



HAL
open science

Asymptotic normality of the mixture density estimator in a disaggregation scheme

Dmitrij Celov, Remigijus Leipus, Anne Philippe

► **To cite this version:**

Dmitrij Celov, Remigijus Leipus, Anne Philippe. Asymptotic normality of the mixture density estimator in a disaggregation scheme. 2008. hal-00242913

HAL Id: hal-00242913

<https://hal.science/hal-00242913>

Preprint submitted on 6 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Asymptotic normality of the mixture density estimator in a disaggregation scheme*

Dmitrij Celov¹, Remigijus Leipus^{1,2} and Anne Philippe³

¹Vilnius University, Lithuania

²Institute of Mathematics and Informatics, Lithuania

³Laboratoire de Mathématiques Jean Leray, Université de Nantes, France

February 6, 2008

Abstract

The paper concerns the asymptotic distribution of the mixture density estimator, proposed by Leipus et al. (2006), in the aggregation/disaggregation problem of random parameter AR(1) process. We prove that, under mild conditions on the (semiparametric) form of the mixture density, the estimator is asymptotically normal. The proof is based on the limit theory for the quadratic form in linear random variables developed by Bhansali et al. (2007). The moving average representation of the aggregated process is investigated. A small simulation study illustrates the result.

Keywords: random coefficient AR(1), long memory, aggregation, disaggregation, mixture density.

*The research was supported by bilateral France-Lithuanian scientific project Gilibert and Lithuanian State Science and Studies foundation (V-07058).

1 Introduction

Aggregated time series data appears in different fields of studies including applied problems in hydrology, sociology, statistics, economics. Considering aggregation as a *time series* object, a number of important questions arise. These comprise the properties of macro level data obtained by small and large-scale aggregation in time, space or both, assumptions of when and how the inverse (disaggregation) problem can be solved, finally, how to apply theoretical results in practice.

Aggregated time series, in fact, can be viewed as a transformation of the underlying time series by some (either linear or non-linear) specific function defined at (in)finite set of individual processes. In this paper we consider a linear aggregation scheme, which is natural in applications. In practice it is found convenient to approximate individual data by simple time series models, such as AR(1), GARCH(1,1) for instance (see Lewbel (1994), Chong (2006), Zaffaroni (2004, 2006)), whereas more complex individual data models do not provide an advantage in accuracy and efficiency of estimates, and usually are very difficult to study from the theoretical point of view.

Aggregation by appropriately averaging the micro level time series models can give intriguing results. It was shown in Granger (1980) that the large-scale aggregation of infinitely many short memory AR(1) models with random coefficients can lead to a long memory fractionally integrated process. It means that the properties of an aggregate time series may in general differ from those of individual data.

It is clear however that the weakest point of the aggregation is a considerable loss of information about individual characteristics of the underlying data. Roughly speaking, an aggregated time series can not be as informative about the attributes of individual data as the micro level processes are. On the other hand, using some special aggregation schemes, which involve, for instance, independent identically distributed “elementary” processes with known structure (such as AR(1)), enables to solve an inverse problem: to recover the properties of individual series with the aggregated data at hand.

This problem is called a *disaggregation problem*.

Different aspects of this problem were investigated in Dacunha-Castelle, Oppenheim (2001), Leipus et al. (2006), Celov et al. (2007). The last two papers deal with the asymptotic statistical theory in the disaggregation problem: they present the construction of the mixture density estimate of the individual AR(1) models, the consistency of an estimate, and some theoretical tools needed here. Resuming the previous research, the major objective of the present paper is to obtain the *asymptotic normality* property of the mixture density estimate, that enlarges the range of applications, solving the accuracy of simulation studies, statistical inference, forecasting and other problems.

Section 2 describes the disaggregation scheme, including the construction of mixture density estimate proposed by Leipus et al. (2006), and formulates the main result of the paper. Important issues about the moving average representation of the aggregated process are discussed in Section 3. The proof of the main theorem and auxiliary results are given respectively in Section 4 and Section 7. Some simulation results are presented in Section 5.

2 Preliminaries and the main result

Consider a sequence of independent processes $Y^{(j)} = \{Y_t^{(j)}, t \in \mathbf{Z}\}$, $j \geq 1$ defined by the random coefficient AR(1) dynamics

$$Y_t^{(j)} = a^{(j)}Y_{t-1}^{(j)} + \varepsilon_t^{(j)}, \quad (2.1)$$

where $\varepsilon_t^{(j)}$, $t \in \mathbf{Z}$, $j = 1, 2, \dots$ are independent identically distributed (i.i.d.) random variables with $E\varepsilon_t^{(j)} = 0$ and $0 < \sigma_\varepsilon^2 = E(\varepsilon_t^{(j)})^2 < \infty$; $a, a^{(j)}$, $j = 1, 2, \dots$ are i.i.d. random variables with $|a| \leq 1$ and satisfying

$$E\left[\frac{1}{1-a^2}\right] < \infty. \quad (2.2)$$

It is assumed that the sequences $\{\varepsilon_t^{(j)}, t \in \mathbf{Z}\}$, $j = 1, 2, \dots$ and $\{a, a^{(j)}, j = 1, 2, \dots\}$ are independent.

Under these conditions, (2.1) admits a stationary solution $Y^{(j)}$ and, according to Oppenheim and Viano (2004), the finite dimensional distributions

of the process

$$X_t^{(N)} = \frac{1}{\sqrt{N}} \sum_{j=1}^N Y_t^{(j)}, \quad t \in \mathbf{Z},$$

weakly converge as $N \rightarrow \infty$ to those of a zero mean stationary Gaussian process $X = \{X_t, t \in \mathbf{Z}\}$, called the *aggregated* process. Suppose that random coefficient a admits a density $\varphi(x)$, absolutely continuous with respect to the Lebesgue measure, which by (2.2) satisfies

$$\int_{-1}^1 \frac{\varphi(x)}{1-x^2} dx < \infty. \quad (2.3)$$

Any density function satisfying (2.3) will be called a *mixture density*.

Note that the covariance function and the spectral density of aggregated process X coincides with those of $Y^{(j)}$ and are given, respectively, by

$$\sigma(h) := \text{Cov}(X_h, X_0) = \sigma_\varepsilon^2 \int_{-1}^1 \frac{x^{|h|}}{1-x^2} \varphi(x) dx \quad (2.4)$$

and

$$f(\lambda) = \frac{\sigma_\varepsilon^2}{2\pi} \int_{-1}^1 \frac{\varphi(x)}{|1-xe^{i\lambda}|^2} dx. \quad (2.5)$$

The *disaggregation problem* deals with finding the individual processes (if they exist) of form (2.1), which produce the aggregated process X with *given* spectral density $f(\lambda)$ (or covariance $\sigma(h)$). This is equivalent to finding $\varphi(x)$ such that (2.5) (or (2.4)) and (2.3) hold. In this case, we say that *the mixture density $\varphi(x)$ is associated with the spectral density $f(\lambda)$* .

