermann (1999).

Proposition 1. *Under model* (2.1),

$$|\Sigma_n(f)| \sim \sigma^{2n}\, n^{a^2+b^2+2\sum_{\ell=1}^{k} c_\ell^2}\, E(f),$$

where the constant $E(f)$ *is*

$$\begin{aligned} E(f) = \exp\Big[\sum_{j=1}^{q} j g_j^2/4 + \sum_{j=1}^{q} g_j\Big\{a + b(-1)^j + 2\sum_{\ell=1}^{k} c_\ell \cos(\omega_\ell j)\Big\}\Big] \\ \cdot 2^{-2ab} \prod_{\ell=1}^{k} \{2-2\cos(\omega_\ell)\}^{-2ac_\ell} \{2+2\cos(\omega_\ell)\}^{-2bc_\ell} \{2-2\cos(2\omega_\ell)\}^{-c_\ell^2} \\ \prod_{m>\ell} \{2-2\cos(\omega_m-\omega_\ell)\}^{-2c_m c_\ell} \{2-2\cos(\omega_m+\omega_\ell)\}^{-2c_m c_\ell} \\ \cdot \frac{G^2(1-a)}{G(1-2a)} \frac{G^2(1-b)}{G(1-2b)} \prod_{\ell=1}^{k} \frac{G^4(1-c_\ell)}{G^2(1-2c_\ell)}, \end{aligned}$$

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

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

with γ *Euler's constant* ≈ 0.57721.

The proof is omitted, as the result directly follows from Böttcher and Silbermann (1999).

These results – asymptotic formulas for moving average coefficients, autocovariances, and the determinant – can be combined into a practical algorithm for computing the autocovariances of (2.1), as Section 4 demonstrates.

4. Computation of the Model

We now turn to the practical issue of computing the autocovariances of (2.1). 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 them, from which the autocovariances are determined in standard fashion. The reason why this approach is computationally efficient is that each extra multiplicative factor in (2.1) 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

$$\theta_j = \frac{1}{\pi}\int_{-\pi}^{\pi} \log f(\lambda)\cos(\lambda j)\,d\lambda \tag{4.1}$$

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 $\log\sigma^2 = (1/2\pi)\int_{-\pi}^{\pi}\log f(\lambda)\,d\lambda$ as usual. The following result tells us how to compute these coefficients (4.1) at once from the parameters in (2.1).

Proposition 2. *Under model* (2.1)*, the coefficients in* (4.1) *are*

$$\theta_j = \frac{2}{j}\left\{a + b(-1)^j + 2\sum_{\ell=1}^{k} c_\ell \cos(\omega_\ell j)\right\} + g_j. \tag{4.2}$$

For Proposition 2 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 $\kappa(z) = \prod_\ell (1-\zeta_\ell z)^{p_\ell}$ for (possibly complex) reciprocal roots ζ_ℓ of the moving average and autoregressive polynomials; here p_ℓ is one if ℓ corresponds to a moving average root, but is negative one if ℓ corresponds to an autoregressive root. Then

$$g_j = 2\sum_{\ell}\frac{p_\ell \zeta_\ell^j}{j}. \tag{4.3}$$

The derivation of (4.3) can be found in McElroy and Holan (2009) and Hsu and Tsai (2009, Appendix B). Note that if ζ_ℓ is complex, there exists a conjugate factor $\overline{\zeta}_\ell$ such that their sum is $|\zeta_\ell|^j 2\cos(j\omega_\ell)$, where ω_ℓ is the angular portion. Thus g_j is always real, and readily computed; due to the exponential decay in (4.3), 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 (2.1) as $|\beta(e^{-i\lambda})|^2\sigma^2$. Then applying the main result of Pourahmadi (1984) – also see Hurvich (2002) for treatment of the FEXP case – we obtain

$$\beta_j = \frac{1}{2j}\sum_{m=1}^{j} m\theta_m \beta_{j-m} \tag{4.4}$$

for $j \geq 1$. This is a convolution, which is expensive to compute; moreover it must be carefully monitored for decay – see Theorem 1 above. 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:

$$\gamma_h = \sigma^2 \sum_{j \geq 0} \beta_j \beta_{j+h} \tag{4.5}$$

for $h \geq 0$ (and $\gamma_{-h} = \gamma_h$). So the algorithm amounts to the application of (4.2), (4.4), and a truncated version of (4.5). Now for processes where the long memory is pronounced (say a, b, or c_l is close to 0.5), the decay of the coefficients β_j is extremely slow, and thus computing γ_h via a truncation in (4.5) 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 $a = 0.49$, 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 using the results of Theorem 1 to compute an asymptotic form for the truncation error. If at least one of the memory parameters (a, b, $c_1, \cdots, c_k$) is positive, then we can use (3.2). However if they are all negative the process is intermediate memory, the β_j decay rapidly, and the problem of truncation error evaporates.

So for some cutoff J, we express (4.5) as

$$\gamma_h = \sigma^2 \sum_{j=0}^{J-1} \beta_j \beta_{j+h} + \sigma^2 \sum_{j=J}^{\infty} \beta_j \beta_{j+h} = B_J(h) + R_J(h). \tag{4.6}$$

The first term is computed using the exact β_js via (4.4), while the second term uses the approximate β_js given in (3.2). This remainder term – denoted $R_J(h)$ – can be written as

$$\sigma^2 \sum_{j=J}^{\infty} \beta_j \beta_{j+h} \sim \frac{\sigma^2}{\Gamma^2(-\alpha)} \sum_{s,\ell=1}^{m} \beta_{-s}(\zeta_s)\beta_{-\ell}(\zeta_\ell) \sum_{j=J}^{\infty} j^{-\alpha-1}(j+h)^{-\alpha-1}\zeta_s^{-j}\zeta_\ell^{-j-h}$$

as $J \to \infty$. This remainder term can be approximated by an easily computed expression, which is given in the proposition below. The 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.