In order to estimate the mixture density $\varphi(x)$ using aggregated observations X_1, \dots, X_n , Leipus et al. (2006) proposed the estimate based on a decomposition of function $\zeta(x) = \varphi(x)(1-x^2)^{-\alpha}$ in the orthonormal $L^2(w^{(\alpha)})$ -basis of Gegenbauer polynomials $\{G_k^{(\alpha)}(x), k = 0, 1, \dots\}$, where $w^{(\alpha)}(x) = (1-x^2)^\alpha$, $\alpha > -1$. This decomposition is valid (i.e. ζ belongs to $L^2(w^{(\alpha)})$) if

$$\int_{-1}^1 \frac{\varphi^2(x)}{(1-x^2)^\alpha} dx < \infty, \quad \alpha > -1. \quad (2.6)$$

Let $G_n^{(\alpha)}(x) = \sum_{j=0}^n g_{n,j}^{(\alpha)} x^j$. The resulting estimate has the form

$$\hat{\varphi}_n(x) = \hat{\sigma}_{n,\varepsilon}^{-2} (1-x^2)^\alpha \sum_{k=0}^{K_n} \hat{\zeta}_{n,k} G_k^{(\alpha)}(x), \quad (2.7)$$

where the $\hat{\zeta}_{n,k}$ are estimates of the coefficients ζ_k in the α -Gegenbauer expansion of the function $\zeta(x) = \sum_{k=0}^{\infty} \zeta_k G_k^{(\alpha)}(x)$ and are given by

$$\hat{\zeta}_{n,k} = \sum_{j=0}^k g_{k,j}^{(\alpha)} (\hat{\sigma}_n(j) - \hat{\sigma}_n(j+2)), \quad (2.8)$$

$\hat{\sigma}_{n,\varepsilon}^2 = \hat{\sigma}_n(0) - \hat{\sigma}_n(2)$ is the consistent estimator of variance σ_ε^2 and $\hat{\sigma}_n(j) = n^{-1} \sum_{i=1}^{n-j} X_i X_{i+j}$ is the sample covariance of the aggregated process. Truncation level K_n satisfies

$$K_n = \lceil \gamma \log n \rceil, \quad 0 < \gamma < (2 \log(1 + \sqrt{2}))^{-1}. \quad (2.9)$$

Leipus et al. (2006) assumed the following semiparametric form of the mixture density:

$$\varphi(x) = (1-x)^{1-2d_1} (1+x)^{1-2d_2} \psi(x), \quad 0 < d_1, d_2 < 1/2, \quad (2.10)$$

where $\psi(x)$ is continuous on $[-1, 1]$ and does not vanish at ± 1 . Then, under conditions above and corresponding relations between α and d_1, d_2 , they showed the consistency of the estimator $\hat{\varphi}_n(x)$ assuming that the variance of the noise, σ_ε^2 , is *known* and equals 1. In more realistic situation of unknown σ_ε^2 , it must be consistently estimated. In order to understand intuitively the construction of estimator $\hat{\sigma}_{n,\varepsilon}^2$, it suffices to note that, by (2.4), $\sigma_\varepsilon^2 = \sigma(0) - \sigma(2)$. Also note that the estimator $\hat{\varphi}_n(x)$ in (2.7) possesses property $\int_{-1}^1 \hat{\varphi}_n(x) dx = 1$, which can be easily verified noting that $\int_{-1}^1 (1-x^2)^\alpha G_k^{(\alpha)}(x) dx = (g_{0,0}^{(\alpha)})^{-1}$ if $k=0$, and = 0 otherwise, implying

$$\int_{-1}^1 (1-x^2)^\alpha \sum_{k=0}^{K_n} \hat{\zeta}_{n,k} G_k^{(\alpha)}(x) dx = \hat{\zeta}_{n,0} / g_{0,0}^{(\alpha)} = \hat{\sigma}_n(0) - \hat{\sigma}_n(2)$$

by (2.8).

In this paper, we further study the properties of the proposed mixture density estimator. In order to formulate the theorem about the asymptotic normality of estimator $\hat{\varphi}_n(x)$, we will assume that aggregated process X_t , $t \in \mathbf{Z}$ admits the following linear representation.

ASSUMPTION A Assume that $X_t, t \in \mathbf{Z}$ is a linear sequence

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad (2.11)$$

where the Z_t are i.i.d. random variables with zero mean, finite fourth moment and the coefficients ψ_j satisfy

$$\psi_j \sim cj^{d-1}, \quad |\psi_j - \psi_{j+1}| = O(j^{d-2}), \quad 0 < d < 1/2 \quad (2.12)$$

with some constant $c \neq 0$.

We also introduce the following condition on the mixture density $\varphi(x)$.

ASSUMPTION B Assume that mixture density φ has a form

$$\varphi(x) = (1-x)^{1-2d}\psi(x), \quad 0 < d < 1/2, \quad (2.13)$$

where $\psi(x)$ is a nonnegative function with $\text{supp}(\psi) \subset [-1, 1]$, continuous at $x = 1$, $\psi(1) \neq 0$.

Note that, omitting in (2.10) the factor responsible for the seasonal part, we thus obtain the corresponding 'long memory' spectral density with singularity at zero (but not necessary at $\pm\pi$) and the corresponding behavior of the coefficients ψ_j in linear representation (3.2).

Theorem 2.1 *Let $X_t, t \in \mathbf{Z}$ be the aggregated process satisfying Assumption A and corresponding to the mixture density given by Assumption B. Assume that (2.6) holds, and d and α satisfy the following condition*

$$-1/2 < \alpha < \frac{5}{2} - 4d. \quad (2.14)$$

Let K_n be given in (2.9) with γ satisfying

$$0 < \gamma < (2 \log(1 + \sqrt{2}))^{-1} \left(1 - \max \left\{ \alpha + 4d - \frac{3}{2}, 0 \right\} \right). \quad (2.15)$$

Then for every fixed $x \in (-1, 1)$, such that $\varphi(x) \neq 0$, it holds

$$\frac{\hat{\varphi}_n(x) - \mathbb{E}\hat{\varphi}_n(x)}{\sqrt{\text{Var}(\hat{\varphi}_n(x))}} \xrightarrow{d} \text{N}(0, 1). \quad (2.16)$$

Proof of the theorem is given in Section 4.

REMARK 2.1 Suppose that $\varphi(x)$ satisfies Assumption B. Then assumption (2.6) is equivalent to $\int_{-1}^1 \psi^2(x)(1+x)^{-\alpha} dx < \infty$ and $\alpha < 3 - 4d$. The last inequality is implied by (2.14).

EXAMPLE 2.1 Assume two mixture densities

$$\varphi(x; d) = C_1(d)x^{d-1}(1-x)^{1-2d}(1+x)\mathbf{1}_{(0,1)}(x), \quad 0 < d < 1/2, \quad (2.17)$$

where $C_1(d) = \frac{\Gamma(3-d)}{2\Gamma(d)\Gamma(2-2d)}$, and

$$\varphi_g(x; \kappa) = C_2(\kappa)|x|^\kappa \mathbf{1}_{[-a_*, 0]}(x), \quad \kappa > 0, \quad (2.18)$$

where $0 < a_* < 1$, $C_2(\kappa) = (\kappa + 1)(a_*)^{-\kappa-1}$.

According to Dacunha-Castelle and Oppenheim (2001), the spectral density corresponding to $\varphi(x; d)$ is FARIMA(0,d,0) spectral density

$$f(\lambda; d) = \frac{1}{2\pi} \left(2 \sin \frac{|\lambda|}{2} \right)^{-2d}. \quad (2.19)$$

Also, since the support of φ_g lies inside $(-1, 1)$, the spectral density $g(\lambda; \kappa)$ corresponding to $\varphi_g(x; \kappa)$ is analytic function (see Proposition 3.3 in Celov et al. (2007)).

Consider the spectral density given by

$$f(\lambda) = f(\lambda; d)g(\lambda; \kappa), \quad \lambda \in [-\pi, \pi]. \quad (2.20)$$

It can be shown that the mixture density $\varphi(x)$ associated with $f(\lambda)$ (2.20) is supported on $[-a_*, 1]$, satisfies Assumption B with $\psi(x)$ which is continuous function on $[-a_*, 1]$ and at the neighborhood of zero satisfies $\psi(x) = O(|x|^d)$. This implies the validity of condition (2.6) needed to obtain the corresponding α -Gegenbauer expansion. For the proof of this example and precise asymptotics of $\psi(x)$ at zero see Appendix A.

Finally, the aggregated process X , obtained using such mixture density $\varphi(x)$, satisfies Assumption A by Proposition 3.2, which shows that assumptions A and B are satisfied under general 'aggregated' spectral density $f(\lambda) = f(\lambda; d)g(\lambda)$, where $g(\lambda)$ is analytic function on $[-\pi, \pi]$ and the associated mixture density is supported on $[-a_*, 0]$ with some $0 < a_* < 1$.