Proposition 3. *The remainder term $R_J(h)$ for fixed h and $J \to \infty$ is*

$$R_J(h) = \left\{ J^{-1-2\alpha} \, \frac{\sigma^2 \, F(1+\alpha; 1+2\alpha; 2+2\alpha; -h/J) A_h}{\Gamma^2(-\alpha)(1+2\alpha)} \right\} \{1 + o(1)\}, \tag{4.7}$$

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

Remark 3. In a sense, (3.4) is a special case of (4.7) 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 (3.4).

Given the previous formulae, the autocovariances can be efficiently computed. To summarize, the procedure is the following:

Step 1: determine all parameters of (2.1);

Step 2: compute a sufficient number of θ_j via (4.2);

Step 3: compute a sufficient number of β_j via (4.4);

Step 4: compute $B_J(h)$ for h small via (4.6);

Step 5: compute A_h for all desired h via (3.3);

Step 6: compute $R_J(h)$ for h small via (4.7);

Step 7: compute γ_h for h large via (3.4).

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., $J \geq 2,000$ and the h cutoff at 100. Additionally, in many cases, the number of desired lags is sufficiently small, such that γ_h can be calculated for all h without appealing to (3.4), 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 possibility 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, Hurvich, and Lu (2006)). We instead propose to utilize the asymptotic result of Proposition 1.

In summary, the preceding algorithm can be used to construct the autocovariance matrix needed to compute the exact Gaussian likelihood. Up to constants, the log Gaussian likelihood is given by

$$\log|\Sigma_n(f)| + X'\Sigma_n^{-1}(f)X,$$

where X represents the sample of time series data of size n, written as a column vector $X = (X_1, X_2, \cdots, X_n)'$. In many 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 1.

Our method is also useful for computing the "exact" log Whittle likelihood

$$g_0 + n^{-1} X' \Sigma_n\Big(\frac{1}{f}\Big) X.$$

(This expression follows at once from the definition of the Whittle likelihood given in Taniguchi and Kakizawa (2000).) Note that no matrix inversion is needed here, but instead we must compute autocovariances for the reciprocal spectrum $1/f$. This is actually trivial, given our preceding discussion and (4.2) – we only need flip the sign of the θ_js and run the same algorithm. Clearly the term involving the determinant is not needed for the Whittle calculation, and avoidance of the matrix inversion can potentially speed up computation for large sample sizes (e.g., $n \gg 1,000$).

We briefly mention two other applications. First, in order to simulate Gaussian processes following (2.1), accurate calculation of autocovariances is a prerequisite. With the autocovariances in hand, one only needs to construct a Cholesky factor (or matrix square root) of $\Sigma_n(f)$ and apply this to a vector of white noise. Alternatively, given the autocovariances, simulation can proceed using the Durbin-Levinson algorithm (e.g., see Brockwell and Davis (1991)) or Davies-Harte algorithm (Davies and Harte (1987)).

Finally, any type of mean squared error optimal projections (cf., Hilbert space treatment in Brockwell and Davis (1991)) involve linear filters and matrix operations applied to the data vector, where the filter coefficients are determined from the second order structure of the time series. In particular, forecasts, backcasts, imputations, and signal extractions are based upon the autocovariances of the model; see McElroy (2008) for an overview.

Given the need for autocovariances in many applications, the utility of our results is far-reaching. Nevertheless, in some cases, quick estimation of model parameters – such as the long memory parameters – may be the only target of interest. In these contexts, use of a periodogram-based Whittle likelihood approximation is less computationally intensive, and therefore requires substantially less CPU time. However, one drawback to using any approximate likelihood approach is a potential loss of accuracy (i.e., introduction of bias). For example, in the simpler SARFIMA$(0, d, 0) \times (0, d_s, 0)_s$ setting Palma (2007, pp.253-258) illustrated a downward bias for both estimated memory parameters when using a Whittle approximation. Thus, in practice, it is important to carefully evaluate the goals of the analysis, along with computation time, before considering the use of any approximate likelihood over an exact likelihood method.

5. Empirical Results

This section discusses the accuracy of the autocovariance approximations and asymptotic formula for the determinant (Proposition 1). Exact quantification of

the error necessarily depends on the ability to compute the true autocovariance sequence and 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 (4.6), the log determinant obtained from Proposition 1, and the log determinant obtained numerically from the exact autocovariances. In all cases, even for $d = 0.45$ (and J=5,000), the difference between all three methods was less than 10^{-5}. (A comprehensive breakdown of these results is available upon request.)

In the case of the FEXP(q) model, the autocovariance sequence cannot 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 (4.6), the log determinant obtained from Proposition 1, 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 $d = 0.45$ (J=5,000), the difference between the three methods was less than 10^{-4}. (A comprehensive breakdown of these results is available upon request.)

In principal, the autocovariances for 1-GEXP case can be computed to any degree of accuracy. However, this requires 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 entails 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 autocovariances from these models are unknown. Additionally, use of the splitting method for the 1-GEXP case requires an explicit formula for the autocovariances associated with the Gegenbauer process. Specifically, this requires the autocovariance function of a GARMA(0,0) model, that has a complicated form (Chung (1996); Lapsa (1997)). Therefore in order to assess the accuracy of our method for these models, we computed the log determinant numerically using the autocovariance obtained from (4.6) and the log determinant obtained from Proposition 1. Although, in principal, the autocovariances for the multiple memory parameter case could be obtained by numerically calculating the inverse Fourier transform of the

Table 1. Log determinant of the autocovariance sequence for a 1-GEXP model, obtained from (4.6) for a given J. The number in parenthesis below the log determinant denotes the difference between the estimate from (4.6) and $\log|\Sigma_{\text{prop}}|$, the log determinant obtained from Proposition 1. Recall that $g_0 = 0$ implies unit innovation variance.