REMARK 2.2 Note that the 'FARIMA mixture density' (2.17), due to factor x^{d-1} , does not satisfy (2.6) and a "compensating" density such as $\varphi_g(x; \kappa)$ in (2.18) is needed in order to obtain the needed integrability in the neighborhood of zero. Obviously, for the same aim, other mixture densities instead of $\varphi_g(x; \kappa)$ (2.18) can be employed.

3 Moving average representation of the aggregated process

In order to obtain the asymptotic normality result in Theorem 2.1, an important assumption is that the aggregated process admits a linear representation with coefficients decaying at an appropriate rate (see Bhansali et al. (2007)). The related issues about the moving average representation of the aggregated process are discussed in this section.

From the aggregating scheme follows that any aggregated process admits an absolutely continuous spectral measure. If, in addition, its spectral density, say, $f(\lambda)$ satisfies

$$\int_{-\pi}^{\pi} \log f(\lambda) d\lambda > -\infty, \quad (3.1)$$

then the function

$$h(z) = \exp \left\{ \frac{1}{4\pi} \int_{-\pi}^{\pi} \frac{e^{i\lambda} + z}{e^{i\lambda} - z} \log f(\lambda) d\lambda \right\}, \quad |z| < 1,$$

is an outer function from the Hardy space H^2 , does not vanish for $|z| < 1$ and $f(\lambda) = |h(e^{i\lambda})|^2$. Then, by the Wold decomposition theorem, corresponding process X_t is purely nondeterministic and has the MA(∞) representation (see Anderson (1971, Ch. 7.6.3))

$$X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}, \quad (3.2)$$

where the coefficients ψ_j are defined from the expansion of normalized outer function $h(z)/h(0)$, $\sum_{j=0}^{\infty} \psi_j^2 < \infty$, $\psi_0 = 1$, and $Z_t = X_t - \hat{X}_t$, $t = 0, 1, \dots$

(\widehat{X}_t is the optimal linear predictor of X_t) is the innovation process, which is zero mean, uncorrelated, with variance

$$\sigma^2 = 2\pi \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f(\lambda) d\lambda \right\}. \quad (3.3)$$

By construction, the aggregated process is Gaussian, implying that the innovations Z_t are i.i.d. $N(0, \sigma^2)$ random variables.

Next we focus on the class of semiparametric mixture densities satisfying Assumption B. As it was mentioned earlier, this form is natural, in particular it covers the mixture densities $\varphi_1(x; d)$ and $\varphi(x)$ in Example 2.1.

Proposition 3.1 *Let the mixture density $\varphi(x)$ satisfies Assumption B. Assume that either*

(i) $\text{supp}(\psi) = [-1, 1]$ and $\tilde{\psi}(x) \equiv \psi(x)(1+x)^{2\tilde{d}-1}$ is continuous at -1 and $\tilde{\psi}(-1) \neq 0$ with some $0 < \tilde{d} < 1/2$,

or

(ii) $\text{supp}(\psi) \subset [-a_*, 1]$ with some $0 < a_* < 1$.

Then the aggregated process admits a moving average representation (3.2), where the Z_t are Gaussian i.i.d. random variables with zero mean and variance (3.3).

PROOF. (i) We have to verify that (3.1) holds. Rewrite $\varphi(x)$ in the form

$$\varphi(x) = (1-x)^{1-2d}(1+x)^{1-2\tilde{d}}\tilde{\psi}(x).$$

Proposition 4.1 in Celov et al. (2007) implies

$$f(\lambda) \sim C_1 |\lambda|^{-2d}, \quad |\lambda| \rightarrow 0,$$

with $C_1 > 0$. Hence $\log f(\lambda) \sim \log C_1 - C_2 \log |1 - e^{i\lambda}|$, $|\lambda| \rightarrow 0$, where $C_2 > 0$. For any $\epsilon > 0$ choose $0 < \lambda_0 \leq \pi/3$, such that

$$-\frac{\log f(\lambda) - \log C_1}{C_2 \log |1 - e^{i\lambda}|} - 1 \geq -\epsilon, \quad 0 < \lambda \leq \lambda_0.$$

Since $-\log |1 - e^{i\lambda}| \geq 0$ for $0 \leq \lambda \leq \pi/3$, we obtain

$$\int_0^{\lambda_0} \log f(\lambda) d\lambda \geq \lambda_0 \log C_1 - C_2(1 - \epsilon) \int_0^{\lambda_0} \log |1 - e^{i\lambda}| d\lambda > -\infty \quad (3.4)$$

using the well known fact that $\int_0^\pi \log |1 - e^{i\lambda}| d\lambda = 0$. Similarly,

$$\int_{\pi-\lambda_0}^\pi \log f(\lambda) d\lambda > -\infty. \quad (3.5)$$

When $\lambda \in [\lambda_0, \pi - \lambda_0]$, there exist $0 < L_1 < L_2 < \infty$ such that

$$L_1 \leq \frac{1}{2\pi|1 - xe^{i\lambda}|^2} \leq L_2$$

uniformly in $x \in (-1, 1)$. Thus, by (2.5), $L_1 \leq f(\lambda) \leq L_2$ for any $\lambda \in [\lambda_0, \pi - \lambda_0]$, and therefore

$$\int_{\lambda_0}^{\pi-\lambda_0} \log f(\lambda) d\lambda > -\infty. \quad (3.6)$$

(3.4)–(3.6) imply inequality (3.1).

The proof in case (ii) is analogous to (i) and, thus, is omitted. \square

Lemma 3.1 *If the spectral density $g(\lambda)$ of the aggregated process $X_t, t \in \mathbf{Z}$ is analytic function on $[-\pi, \pi]$, then X_t admits representation*

$$X_t = \sum_{j=0}^{\infty} g_j Z_{t-j},$$

where the Z_t are i.i.d. Gaussian random variables with zero mean and variance

$$\sigma_g^2 = 2\pi \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \log g(\lambda) d\lambda \right\} \quad (3.7)$$

and the g_j satisfy $|\sum_{j=0}^{\infty} g_j| < \infty$, $g_0 = 1$.

PROOF. From Proposition 3.3 in Celov et al. (2007) it follows that there exists $0 < a_* < 1$ such that

$$g(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} \int_{-a_*}^{a_*} \frac{\varphi_g(x)}{|1 - xe^{i\lambda}|^2} dx. \quad (3.8)$$

For all $x \in [-a_*, a_*]$ and $\lambda \in [0, \pi]$ we have

$$\frac{1}{|1 - xe^{i\lambda}|^2} \geq C_3 > 0,$$

where $C_3 = C_3(a_*)$. This and (3.8) imply $\int_0^\pi \log g(\lambda) d\lambda > -\infty$. Finally, $|\sum_{j=0}^\infty g_j| < \infty$ follows from representation

$$g(\lambda) = \frac{\sigma_g^2}{2\pi} \left| \sum_{j=0}^\infty g_j e^{ij\lambda} \right|^2$$

and the assumption of analyticity of g . \square

Proposition 3.2 *Let $X_t, t \in \mathbf{Z}$ be an aggregated process with spectral density*

$$f(\lambda) = f(\lambda; d)g(\lambda), \quad (3.9)$$

where $f(\lambda; d)$ is FARIMA spectral density (2.19) and $g(\lambda)$ is analytic spectral density. Then:

(i) if mixture density $\varphi_g(x)$ associated with $g(\lambda)$ satisfies $\text{supp}(\varphi_g) \subset [-a_*, 0]$ with some $0 < a_* < 1$, then $\varphi(x)$, associated with $f(\lambda)$, satisfies Assumption B.

(ii) X_t admits a linear representation (3.2), where the Z_t are Gaussian i.i.d. random variables with zero mean and variance

$$\sigma^2 = 2\pi \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^\pi \log f(\lambda) d\lambda \right\} = \exp \left\{ \frac{1}{2\pi} \int_{-\pi}^\pi \log g(\lambda) d\lambda \right\} = \frac{\sigma_g^2}{2\pi}$$

and the coefficients ψ_j satisfy

$$\psi_j \sim \frac{\sum_{k=0}^\infty g_k}{\Gamma(d)} j^{d-1}, \quad |\psi_j - \psi_{j+1}| = O(j^{d-2}), \quad (3.10)$$

where $\psi_0 = 1$. (Here, the g_k are given in Lemma 3.1.)

PROOF. (i) By Corollary 3.1 in Celov et al. (2007), the mixture density associated with the "product" spectral density (3.9) exists and has a form

$$\varphi(x) = C_*^{-1} \left(\varphi(x; d) \int_{-a_*}^0 \frac{\varphi_g(y) dy}{(1-xy)(1-y/x)} + \varphi_g(x) \int_0^1 \frac{\varphi(y; d) dy}{(1-xy)(1-y/x)} \right), \quad (3.11)$$

with

$$C_* := \int_0^1 \left(\int_{-a_*}^0 \frac{\varphi(x; d)\varphi_g(y)}{1-xy} dy \right) dx, \quad (3.12)$$

where $\varphi(x; d)$ is given in (2.17) and is associated with the spectral density $f(\lambda; d)$, and $\varphi_g(x)$ is associated with the spectral density $g(\lambda)$. Clearly, this implies that Assumption B is satisfied.

(ii) We have

$$f(\lambda; d) = \frac{1}{2\pi} \left| \sum_{j=0}^{\infty} h_j e^{ij\lambda} \right|^2 \quad \text{with} \quad h_j = \frac{\Gamma(j+d)}{\Gamma(j+1)\Gamma(d)}$$

and, recall,

$$g(\lambda) = \frac{\sigma_g^2}{2\pi} \left| \sum_{j=0}^{\infty} g_j e^{ij\lambda} \right|^2, \quad \sum_{j=0}^{\infty} g_j^2 < \infty$$

since, by Lemma 3.1, $\int_{-\pi}^{\pi} \log g(\lambda) d\lambda > -\infty$. On the other hand, $\int_{-\pi}^{\pi} \log f(\lambda) d\lambda > -\infty$ implies

$$f(\lambda) = \frac{1}{2\pi} \left| \sum_{j=0}^{\infty} \tilde{\psi}_j e^{ij\lambda} \right|^2, \quad \sum_{j=0}^{\infty} \tilde{\psi}_j^2 < \infty$$

and, by uniqueness of the representation,

$$\tilde{\psi}_k = \frac{\sigma_g}{\sqrt{2\pi}} \sum_{j=0}^k h_{k-j} g_j.$$

It easy to see that,

$$\sum_{j=0}^k h_{k-j} g_j \sim h_k \sum_{j=0}^{\infty} g_j \sim C_4 k^{d-1}, \quad (3.13)$$

where $C_4 = \Gamma^{-1}(d) \sum_{j=0}^{\infty} g_j$. Indeed, taking into account that $h_k \sim \Gamma^{-1}(d) k^{d-1}$, we can write

$$\sum_{j=0}^k h_{k-j} g_j = \Gamma^{-1}(d) k^{d-1} \sum_{j=0}^{\infty} a_{k,j} g_j,$$

where $a_{k,j} = h_{k-j} \Gamma(d) k^{1-d} \mathbf{1}_{\{j \leq k\}} \rightarrow 1$ as $k \rightarrow \infty$ for each j . On the other hand, we have $|a_{k,j}| \leq C(1+j)^{1-d}$ uniformly in k and, since the g_j decay

exponentially fast, the sum $\sum_{j=0}^{\infty}(1+j)^{1-d}|g_j|$ converges and the dominated convergence theorem applies to obtain (3.13).

Hence, we can write

$$f(\lambda) = \frac{\sigma_g^2}{(2\pi)^2} \left| \sum_{j=0}^{\infty} \psi_j e^{ij\lambda} \right|^2, \quad \psi_0 = 1,$$

where $\psi_j = \tilde{\psi}_j \sqrt{2\pi}/\sigma_g \sim C_4 j^{d-1}$. Thus, representation (3.2) and the first relation in (3.10) follows.

Finally, in order to check the second relation in (3.10), it suffices to note that

$$\psi_j - \psi_{j+1} = \sum_{i=0}^j (h_{j-i} - h_{j+1-i}) g_i - g_{j+1},$$

where $h_j - h_{j+1} \sim C_5 j^{d-2}$ and the g_j decay exponentially fast. \square

4 Proof of main result

In order to prove Theorem 2.1, we use the result of Bhansali et al. (2007), who considered the following quadratic form

$$Q_{n,X} = \sum_{t,s=1}^n d_n(t-s) X_t X_s,$$

where the X_t are linear sequences satisfying Assumption A and the function $d_n(k)$ satisfies the following assumption.

ASSUMPTION C Suppose that

$$d_n(k) = \int_{-\pi}^{\pi} \eta_n(\lambda) e^{ik\lambda} d\lambda$$

with some even real function $\eta_n(\lambda)$, such that, for some $-1 < \beta < 1$ and a sequence of constants $m_n \geq 0$, it holds

$$|\eta_n(\lambda)| \leq m_n |\lambda|^{-\beta}, \quad \lambda \in [-\pi, \pi]. \quad (4.1)$$

Denote by E_n a matrix $(e_n(t-s))_{t,s=1,\dots,n}$, where

$$e_n(t-s) = \int_{-\pi}^{\pi} \eta_n(\lambda) f(\lambda) e^{i\lambda(t-s)} d\lambda \quad (4.2)$$

and let $\|E_n\|^2 = \sum_{t,s=1}^n e_n^2(t-s)$.

Theorem 4.1 [Bhansali et al. (2007)] *Suppose that assumptions A and C are satisfied. If $2d + \beta < 1/2$ and*

$$r_n = o(\|E_n\|), \quad (4.3)$$

where

$$r_n = \begin{cases} m_n n^{\max(0, 2d+\beta)} & \text{if } 2d + \beta \neq 0, \\ m_n \log n & \text{if } 2d + \beta = 0, \end{cases} \quad (4.4)$$

then, as $n \rightarrow \infty$, it holds

$$\text{Var}(Q_{n,X}) \asymp \|E_n\|^2$$

and

$$\frac{Q_{n,X} - \mathbb{E}Q_{n,X}}{\sqrt{\text{Var}(Q_{n,X})}} \xrightarrow{d} \mathbb{N}(0, 1).$$

(Here for $a_n, b_n \geq 0$, $a_n \asymp b_n$ means that $C_6 b_n \leq a_n \leq C_7 b_n$ for some $C_6, C_7 > 0$.)

PROOF OF THEOREM 2.1. First of all, note that

$$\hat{\sigma}_{n,\varepsilon}^2 \xrightarrow{P} \sigma_\varepsilon^2,$$

which easily follows using Theorem 3 in Hosking (1996). Hence, to obtain convergence (2.16), we can replace the factor $\hat{\sigma}_{n,\varepsilon}^2$ by σ_ε^2 in the definition of $\hat{\varphi}_n(x)$. Without loss of generality assume that $\sigma_\varepsilon^2 = 1$.

Rewrite the estimate $\hat{\varphi}_n(x)$ in a form

$$\begin{aligned} \hat{\varphi}_n(x) &= (1-x^2)^\alpha \sum_{k=0}^{K_n} \sum_{j=0}^k g_{k,j}^{(\alpha)} (\hat{\sigma}_n(j) - \hat{\sigma}_n(j+2)) G_k^{(\alpha)}(x) \\ &= (1-x^2)^\alpha \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \sum_{j=0}^k g_{k,j}^{(\alpha)} \int_{-\pi}^{\pi} (e^{i\lambda j} - e^{i\lambda(j+2)}) I_n(\lambda) d\lambda \\ &= \int_{-\pi}^{\pi} \eta_n(\lambda; x) I_n(\lambda) d\lambda, \end{aligned} \quad (4.5)$$

where

$$\eta_n(\lambda; x) := (1 - x^2)^\alpha \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \sum_{j=0}^k g_{k,j}^{(\alpha)}(e^{i\lambda j} - e^{i\lambda(j+2)}) \quad (4.6)$$

and $I_n(\lambda) = (2\pi n)^{-1} |\sum_{j=1}^n X_j e^{ij\lambda}|^2$, $\lambda \in [-\pi, \pi]$ is the periodogram.