1-GEXP: $g = (0,\ 0.75)$**;** $\omega = 0.56$**;** $n = 500$					
$c = c_0$ $(\log\|\Sigma_{\text{prop}}\|)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = 0.1$ (0.429917)	0.429932 (1.568073e-5)	0.429934 (1.779021e-5)	0.429934 (1.741245e-5)	0.429934 (1.748945e-5)	0.429934 (1.746436e-5)
$c = 0.25$ (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)
$c = 0.35$ (3.058414)	3.055264 (-3.149895e-3)	3.058549 (1.351969e-4)	3.058041 (-3.730171e-4)	3.058518 (1.033682e-4)	3.058331 (-8.284450e-5)
$c = 0.45$ (5.973976)	5.939837 (-3.413864e-2)	5.972566 (-1.410034e-3)	5.968087 (-5.889043e-3)	5.975760 (-1.783619e-3)	5.972568 (-1.407956e-3)

Table 2. Log determinant of the autocovariance sequence for a 1-GEXP model, obtained from (4.6) for a given J. The number in parenthesis below the log determinant denotes the difference between the estimate from (4.6) and $\log|\Sigma_{\text{prop}}|$, the log determinant obtained from Proposition 1. Recall that $g_0 = 0$ implies unit innovation variance.

1-GEXP: $g = (0,\ 0.75)$**;** $\omega = 0.56$**;** $n = 1,000$					
$c = c_0$ $(\log\|\Sigma_{\text{prop}}\|)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = 0.1$ (0.443779)	0.443785 (5.240159e-6)	0.443789 (9.415287e-6)	0.443788 (8.660819e-6)	0.443783 (8.817557e-6)	0.443788 (8.765210e-6)
$c = 0.25$ (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)
$c = 0.35$ (3.228235)	3.221981 (-6.254165e-3)	3.228549 (3.133181e-4)	3.227534 (-7.010782e-4)	3.228488 (2.526479e-4)	3.228115 (-1.201864e-4)
$c = 0.45$ (6.254700)	6.186600 (-6.804101e-2)	6.252189 (-2.511493e-3)	6.243225 (-1.147568e-2)	6.258590 (3.889263e-3)	6.252199 (-2.501926e-3)

spectrum, this numerical integration can be prohibitively slow and becomes unstable, even for "moderate" size memory parameters. As a consequence, this estimate is subject to numerical error and often not obtainable; thus, it is not included in our comparison.

As shown in Tables 1 and 2, the approximations for the 1-GEXP agree closely – for $J \geq 10,000$ and $c \leq 0.35$ the difference is on the order of 10^{-4}. In fact by increasing J, even for $c = 0.45$, 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, unless both c_1 and c_2 are greater than or equal to 0.4, the difference in the approximations can be made

Table 3. Log determinant of the autocovariance sequence for a 2-GEXP model, obtained from (4.6) for a given J. The number in parenthesis below the log determinant denotes the difference between the estimate from (4.6) and $\log|\Sigma_{\text{prop}}|$, the log determinant obtained from Proposition 1. Recall that $g_0 = 0$ implies unit innovation variance.

2-GEXP: $g = (0, 0.75)$; $\omega = (\omega_1, \omega_2) = (0.1, 0.56)$; $n = 500$					
$c = (c_1, c_2)$ $(\log\|\Sigma_{\text{prop}}\|)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = (0.1, 0.2)$ (1.528072)	1.523229 (-4.843934e-3)	1.525410 (-2.662391e-3)	1.526801 (-1.271907e-3)	1.527357 (-7.154246e-4)	1.527665 (-4.072432e-4)
$c = (0.1, 0.3)$ (2.714834)	2.708252 (-6.581986e-3)	2.711868 (-2.966315e-3)	2.713324 (-1.509996e-3)	2.714080 (-7.542180e-4)	2.714331 (-5.028845e-4)
$c = (0.1, 0.45)$ (6.538299)	6.486128 (-5.217059e-2)	6.519585 (-1.871311e-2)	6.527094 (-1.120481e-2)	6.538506 (2.078213e-4)	6.535380 (-2.918457e-4)
$c = (0.45, 0.2)$ (8.134616)	7.668759 (-0.465857)	7.90187 (-0.232746)	8.190575 (5.595938e-2)	8.13965 (0.0050335)	8.115619 (-1.899707e-2)
$c = (0.45, 0.3)$ (9.492050)	8.550800 (-0.941250)	8.945401 (-0.546649)	9.494429 (2.378529e-3)	9.444705 (-4.734516e-2)	9.411312 (-8.073868e-2)
$c = (0.45, 0.4)$ (11.682510)	9.423213 (-2.259298)	9.978965 (-1.703546)	11.261940 (-0.420568)	11.273010 (-0.409498)	11.220010 (-0.462501)

Table 4. Log determinant of the autocovariance sequence for a 2-GEXP model, obtained from (4.6) for a given J. The number in parenthesis below the log determinant denotes the difference between the estimate from (4.6) and $\log|\Sigma_{\text{prop}}|$, the log determinant obtained from Proposition 1. Recall that $g_0 = 0$ implies unit innovation variance.