Now the proof follows from Assumption A and the results obtained in Lemma 4.1 and Lemma 4.2 below, which imply that, under appropriate choice of m_n and β , all the assumptions in Theorem 4.1 are satisfied. In particular, by Lemma 4.1, the following bound for the kernel $\eta_n(\lambda; x)$ holds

$$|\eta_n(\lambda; x)| \leq m_n |\lambda|^{-\beta}, \quad (4.7)$$

where

$$m_n = C_8 n^{\gamma \log(1+\sqrt{2})}, \quad \beta = \frac{\alpha}{2} - \frac{3}{4}, \quad (4.8)$$

C_8 is a positive constant, depending on x and α . Clearly, (2.14) implies that $-1 < \beta \leq \frac{1}{2} - 2d < \frac{1}{2}$ and $2d + \beta < \frac{1}{2}$.

Consider the cases $2d + \beta \leq 0$ or $0 < 2d + \beta < 1/2$. In the case $2d + \beta \leq 0$, from (4.4), (4.8) we obtain

$$r_n = C_8 \begin{cases} n^{\gamma \log(1+\sqrt{2})} & \text{if } 2d + \frac{\alpha}{2} - \frac{3}{4} < 0, \\ n^{\gamma \log(1+\sqrt{2})} \log n & \text{if } 2d + \frac{\alpha}{2} - \frac{3}{4} = 0. \end{cases}$$

Hence, by Lemma 4.2, $r_n \|E_n\|^{-1} \rightarrow 0$ because $\gamma \log(1 + \sqrt{2}) < 1/2$.

Assume now $2d + \beta > 0$. Then

$$r_n = C_8 n^{\gamma \log(1+\sqrt{2}) + 2d + \frac{\alpha}{2} - \frac{3}{4}}$$

and $r_n \|E_n\|^{-1} \rightarrow 0$ by (2.15). \square

The following lemma shows that the kernel $\eta_n(\lambda; x)$ given in (4.6) satisfies inequality (4.7) with m_n and β given in (4.8).

Lemma 4.1 *For quantity $\eta_n(\lambda; x)$ given in (4.6) and for every fixed $x \in (-1, 1)$, $0 < |\lambda| < \pi$ it holds*

$$|\eta_n(\lambda; x)| \leq C_9 n^{\gamma \log(1+\sqrt{2})} |\lambda|^{(3-2\alpha)/4} \begin{cases} (1 - x^2)^{\alpha/2 - 1/4} & \text{if } \alpha > -1/2, \\ (1 - x^2)^\alpha & \text{if } -1 < \alpha < -1/2, \end{cases}$$

where C_9 depends on α , and γ is given in (2.9).

Lemma 4.2 *Assume that a mixture density $\varphi(x)$ satisfies condition (2.6) and let $K_n \rightarrow \infty$. Then for every $x \in (-1, 1)$, such that $\varphi(x) \neq 0$ it holds*

$$\|E_n\|^2 \geq C_{10}n(1 + o(1)), \quad (4.9)$$

where $C_{10} > 0$ is positive constant depending on α and x .

Proof of these two lemmas are given in Appendix B.

5 A simulation study

In order to gain further insight into the asymptotic normality property of the mixture density estimator (2.7), in this section we conduct a Monte-Carlo simulation study. Several examples are considered, which correspond to the mixture densities having different shapes (here we do not pose a question which rigorous aggregating schemes lead to the latter).

The following two families of mixture densities

$$\varphi(x) = w\varphi_1(x) + (1 - w)\varphi_2(x), \quad 0 < w < 1,$$

are considered:

- Beta-type mixture densities defined by

$$\varphi_1(x) \propto x^{p_1-1}(1-x)^{q_1-1}\mathbf{1}_{[0,1]}(x), \quad p_1 > 0, \quad q_1 > 0,$$

$$\varphi_2(x) \propto |x|^{p_2-1}(a_* + x)^{q_2-1}\mathbf{1}_{[-a_*,0]}(x), \quad p_2 > 0, \quad q_2 > 0, \quad 0 < a_* < 1;$$

- mixed (Beta and Uniform)-type mixture densities defined by

$$\varphi_1(x) \propto x^{p_3-1}(1-x)^{q_3-1}\mathbf{1}_{[0,1]}(x), \quad p_3 > 0, \quad q_3 > 0,$$

$$\varphi_2(x) = a_*^{-1}\mathbf{1}_{[-a_*,0]}(x), \quad 0 < a_* < 1.$$

In order to construct the mixture density estimator, in the first step, the parameters K_n and α must be chosen. Preliminary Monte-Carlo simulations showed that the estimator $\hat{\varphi}_n(x)$ has the minimal mean integrated square error (MISE) when the parameter α is chosen to be equal $1 - 2d$. The justification of this interesting conjecture remains an open problem. This rule

also ensures that (2.14) is satisfied. The number of Gegenbauer polynomials K_n is chosen according to (2.9). Note that, by construction, the estimator $\hat{\varphi}_n(x)$ is not necessary positive, though it integrates to one.

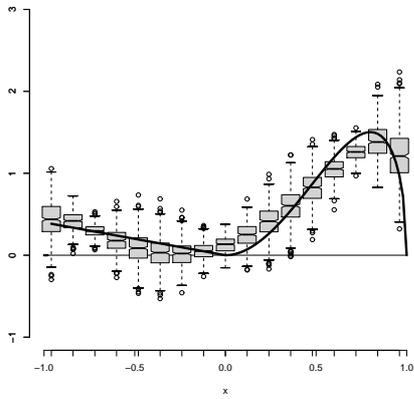
In Figure 1, we present three graphs and corresponding box plots for the mixture densities of the form above. Cases 1 and 2 correspond to the Beta-type mixture densities, Case 3 corresponds to the mixed (Beta and Uniform)-type mixture density. The parameter values are presented in Table 1. The box plots are obtained by a Monte-Carlo procedure based on $M = 500$ independent replications with sample size $n = 1500$ and bandwidth $K_n = 3$ (we aggregate $N = 5000$ i.i.d. AR(1) processes). Individual innovations $\varepsilon_t^{(j)}$ are i.i.d. $N(0, 1)$. Note that the mixture density in Case 2 corresponds to Example 2.1 with the parameters $d = 0.2$, $\kappa = 0.1$ (in the sense of behavior at zero).

	w	a_*	(p_1, q_1)	(p_2, q_2)	(p_3, q_3)	d	α
Case 1	0.8	0.95	(3.0, 1.5)	(2.0, 1.0)	–	0.25	0.5
Case 2	0.8	0.80	(1.2, 1.6)	(1.3, 2.5)	–	0.20	0.6
Case 3	0.8	0.90	–	–	(2.0, 1.2)	0.40	0.2

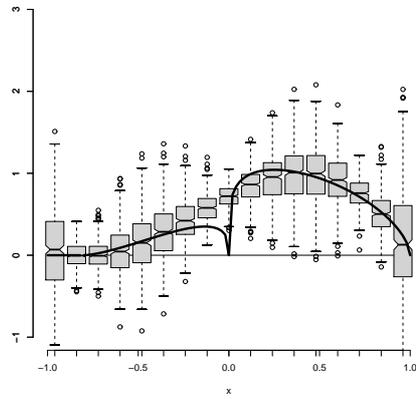
Table 1: Parameter values in cases 1–3.