2-GEXP: $g = (0, 0.75)$; $\omega = (\omega_1, \omega_2) = (0.1, 0.56)$; $n = 1,000$					
$c = (c_1, c_2)$ $(\log\|\Sigma_{\text{prop}}\|)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$c = (0.1, 0.2)$ (1.597387)	1.588833 (-8.554216e-3)	1.592713 (-4.673783e-3)	1.595149 (-2.238315e-3)	1.596123 (-1.263973e-3)	1.596659 (-7.277024e-3)
$c = (0.1, 0.3)$ (2.853464)	2.841542 (-1.192180e-2)	2.848285 (-5.178798e-3)	2.850855 (-2.608591e-3)	2.852228 (-1.235417e-3)	2.852651 (-8.126799e-4)
$c = (0.1, 0.45)$ (6.832886)	6.730148 (-0.102738)	6.796645 (-3.624135e-2)	6.811344 (-2.154171e-2)	6.834062 (1.175967e-3)	6.827721 (-5.165404e-3)
$c = (0.45, 0.2)$ (8.470792)	7.557704 (-0.913088)	8.018156 (-0.452637)	8.591220 (0.120428)	8.486669 (1.587655e-2)	8.436870 (-3.392270e-2)
$c = (0.45, 0.3)$ (9.897542)	8.142808 (-1.754733)	8.902406 (-0.995136)	9.972420 (7.487789e-2)	9.856134 (-4.140792e-2)	9.776608 (-0.120933)
$c = (0.45, 0.4)$ (12.185040)	8.318515 (-3.866528)	9.338821 (-2.846222)	11.807010 (-0.378036)	11.764170 (-0.420872)	11.602530 (-0.582514)

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 is $d + D$, whereas the seasonal long-memory parameter is D. As shown in Tables 5 and 6, when both the trend $(d + D)$ and

Table 5. Log determinant of the autocovariance sequence for a SFEXP model, obtained from (4.6) for a given J. The number in parenthesis below the log determinant denotes the difference between the estimate from (4.6) and $\log|\Sigma_{\text{prop}}|$, the log determinant obtained from Proposition 1. Recall that $g_0 = 0$ implies unit innovation variance.

SFEXP: $g = (0,\ 0.75)$**;** $n = 500$					
$(d,\ D) = (d_0,\ D_0)$ $(\log\lvert\Sigma_{\text{prop}}\rvert)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$(d,\ D) = (0.1,\ 0.2)$ (3.606141)	3.376883 (-0.229258)	3.455379 (-0.150762)	3.518562 (-8.757812e-2)	3.547613 (-5.852793e-2)	3.566701 (-3.943965e-2)
$(d,\ D) = (0.1,\ 0.3)$ (8.584982)	7.459553 (-1.125429)	7.747775 (-0.837207)	8.011780 (-0.573202)	8.151985 (-0.432998)	8.256607 (-0.325376)
$(d,\ D) = (0.1,\ 0.35)$ (12.714770)	10.422228 (-2.292494)	10.905550 (-1.809218)	11.373010 (-1.341755)	11.637300 (-1.077465)	11.846090 (-0.868684)
$(d,\ D) = (0.2,\ 0.1)$ (1.648142)	1.622271 (-2.587158e-2)	1.633137 (-1.500525e-2)	1.640714 (-7.428354e-3)	1.643690 (-4.452358e-3)	1.645398 (-2.744174e-3)
$(d,\ D) = (0.3,\ 0.1)$ (2.599647)	2.573756 (-2.589099e-2)	2.584518 (-1.512898e-2)	2.592081 (-7.565892e-3)	2.595056 (-4.590889e-3)	2.596761 (-2.886382e-3)
$(d,\ D) = (0.35,\ 0.1)$ (3.472745)	3.446911 (-2.583382e-2)	3.457536 (-1.520905e-2)	3.465090 (-7.655405e-3)	3.468067 (-4.678368e-3)	3.469766 (-2.979346e-3)

Table 6. Log determinant of the autocovariance sequence for a SFEXP model, obtained from (4.6) for a given J. The number in parenthesis below the log determinant denotes the difference between the estimate from (4.6) and $\log|\Sigma_{\text{prop}}|$, the log determinant obtained from Proposition 1. Recall that $g_0 = 0$ implies unit innovation variance.

SFEXP: $g = (0,\ 0.75)$**;** $n =$**1,000**					
$(d,\ D) = (d_0,\ D_0)$ $(\log\lvert\Sigma_{\text{prop}}\rvert)$	$J = 5,000$	$J = 10,000$	$J = 25,000$	$J = 50,000$	$J = 100,000$
$(d,\ D) = (0.1,\ 0.2)$ (3.973509)	3.622366 (-0.351143)	3.745153 (-0.228356)	3.842621 (-0.130888)	3.887080 (-8.642860e-2)	3.916191 (-5.731751e-2)
$(d,\ D) = (0.1,\ 0.3)$ (9.382102)	7.848874 (-1.53227)	8.257978 (-1.124124)	8.622112 (-0.759990)	8.812097 (-0.570004)	8.952638 (-0.429463)
$(d,\ D) = (0.1,\ 0.35)$ (13.789150)	10.834780 (-2.954365)	11.496190 (-2.292958)	12.111980 (-1.677172)	12.451690 (-1.337460)	12.716670 (-1.072474)
$(d,\ D) = (0.2,\ 0.1)$ (1.786772)	1.741965 (-4.480619e-2)	1.761085 (-2.568656e-2)	1.774356 (-1.241565e-2)	1.779555 (-7.216655e-3)	1.782536 (-4.235872e-3)
$(d,\ D) = (0.3,\ 0.1)$ (2.786797)	2.742162 (-4.463461e-2)	2.761073 (-2.572414e-2)	2.774320 (-1.247751e-2)	2.779518 (-7.279237e-3)	2.782491 (-4.305617e-3)
$(d,\ D) = (0.35,\ 0.1)$ (3.689353)	3.644970 (-4.438323e-2)	3.663606 (-2.574781e-2)	3.676836 (-1.251715e-2)	3.682040 (-7.313700e-3)	3.685002 (-4.350972e-3)

seasonal (D) long-memory parameters are larger than 0.3, 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 – a difference on the order of 10^{-2} for $J \geq 5,000$ – indicating excellent accuracy.

As expected, when models have multiple memory parameters approaching the nonstationary region of 0.5, 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 mem-

ory 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. 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, the latter being impossible to calculate directly in many situations. Finally, by coupling (4.6) with (3.4), computational efficiency can be increased for large sample sizes and/or substantial memory parameters.

6. 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)). These data were previously analyzed by Woodward, Cheng, and Gray (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. Using $J = 25,000$ to compute the autocovariances, we used the *optim* command in R (R Development Core Team (2010)), with the "L-BFGS-B" option, to numerically determine the maximum of the likelihood surface. This procedure is carefully monitored to insure convergence and that none of the parameters lie on the boundary of the parameter space.

The goal of our analysis was to demonstrate the effectiveness of embedding our autocovariance computations into a maximum likelihood analysis. As such, no efforts were 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 $\gamma = 0.57211$ 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 $\boldsymbol{\omega} = (\omega_1, \omega_2) = (0.5239, 1.048)$. These correspond to 12- and 6-month cycles and corroborate the analysis of Woodward, Cheng, and Gray (1998). Additionally, the corresponding memory parameters $\boldsymbol{c} = (c_1, c_2) = (0.4972, 0.4970)$ also coincide with the analysis of Woodward, Cheng, and Gray (1998). The associated standard errors can be obtained

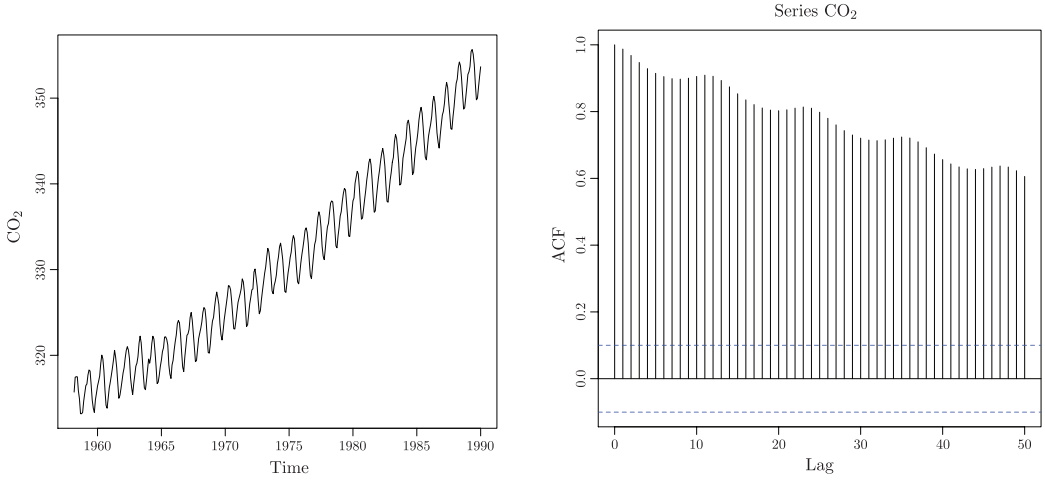

Figure 1. (a) Mauna Loa CO_2 data, 382 monthly observations beginning March 1958; (b) Autocorrelation function of the Mauna Loa CO_2 data.

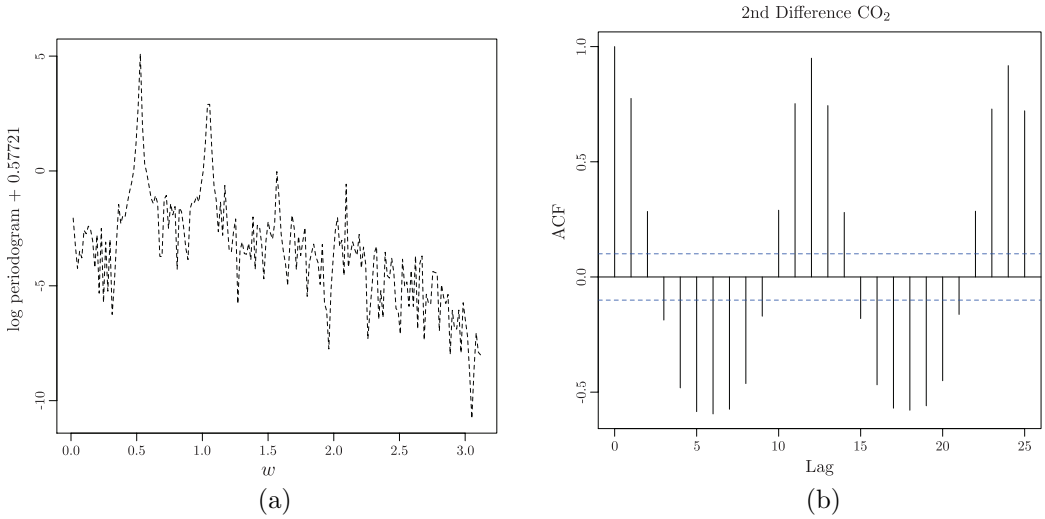

Figure 2. (a) Log periodogram + 0.57721 of the twice differenced Mauna Loa CO_2 data; (b) Autocorrelation function of the twice differenced Mauna Loa CO_2 data.