Box plots in Figure 1 show that $\hat{\varphi}_n$ approximates the mixture density well when n is sufficiently large. However, when the sample size is relatively small it is difficult to estimate the mixture density of the shape as in cases 2–3. This can be explained by the construction of the estimator which assumes rather smooth form of the mixture density around zero. On the other hand, it is clear that the AR(1) parameter values which are close to zero does not affect the long memory property. For our purposes, an important fact is that the estimator correctly approximates the density at the neighborhood of $x = 1$. This enables us to estimate the unknown (in real applications) parameter d using a log–log regression on periodogram at the neighborhood of this point (for example Geweke and Porter-Hudak or Whittle-type estimators).

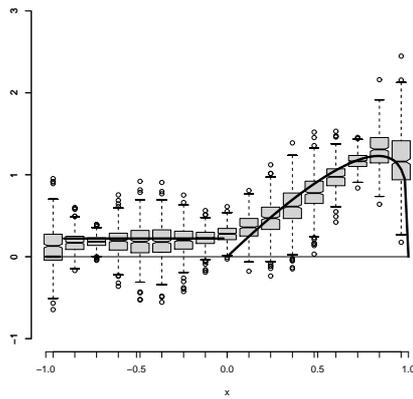
Figure 2 supplements the earlier findings and shows that the distribution



(a) Case 1



(b) Case 2



(c) Case 3

Figure 1: True mixture densities (solid line) and the box plots of the estimates. Number of replications $M = 500$, sample size $n = 1500$.

of estimator is approximately normal.¹ QQ-plots and histograms are given

¹The Shapiro-Wilk test confirms that in most cases normality hypothesis is consistent with the data.

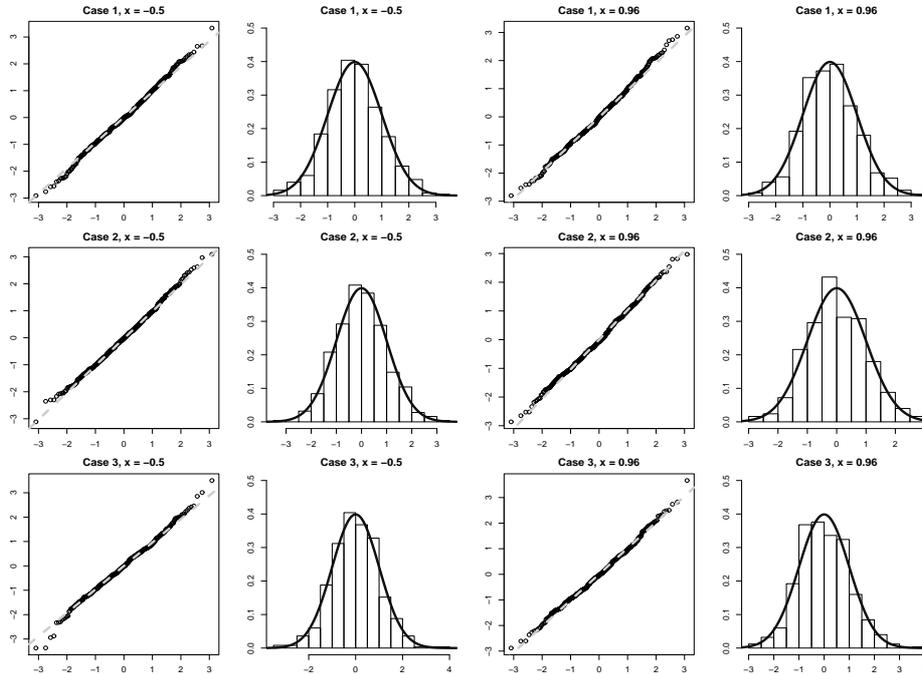


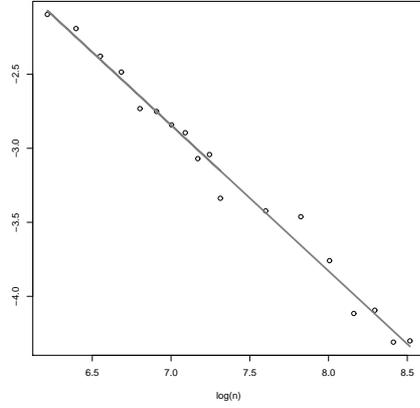
Figure 2: QQ plots and histograms of the estimates at points $x = -0.5$ and $x = 0.96$. Number of replications $M = 500$, sample size $n = 1500$.

for fixed values $x = -0.5$ and $x = 0.96$ correspondingly. We use the same number of replications $M = 500$ and sample size $n = 1500$.

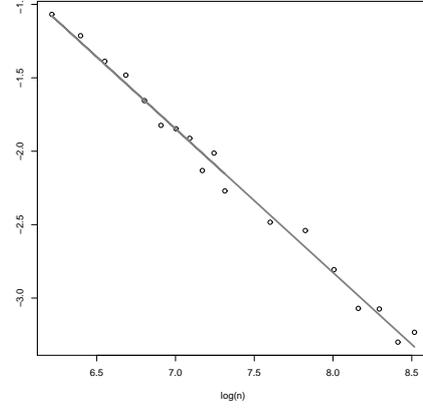
The last Monte-Carlo experiment aims to show that the decay rate of $\text{Var}(\hat{\varphi}_n(x))$ is $n^{-\gamma}$ with $\gamma = 1$. This ensures that the variance is decreasing fast enough. To do this, we calculate the log-log regression of variance on the length of time series $n \in \{500, 600, \dots, 1400, 1500, 2000, \dots, 5000\}$. Figure 3 demonstrates the corresponding parameter estimates at different points and shows that $\hat{\gamma} \approx 1$.

6 Appendix A. Proof of Example 2.1

By Corollary 3.1 in Celov et al. (2007), the mixture density $\varphi(x)$, $x \in [-a_*, 1]$ associated with $f(\lambda)$ (2.20) is given by equality (3.11), where $\varphi_g(x) \equiv$



(a) Case 1, $x = -0.5$



(b) Case 1, $x = 0.96$

Figure 3: log-log scale regression of the variance of $\hat{\varphi}_n(x)$ as a function of n . The variance is estimated using $M = 500$ independent replications.

$\varphi_g(x; \kappa)$. Clearly, in this case, (3.11) can be rewritten in form (2.13) with

$$\psi(x) = \tilde{C}(\psi_1(x) + \psi_2(x)), \quad (6.1)$$

where $\tilde{C} = C_1(d)C_2(\kappa)C_*^{-1}$ is positive constant,

$$\psi_1(x) := x^{d-1}(1+x)\mathbf{1}_{(0,1]}(x) \int_{-a_*}^0 \frac{|y|^\kappa}{(1-xy)(1-y/x)} dy, \quad (6.2)$$

$$\psi_2(x) := |x|^\kappa(1-x)^{2d-1}\mathbf{1}_{[-a_*,0]}(x) \int_0^1 \frac{y^{d-1}(1-y)^{1-2d}(1+y)}{(1-xy)(1-y/x)} dy \quad (6.3)$$

Denote by $F(a, b; c; x)$ a hypergeometric function

$$F(a, b; c; x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-tx)^{-a} dt,$$

with $c > b > 0$ if $x < 1$ and, in addition, $c - a - b > 0$ if $x = 1$. Then the corresponding integrals in $\psi_1(x)$ and $\psi_2(x)$ can be rewritten as

$$\begin{aligned} & \int_{-a_*}^0 \frac{|y|^\kappa}{(1-xy)(1-y/x)} dy, \\ &= \frac{a_*^{\kappa+1}}{\kappa+1} \frac{x(F(1, \kappa+1; \kappa+2; -a_*x) - F(1, \kappa+1; \kappa+2; -a_*/x))}{1-x^2} \\ &\sim \frac{a_*^{\kappa+1}}{\kappa+1} x, \quad \text{as } x \rightarrow 0+, \end{aligned}$$

and

$$\begin{aligned} & \int_0^1 \frac{y^{d-1}(1-y)^{1-2d}(1+y)}{(1-xy)(1-y/x)} dy \\ &= \frac{\Gamma(d)\Gamma(2-2d)}{\Gamma(2-2d)} \frac{F(1, d; 2-d; 1/x) - xF(1, d; 2-d; x)}{1-x} \\ &\sim \Gamma(d)\Gamma(1-d)|x|^d, \quad \text{as } x \rightarrow 0-, \end{aligned}$$

where the last asymptotics follow from the well known properties of the hypergeometric functions (see Abramovitz and Stegun (1965)).