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 $(g_0, g_1, g_2, g_3, g_4) = (-2.162, -1.188, -1.175, 1.150, 0.161)$ with standard error

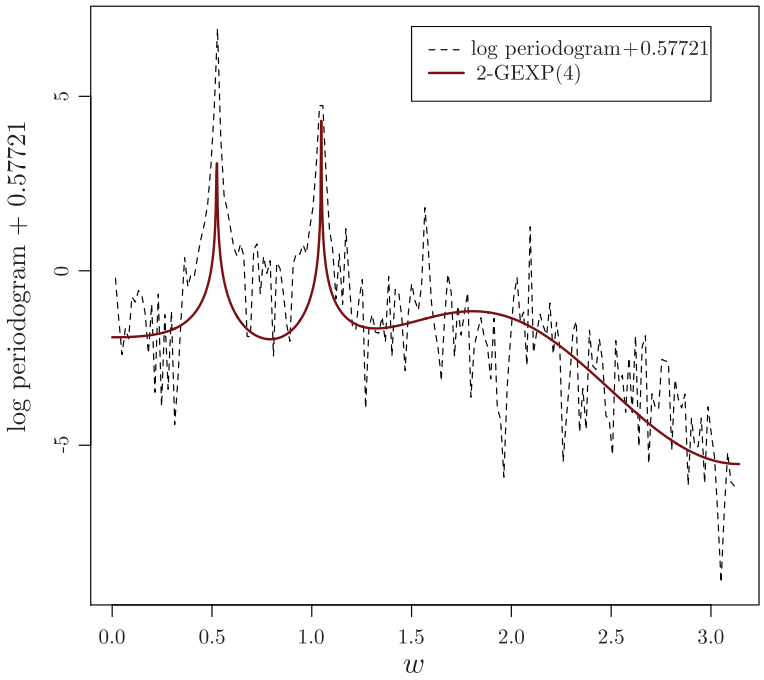

Figure 3. Log periodogram + 0.57721 of the twice differenced Mauna Loa CO_2 data with the estimated 2-GEXP(4) model superimposed.

$(1.54\times10^{-2}, 4.99\times10^{-2}, 3.17\times10^{-2}, 1.70\times10^{-2}, 1.22\times10^{-2})$ and the estimated mean (of the twice differenced data) is 0.198, with standard error 3.98×10^{-4}.

7. Conclusion

Flexible modeling of seasonal long-range dependent processes has been severely hampered by 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, generalized Gegenbauer processes have experienced limited use. Further, in their limited usage, these models have been necessarily estimated using log periodogram regression or periodogram-based 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 were achieved using the estimated autocovariance sequence directly and through the proposed asymptotic approximation to the log determinant. In general, we found that the estimated autocovariance sequences produced log determinants that agreed with high accuracy to their asymptotic approximation counterpart.

Additionally, we illustrated 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, Cheng, and Gray (1998).

Acknowledgements

Holan's research was partially supported through the "Summer at Census" program of the United States Census Bureau. The authors would like to thank two anonymous reviewers for providing many helpful suggestions to improve the article.

Appendix - Proofs of Results

Proof of Theorem 1. As a first step, consider the single factor case $\Theta(z) = (1-\zeta^{-1}z)^{\alpha}$. 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 (3.1). 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

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

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 $0 \leq \ell < \tau(j) - 1$ (part I); sum over $\tau(j) \leq \ell < j/2$ (part II); sum over $j/2 \leq \ell < j - \tau(j)$ (part III); sum over $j - \tau(j) \leq \ell \leq j$ (part IV). For part I, we have $\phi_{j-\ell}$ is asymptotic to ϕ_j times a function of exponential decay, so that $\phi_{j-\ell}/\phi_j \sim r^{-\ell}$ (for some $r < 1$) times other bounded functions. Then the sum in part I is asymptotic to $\phi_j \sum_{\ell=0}^{\tau(j)}\pi_\ell(\alpha)\zeta^{-\ell}r^{-\ell}$ (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 $\pi_\ell(\alpha)$'s asymptotic form (3.1) can be substituted since $\ell \geq \tau(j) \to \infty$. The resulting sum converges (because $\phi_{j-\ell}$ 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 $\pi_{j-\ell}(\alpha)\zeta^{\ell-j}\phi_\ell$. Since $\pi_{j-\ell}(\alpha)/\pi_j(\alpha) \to 1$ for $\ell \leq \tau(j)$, 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 $\pi_j(\alpha)\zeta^{-j}\sum_{\ell=0}^{\tau(j)}\zeta^\ell\phi_\ell$; this sum tends to $\Phi(\zeta)$, which yields (3.2).

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 (3.2): $\delta_j \sim \pi_j(\alpha)\zeta^{-j}\Phi(\zeta)$. The second term $(1-\xi^{-1}z)^\gamma$ has coefficients $\rho_j = \pi_j(\gamma)\xi^{-j}$. Now similar to (A.1), $\beta_j = \sum_{\ell=0}^{j}\delta_\ell\rho_{j-\ell}$, 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 $\rho_\ell\zeta^{-\ell}\xi^{-\ell}$, as ξ and ζ are unit roots – to conclude that the first three parts are $o(\pi_j(\alpha))$. Part IV is asymptotic to $\delta_j\sum_{\ell=0}^{\tau(j)}\xi^\ell\rho_\ell \sim \pi_j(\alpha)\zeta^{-j}\Phi(\zeta)(1-\xi^{-1}\zeta)^\gamma$, as desired.

Finally, we induct on the number of factors in $\Theta(z)$. We have established the base case of one factor, so suppose that (3.2) holds for m factors: $\beta(z) = \Theta(z)(1-\zeta_{m+1}^{-1}z)^\alpha$, and θ_j satisfies (3.2). Since the coefficients of this latter factor are exactly $\pi_j(\alpha)\zeta_{m+1}^{-j}$, (A.1) yields $\beta_j = \sum_{\ell=0}^{j}\theta_\ell\pi_{j-\ell}(\alpha)\zeta_{m+1}^{\ell-j}$. 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 have $\zeta_{m+1}^{-j}\pi_j(\alpha)\sum_{\ell=0}^{\tau(j)}\theta_\ell\zeta_{m+1}^\ell \sim \pi_j(\alpha)\zeta_{m+1}^{-j}\Theta(\zeta_{m+1})$. For part IV we first observe that

$$\frac{\theta_{j-\ell}}{\theta_j} \sim \frac{\sum_{s=1}^m \zeta_s^{-j+\ell}\Theta_{-s}(\zeta_s)}{\sum_{s=1}^m \zeta_s^{-j}\Theta_{-s}(\zeta_s)},$$

and this ratio is asymptotic to $\theta_j\sum_{\ell=0}^{\tau(j)}\sum_{s=1}^m\zeta_s^{-j+\ell}\Theta_{-s}(\zeta_s)\pi_\ell(\alpha)\zeta_{m+1}^{-\ell}/\sum_{s=1}^m\zeta_s^{-j}$ $\Theta_{-s}(\zeta_s)$, which is $\pi_j(\alpha)\sum_{s=1}^m\zeta_s^\ell\Theta_{-s}(\zeta_s)(1-\zeta_s/\zeta_{m+1})^\alpha$. Putting parts I and IV together yields

$$\frac{\beta_j}{\pi_j(\alpha)} \sim \zeta_{m+1}^{-j}\Theta(\zeta_{m+1}) + \sum_{s=1}^m\zeta_s^\ell\Theta_{-s}(\zeta_s)\Big(1-\frac{\zeta_s}{\zeta_{m+1}}\Big)^\alpha = \sum_{s=1}^{m+1}\zeta_s^\ell\beta_{-s}(\zeta_s).$$

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.

Proof of Theorem 2 and Proposition 3. 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: $j \le \tau(h)$ and $j > \tau(h)$. The first portion is

$$\begin{aligned}\sum_{j=0}^{\tau(h)} \beta_j \beta_{j+h} &= \sum_{j=0}^{\tau(h)} \beta_j (j+h)^{-(1+\alpha)} \Gamma^{-1}(-\alpha) E(j+h)\{1+o(1)\} \\ &= h^{-(1+\alpha)} \Gamma^{-1}(-\alpha) \sum_{j=0}^{\tau(h)} \beta_j \Big(1+\frac{j}{h}\Big)^{-(1+\alpha)} E(j+h)\{1+o(1)\}.\end{aligned}$$

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 $\sum_{j>\tau(h)} j^{-(1+\alpha)}(j+h)^{-(1+\alpha)} \zeta_s^{-j} \zeta_\ell^{-j}$ for unit roots ζ_s, ζ_ℓ. If $\zeta_s \zeta_\ell \ne 1$, the sum is oscillatory, fostering convergence, such that we obtain a bound of $O(h^{-(1+\alpha)})$; else if $\zeta_s \zeta_\ell = 1$ the sum decays at the slower rate of $h^{-(1+2\alpha)}$, as shown below. Letting $\ell_0 = \tau(h)$ and $\ell_p = hp$ for $p \ge 1$, we can rewrite as

$$\begin{aligned}&h^{-2(1+\alpha)} \Big\{ \sum_{p\ge 1} \sum_{j=\ell_{p-1}+1}^{\ell_p} (j/h)^{-(1+\alpha)} (1+j/h)^{-(1+\alpha)} \Big\} \\ &\quad \sim h^{-(1+2\alpha)} \sum_{p\ge 1} \int_{p-1}^{p} x^{-(1+\alpha)} (1+x)^{-(1+\alpha)}\, dx,\end{aligned}$$

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 determine how many unit root pairs satisfy $\zeta_s \zeta_\ell = 1$. 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 have $\zeta_s \zeta_\ell = 1$ for all $1 \le s, \ell \le m$, which is

$$\sum_{s,\ell=1}^{m} 1_{\{\zeta_s \zeta_\ell = 1\}} \beta_{-s}(\zeta_s) \beta_{-\ell}(\zeta_\ell) \zeta_\ell^{-h} = \sum_{\ell=1}^{m} |\beta_{-\ell}(\zeta_\ell)|^2 \zeta_\ell^{-h} = A_h.$$

Putting this together with $h^{-(1+2\alpha)}\Gamma(-\alpha)\Gamma(1+2\alpha)/\Gamma(1+\alpha)$ yields (3.4).

Here we also prove Proposition 3 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 $\zeta_\ell \zeta_s = 1$, and the summation for such ℓ, s is

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

This is the Riemann integration approximation for bounded integrands, so the error is $O(J^{-2-2\alpha})$. However, the approximation error from discounting the $\zeta_\ell \zeta_s \neq 1$ terms is lower order, namely $o(J^{-1-2\alpha})$. Using 3.194.2 of Gradshteyn and Ryzhik (1994), the integral is

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

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

Proof of Proposition 2. Taking the logarithm of (2.1) yields

$$\begin{aligned}\log f(\lambda) = &-a \log(2 - 2\cos\lambda) - b \log(2 + 2\cos\lambda) \\ &- \sum_{\ell=1}^{k} c_\ell \log\left[\{2 - 2\cos(\omega_\ell - \lambda)\}\{2 - 2\cos(\omega_\ell + \lambda)\}\right] + \log g(\lambda).\end{aligned}$$

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 is broken into two integrals over $[-\pi, 0)$ and $(0, \pi]$ and re-assembled, which shows 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

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

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

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

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

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

The integrand has a simple pole at $e^{-i\omega}$, unless $e^{i\omega 2j}=1$ (in which case there is a cancellation). 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 (2.1) always occur in pairs, which accounts for the doubling of these terms in (4.2). This concludes the derivation.