Thus, from (6.2)–(6.3) we obtain that

$$\psi_1(x) \sim \frac{a_*^{\kappa+1}}{\kappa+1} x^d, \quad \text{as } x \rightarrow 0+, \quad (6.4)$$

$$\psi_2(x) \sim \Gamma(d)\Gamma(1-d)|x|^{\kappa+d}, \quad \text{as } x \rightarrow 0-. \quad (6.5)$$

(6.1) and relations (6.4)–(6.5) complete the proof. \square

7 Appendix B. Proofs of lemmas 4.1–4.2

PROOF OF LEMMA 4.1. By (4.6),

$$\begin{aligned} (1-x^2)^{-\alpha} \eta_n(\lambda; x) &= \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \sum_{j=0}^k g_{k,j}^{(\alpha)} (e^{i\lambda j} - e^{i\lambda(j+2)}) \\ &= (1 - e^{2i\lambda}) \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \sum_{j=0}^k g_{k,j}^{(\alpha)} e^{i\lambda j} \\ &= (1 - e^{2i\lambda}) \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) G_k^{(\alpha)}(e^{i\lambda}). \end{aligned}$$

This and Lemma 7.1 below implies

$$(1-x^2)^{-\alpha} |\eta_n(\lambda; x)| \leq C_{11} |\lambda|^{-(2\alpha-3)/4} \sum_{k=0}^{K_n} |G_k^{(\alpha)}(x)| (1+\sqrt{2})^k. \quad (7.1)$$

Now, using the fact that for all $-1 < x < 1$

$$|G_k^{(\alpha)}(x)| \leq \begin{cases} C_{12} (1-x^2)^{-\frac{\alpha}{2}-\frac{1}{4}} & \text{if } \alpha > -1/2 \\ C_{12} & \text{if } \alpha < -1/2, \alpha \neq -3/2, -5/2, \dots \end{cases}$$

(see inequality (7.33.6) in Szegö (1967) and (3.9) in Leipus et al. (2006)) and (2.9), we get from (7.1)

$$\begin{aligned} (1-x^2)^{-\alpha} |\eta_n(\lambda; x)| &\leq C_{13} |\lambda|^{-(2\alpha-3)/4} (1+\sqrt{2})^{K_n} \\ &= C_{13} |\lambda|^{-(2\alpha-3)/4} e^{K_n \log(1+\sqrt{2})} \\ &\leq C_9 |\lambda|^{-(2\alpha-3)/4} n^{\gamma \log(1+\sqrt{2})}. \end{aligned}$$

□

Lemma 7.1 For all $k \geq 0$, $\alpha > -1$, ($\alpha \neq -1/2$) and $0 < |\lambda| < \pi$ it holds

$$|(1 - e^{2i\lambda}) G_k^{(\alpha)}(e^{i\lambda})| \leq C_{11} (1 + \sqrt{2})^k |\lambda|^{-(2\alpha-3)/4},$$

where constant C_{11} depends on α .

PROOF. Theorem 8.21.10 of Szegö (1967) implies that for the usual (non-normalized) Gegenbauer polynomials with $\alpha > -1$, $\alpha \neq -1/2$ it holds

$$C_k^{(\alpha+1/2)}(e^{i\lambda}) = \frac{\Gamma(k + \alpha + \frac{1}{2})}{\Gamma(k+1)\Gamma(\alpha + \frac{1}{2})} z^k (1-z^{-2})^{-\alpha-1/2} + O(k^{\alpha-3/2} |z|^k), \quad (7.2)$$

where the complex numbers $w = e^{i\lambda}$ and z are connected by the elementary conformal mapping

$$w = \frac{1}{2}(z + z^{-1}), \quad z = w + (w^2 - 1)^{1/2}, \quad (7.3)$$

and z satisfies $|z| > 1$ (thus, $\lambda \neq 0, \pm\pi$).

Recall that the normalized Gegenbauer polynomials $G_k^{(\alpha)}(z)$ are linked to $C_k^{(\alpha+1/2)}(z)$ by equality

$$G_k^{(\alpha)}(z) = \gamma_k^{-1/2} C_k^{(\alpha+1/2)}(z), \text{ where } \gamma_k = \frac{\pi}{2^{2\alpha}} \frac{\Gamma(k+2\alpha+1)}{(k+\alpha+\frac{1}{2})\Gamma^2(\alpha+\frac{1}{2})\Gamma(k+1)}.$$

Therefore, in terms of the normalized Gegenbauer polynomials, (7.2) reads as follows

$$G_k^{(\alpha)}(e^{i\lambda}) = \frac{\operatorname{sgn}(\alpha+1/2)2^\alpha}{\pi^{1/2}} b_k z^k (1-z^{-2})^{-\alpha-1/2} + O(k^{-1}|z|^k), \quad (7.4)$$

where

$$b_k = \frac{(k+\alpha+1/2)^{1/2}\Gamma(k+\alpha+1/2)}{\Gamma^{1/2}(k+1)\Gamma^{1/2}(k+2\alpha+1)} \rightarrow 1 \text{ as } k \rightarrow \infty.$$

From (7.3) we obtain for $w = e^{i\lambda}$

$$w^2 - 1 = \frac{1}{4} z^2 (1 - z^{-2})^2,$$

which together with (7.4) yields

$$(1 - e^{2i\lambda})G_k^{(\alpha)}(e^{i\lambda}) = -\frac{\operatorname{sgn}(\alpha+1/2)2^\alpha}{4\pi^{1/2}} b_k z^{k+2} (1 - z^{-2})^{-\alpha+3/2} + O(k^{-1}|z|^k).$$

Since $|z| > 1$ and $z^2 - 1 = 2(e^{2i\lambda} - 1) + 2e^{3i\lambda/2}(e^{i\lambda} - e^{-i\lambda})^{1/2}$, we have

$$\begin{aligned} |1 - z^{-2}| &\leq |z^2 - 1| \\ &\leq 2|e^{2i\lambda} - 1| + 2|e^{i\lambda} - e^{-i\lambda}|^{1/2} \\ &= 4|\sin \lambda| + 2\sqrt{2}|\sin \lambda|^{1/2} \\ &\leq (4 + 2\sqrt{2})|\lambda|^{1/2}. \end{aligned} \quad (7.5)$$

So that, by (7.4)–(7.5),

$$|(1 - e^{2i\lambda})G_k^{(\alpha)}(e^{i\lambda})| \leq C_{14} b_k |z|^k |\lambda|^{-(2\alpha-3)/4}, \quad (7.6)$$

where $C_{14} = C_{14}(\alpha)$.