References

Beran, J. (2010). Long-range dependence. *WIREs Comput. Stat.* **2**, 26-35.

Bertelli, S. and Caporin, M. (2002). A note on calculating autocovariances of long-memory processes. *J. Time Ser. Anal.* **23**, 503-508.

Bisaglia, L., Bordignon, S. and Lisi, F. (2003). k-Factor GARMA models for intraday volatility forecasting. *Appl. Econom. Lett.* **10**, 251-254.

Bisognin, C. and Lopes, S. R. C. (2009). Properties of seasonal long memory processes. *Math. Comput. Modelling* **49**, 1837-1851.

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

Böttcher, A. and Silbermann, B. (1999). *Introduction to Large Truncated Toeplitz Matrices.* Springer, New York.

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

Chan, N. H. and Palma, W. (2005). Estimation of Long-Memory Time Series Models: a Survey of Different Likelihood-Based Methods. In *Econometric Analysis of Financial and Economic Time Series*, (Edied by Thomas B. Fomby, Dek Terrell), Emerald Group Publishing Limited, 89-121.

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. *J. Amer. Statist. Assoc.* **101**, 812-822.

Chung, C. (1996). Estimating a generalized long memory process. *J. Econometrics* **73**, 237-259.

Davies, R. B. and Harte, D. S. (1987). Tests for the Hurst effect. *Biometrika* **74**, 95-101.

Doornik, J. and Ooms, M. (2003). Computational aspects of maximum likelihood estimation of autoregressive fractionally integrated moving average models. *Comput. Statist. Data Anal.* **42**, 333-348.

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

Gil-Alana, L. A. (2008). Cyclical long-range dependence and the warming effect in a long temperature time series. *Internat. J. Climatology* **28**, 1435-1443.

Giraitis, L. and Leipus, R. (1995). A generalized fractionally differencing approach in long-memory modeling. *Lithuanian Math. J.* **35**, 53-65.

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

Gray, H., Zhang, N. and Woodward, W. (1989). On generalized fractional processes. *J. Time Ser. Anal.* **10**, 233-257.

Golub, G. H. and Van Loan, C. F. (1996). *Matrix Computations.* 3rd edition. Johns Hopkins University Press, Baltimore.

Henrici, P. (1974). *Applied and Computational Complex Analysis*, Volume I. Wiley, New York.

Holan, S., McElroy, T. and Chakraborty, S. (2009). A Bayesian approach to estimating the long memory parameter. *Bayesian Analysis* **4**, 159-190.

Hosking, J. (1981). Fractional differencing. *Biometrika* **68**, 165-176.

Hsu, N. and Tsai, H. (2009). Semiparametric estimation for seasonal long-memory time series using generalized exponential models. *J. Statist. Plann. Inference* **139**, 1992-2009.

Hurvich, C. (2002). Multistep forecasting of long memory series using fractional exponential models. *Internat. J. Forecasting* **18**, 167-179.

Keeling, C. D., Bacastow, R. B., Carter, A. F., Piper, S. C. and Whorf, T. P. (1989). Aspects of climate variability in the Pacific and the Western Americas. *Am. Geophys. Union, Geophys. Monogr.* **55**, 165-236.

Ko, K. and Vannucci, M. (2006). Bayesian wavelet-based methods for the detection of multiple changes of the long memory parameter. *IEEE Trans. Signal Process.* **54**, 4461-4470.

Lapsa, P. (1997). Determination of Gegenbauer-type random process models. *Signal Processing* **63**, 73-90.

Mathworks, Inc. (2010). Matlab. `www.mathworks.com`.

McElroy, T. (2008). Matrix formulas for nonstationary ARIMA signal extraction. *Econometric Theory* **24**, 1-22.

McElroy, T. and Holan, S. (2009). Spectral domain diagnostics for testing model proximity and disparity in time series data. *Statistical Methodology* **6**, 1-20.

Pai, J. and Ravishanker, N. (1998). Bayesian analysis of autoregressive fractionally integrated moving-average processes. *J. Time Ser. Anal.* **19**, 99-112.

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

Palma, W. and Chan, N. H. (2005). Efficient estimation of seasonal long range dependent processes. *J. Time Ser. Anal.* **26**, 863-892.

Percival, D. and Walden, A. (2000). *Wavelet Methods for Time Series Analysis.* University Press, Cambridge.

Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. *J. Amer. Statist. Assoc.* **85**, 338-344.

Pourahmadi, M. (1984). Taylor expansion of $\exp(\sum_{k=0}^{\infty} a_k z^k)$ and some applications. *Amer. Math. Monthly* **91**, 303-307.

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

Reisen, V., Rodriguez, A. and Palma, W. (2006). Estimation of seasonal fractionally integrated processes. *Comput. Statist. Data Anal.* **50**, 568-582.

Soares, L. J. and Souza, L. R. (2006). Forecasting electricity demand using generalized long memory. *Internat. J. Forecasting* **22**, 17-28.

Talamantes, J., Behseta, S. and Zender, C. (2007). Statistical modeling of valley fever data in Kern County, California. *Internat. J. Biometeorology* **51**, 307-313.

Taniguchi, M. and Kakizawa, Y. (2000). *Asymptotic Theory of Statistical Inference for Time Series.* Springer-Verlag, New York.

Woodward, W., Cheng, Q. and Gray, H. (1998). A k-factor GARMA long memory model. *J. Time Ser. Anal.* **19**, 485-504.

Center for Statistical Research and Methodology, U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-9100, USA.

E-mail: tucker.s.mcelroy@census.gov

Department of Statistics, University of Missouri, 146 Middlebush Hall, Columbia, MO, 65211-6100, USA.

E-mail: holans@missouri.edu

(Received August 2010; accepted October 2011)