Finally, the straightforward verification shows that

$$\sup_{\lambda \in [-\pi, \pi]} |e^{i\lambda} + (e^{2i\lambda} - 1)^{1/2}| = 1 + \sqrt{2}.$$

This completes the proof of lemma. \square

PROOF OF LEMMA 4.2. Using (4.2), (4.6) rewrite the coefficients of E_n

$$e_n(t-s) = (1-x^2)^\alpha \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \sum_{j=0}^k g_{k,j}^{(\alpha)} \int_{-\pi}^{\pi} f(\lambda) (e^{i\lambda(t-s+j)} - e^{i\lambda(t-s+j+2)}) d\lambda.$$

Using the expression of the covariance function of an aggregated process, we have for $t-s+j \geq 0$

$$\begin{aligned} \int_{-\pi}^{\pi} f(\lambda) (e^{i\lambda(t-s+j)} - e^{i\lambda(t-s+j+2)}) d\lambda &= \sigma(t-s+j) - \sigma(t-s+j+2) \\ &= \sigma_\varepsilon^2 \int_{-1}^1 y^{t-s+j} \varphi(y) dy. \end{aligned}$$

Thus, assuming $\sigma_\varepsilon^2 = 1$, for $t-s \geq 0$ we have

$$\begin{aligned} e_n(t-s) &= (1-x^2)^\alpha \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \sum_{j=0}^k g_{k,j}^{(\alpha)} \int_{-1}^1 y^{t-s+j} \varphi(y) dy \\ &= (1-x^2)^\alpha \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \int_{-1}^1 y^{t-s} \varphi(y) \sum_{j=0}^k g_{k,j}^{(\alpha)} y^j dy \\ &= (1-x^2)^\alpha \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \int_{-1}^1 y^{t-s} \varphi(y) G_k^{(\alpha)}(y) dy. \end{aligned}$$

Integral $\int_{-1}^1 y^m \varphi(y) G_k^{(\alpha)}(y) dy$ (m is a nonnegative integer), appearing in the last expression is nothing else but the k th coefficient, $\psi_{m,k}$, in the α -Gegenbauer expansion of the function

$$\psi_m(x) = \frac{x^m \varphi(x)}{(1-x^2)^\alpha}, \quad (7.7)$$

which obviously satisfies $\psi_m \in L^2(w^{(\alpha)})$. Therefore,

$$\begin{aligned} e_n(t-s) &= (1-x^2)^\alpha \sum_{k=0}^{K_n} G_k^{(\alpha)}(x) \psi_{|t-s|,k} \\ &= (1-x^2)^\alpha \left(\psi_{|t-s|}(x) - \sum_{k=K_n+1}^{\infty} G_k^{(\alpha)}(x) \psi_{|t-s|,k} \right) \end{aligned}$$

and, denoting $R_n(m) := \sum_{k=K_n+1}^{\infty} G_k^{(\alpha)}(x)\psi_{|m|,k}$, $|m| < n$, we have

$$\begin{aligned} (1-x^2)^{-2\alpha}\|E_n\|^2 &= \sum_{|m|<n} (n-|m|) \left(\psi_{|m|}(x) - \sum_{k=K_n+1}^{\infty} G_k^{(\alpha)}(x)\psi_{|m|,k} \right)^2 \\ &= \sum_{|m|<n} (n-|m|)\psi_{|m|}^2(x) - 2 \sum_{|m|<n} (n-|m|)\psi_{|m|}(x)R_n(m) \\ &\quad + \sum_{|m|<n} (n-|m|)R_n^2(m) =: A_{1,n} - 2A_{2,n} + A_{3,n}. \end{aligned}$$

Now, we prove that, as $n \rightarrow \infty$,

$$A_{1,n} \sim C_{15}n, \quad (7.8)$$

where $C_{15} = C_{15}(x) > 0$ is some positive constant, and

$$A_{2,n} = o(n). \quad (7.9)$$

Since the last term $A_{3,n}$ is nonnegative by construction, this will prove (4.9).

At points x where $\varphi(x) \neq 0$ we have

$$\begin{aligned} A_{1,n} &= \frac{\varphi^2(x)}{(1-x^2)^{2\alpha}} \sum_{|m|<n} (n-|m|)x^{2|m|} \\ &\sim n \frac{\varphi^2(x)(1+x^2)}{(1-x^2)^{2\alpha+1}}, \quad \text{as } n \rightarrow \infty, \end{aligned}$$

which gives (7.8). Consider term $A_{2,n}$. By (7.7),

$$\begin{aligned} A_{2,n} &= \sum_{|m|<n} (n-|m|)\psi_{|m|}(x) \sum_{k=K_n+1}^{\infty} G_k^{(\alpha)}(x)\psi_{|m|,k} \\ &= \frac{\varphi(x)}{(1-x^2)^\alpha} \sum_{k=K_n+1}^{\infty} G_j^{(\alpha)}(x) \int_{-1}^1 \varphi(y)G_k^{(\alpha)}(y) \sum_{|m|<n} (n-|m|)(xy)^{|m|} dy \\ &= \frac{\varphi(x)}{(1-x^2)^\alpha} (B_{1,n} - B_{2,n} - B_{3,n}), \end{aligned}$$

where

$$\begin{aligned}
B_{1,n} &:= n \sum_{k=K_n+1}^{\infty} G_j^{(\alpha)}(x) \int_{-1}^1 \varphi(y) G_k^{(\alpha)}(y) \sum_{m=-\infty}^{\infty} (xy)^{|m|} dy, \\
&= n \sum_{k=K_n+1}^{\infty} G_j^{(\alpha)}(x) \int_{-1}^1 \varphi(y) G_k^{(\alpha)}(y) \frac{1+xy}{1-xy} dy, \\
B_{2,n} &:= \sum_{k=K_n+1}^{\infty} G_j^{(\alpha)}(x) \int_{-1}^1 \varphi(y) G_k^{(\alpha)}(y) \sum_{|m|<n} |m| (xy)^{|m|} dy, \\
B_{3,n} &:= n \sum_{k=K_n+1}^{\infty} G_j^{(\alpha)}(x) \int_{-1}^1 \varphi(y) G_k^{(\alpha)}(y) \sum_{|m|\geq n} (xy)^{|m|} dy.
\end{aligned}$$

Since, by (2.6),

$$\tilde{\varphi}_x(y) \equiv \frac{\varphi(y)}{(1-y^2)^\alpha} \frac{1+xy}{1-xy}$$

satisfies $\tilde{\varphi}_x \in L^2(w^{(\alpha)})$ and $K_n \rightarrow \infty$, the sum $\sum_{k=K_n+1}^{\infty}$ in $B_{1,n}$ vanishes (as the tail of the convergent series). So that, $B_{1,n} = o(n)$ and, similarly, $B_{3,n} = o(n)$.

Finally,

$$B_{2,n} \sim \sum_{k=K_n+1}^{\infty} G_j^{(\alpha)}(x) \int_{-1}^1 \varphi(y) G_k^{(\alpha)}(y) \frac{2xy}{(1-xy)^2} dy = o(1)$$

using the similar argument as in the case of term $B_{1,n}$. This completes the proof of (7.9) and of the lemma. \square

References

- Abramovitz, M. and Stegun, I. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. New York, Dover Publications.
- Anderson, T. (1971). *The Statistical Analysis of Time Series*. Wiley Series in Probability and Mathematical Statistics, New York.
- Bhansali, R. J., Giraitis, L., and Kokoszka, P. S. (2007). Approximations and limit theory for quadratic forms of linear processes. *Stochastic Processes and their Applications*, 117:71–95.

- Celov, D., Leipus, R., and Philippe, A. (2007). Time series aggregation, disaggregation and long memory. *Lithuanian Mathematical Journal* (forthcoming).
- Chong, T. T. (2006). The polynomial aggregated AR(1) model. *Econometrics Journal*, 9:98–122.
- Hosking, J. (1996). Asymptotic distributions of the sample mean, autocovariances and autocorrelations of long-memory time series. *Journal of Econometrics*, 73:261–264.
- Leipus, R., Oppenheim, G., Philippe, A., and Viano, M.-C. (2006). Orthogonal series density estimation in a disaggregation scheme. *Journal of Statistical Planning and Inference*, 136:2547–2571.
- Lewbel, A. (1994). Aggregation and simple dynamics. *American Economic Review*, 84:905–918.
- Oppenheim, G. and Viano, M.-C. (2004). Aggregation of random parameters Ornstein-Uhlenbeck or AR processes: some convergence results. *Journal of Time Series Analysis*, 25:335–350.
- Szegő, G. (1967). *Orthogonal Polynomials*. American Mathematical Society, New York.
- Zaffaroni, P. (2004). Contemporaneous aggregation of linear dynamic models in large economies. *Journal of Econometrics*, 120:75–102.
- Zaffaroni, P. (2006). Contemporaneous aggregation of GARCH processes. *Journal of Time Series Analysis*, 28:521–544.