PUB. IRMA, LILLE 2003
Vol. 60, N° IX

Orthogonal series density estimation in a disaggregation scheme*

Remigijus Leipus[1], George Oppenheim[2], Anne Philippe[3], Marie-Claude Viano[3]

[1] *Institute of Mathematics and Informatics*
Vilnius University,
Naugarduko 24, 2600 Vilnius Lituanie
[2] *Laboratoire de Modélisation Stochastique et Statistique,*
Bât 425, Université de Paris Sud, 91 405 Orsay Cedex, FRANCE,
[3] *Laboratoire de Mathématiques Appliquées CNRS FRE 2222*
Bat M2, Université des Sciences et Technologies,
59655 Villeneuve d'Ascq Cedex, FRANCE

Abstract

In this paper we consider long-memory processes obtained by aggregation of independent random parameter AR(1) processes. We propose an estimator of the density of the underlying random parameter. This estimator is based on the expansion of the density function on the basis of Gegenbauer polynomials. Rate of convergence to zero of the mean integrated square error (MISE) and of the uniform error are obtained. The results are illustrated by Monte-Carlo simulations.

Résumé

Ce papier est consacré à l'étude des processus obtenus par agrégation de processus AR(1) à paramètre aléatoire. Pour cette classe de processus, nous proposons une méthode d'estimation de la loi de mélange. Les estimateurs sont construits à partir du développement de la densité sur la base des polynômes de Gegenbauer. Nous obtenons la convergence du MISE et la convergence uniforme. Quelques simulations illustrent ces résultats.

keywords: Aggregation, mixing density, long memory, random coefficient AR(1), orthogonal Gegenbauer polynomials, problem of the moments.

---

*preprint

R. Leipus, A. Philippe and M.-C. Viano are supported by a cooperation agreement CNRS/LITHUANIA (4714)

1 Introduction

In his seminal paper, Granger (1980) showed that aggregation of random parameter AR(1) processes can produce long-memory. The idea that accumulation of short-memory random variables can lead to long-memory macro-phenomena provides an explanation for the appearance of long-memory in various fields like astronomy, hydrology, economics and finance. For a discussion of long-memory property in the aggregate output of economic time series see Haubrich and Lo (2001) for instance.

Since Granger's work, a number of papers were devoted to theoretical aspects of this problem. For reviews of various aggregation procedures we refer the reader to Oppenheim and Viano (2003) and to Lippi and Zaffaroni (1998). The importance of Granger's idea was recently emphasized by Dacunha-Castelle and Oppenheim (2001), who proved that a large class of long-memory processes are obtained via aggregation of short-memory processes having random spectral densities. In other words, the phenomenon noticed by Granger is far from being exotic.

A question, which is none the less interesting and important, is the so-called *disaggregation* problem: having data from an aggregated process at hand, and assuming that these data are accumulated from i.i.d. short memory dynamics, to recover the distribution of the individual processes. This is the problem considered in our paper.

Let us precise the construction of aggregated processes used in the sequel. Consider a sequence of independent processes $Y^{(j)}$, $j = 1, 2, \ldots$ defined by the random AR(1) dynamics

$$Y_t^{(j)} = a^{(j)} Y_{t-1}^{(j)} + \varepsilon_t^{(j)}, \qquad t \in \mathbb{Z}, \tag{1.1}$$

where

1. $(\varepsilon^{(j)})_{j\geq 1}$ is a sequence of independent copies of a zero mean unit variance strong white noise (by strong white noise we mean that for fixed j the variables $\varepsilon_t^{(j)}$, $t \in \mathbb{Z}$ are i.i.d.);
2. the $a^{(j)}$'s are independent copies of a variable a supported by [-1,1] and satisfying
$$\mathbb{E}\left[\frac{1}{1-a^2}\right] < \infty; \tag{1.2}$$
3. the sequences $(a^{(j)})_{j\geq 1}$ and $(\varepsilon^{(j)})_{j\geq 1}$ are independent.

Under these conditions (1.1) admits a stationary solution $Y_t^{(j)}$ and the finite dimensional distributions of the process

$$X_t^{(N)} = \frac{1}{\sqrt{N}} \sum_{j=1}^{N} Y_t^{(j)}, \qquad t \in \mathbb{Z}, \tag{1.3}$$

weakly converges, as N tends to infinity, to those of a zero mean Gaussian process X_t, called the *aggregated* process (see Oppenheim and Viano (2003) for details).

Hereafter, the distribution of the coefficients $a^{(i)}$'s shall be called the *mixture distribution*. We suppose that this distribution admits a *mixture density* φ, which by (1.2) satisfies

$$\int_{-1}^{1} \frac{\varphi(x)}{1-x^2}\, \mathrm{d}x < \infty. \tag{1.4}$$

In the paper of Oppenheim and Viano (2003) the following form of the mixture density, which clearly satisfies (1.4), was considered

$$\varphi(x) = (1-x)^{d_1}(1+x)^{d_2}\psi(x), \quad d_1 > 0,\ d_2 > 0, \tag{1.5}$$

where $\psi(x)$ is continuous on $[-1,1]$, and does not vanish at ± 1.

Note that both the underlying process $Y_t^{(j)}$ and the aggregated process X_t have the same covariance function

$$\sigma(h) := \mathrm{Cov}(Y_h^{(j)}, Y_0^{(j)}) = \mathrm{Cov}(X_h, X_0) = \int_{-1}^{1} \frac{x^{|h|}}{1-x^2} \varphi(x)\ \mathrm{d}x \tag{1.6}$$

and a spectral density given by

$$f(\lambda) = \frac{1}{2\pi} \int_{-1}^{1} \frac{1}{|1 - xe^{i\lambda}|^2}\ \varphi(x)\ \mathrm{d}x. \tag{1.7}$$

Long memory, that is $\sum_{h=-\infty}^{\infty} |\sigma(h)| = \infty$, is obtained if and only if

$$\int_{-1}^{1} \frac{\varphi(x)}{(1-x^2)^2}\ \mathrm{d}x = \infty, \tag{1.8}$$

or, under the semi-parametric assumption (1.5), whenever d_1 or d_2 is less than 1.

It should be noted that Robinson (1978) presents the same type of models and gives a method to estimate the density φ. However, Robinson's concern is quite different from our's: he assumes that the density belongs to a parametric family $\{\varphi_\theta\}$ and he constructs the estimator of θ from samples of size T of N elementary processes $X_t^{(j)},\ j = 1, \ldots, N$. In the present paper, we estimate φ from a sample of the *limiting* aggregated process X_t, while the samples of the elementary processes $X_t^{(j)}$ $(j = 1, 2, \cdots)$ remain unobserved.

We note that (1.6) means that the covariances are nothing else but the moments of the finite measure having density $(1-x^2)^{-1}\varphi(x)$. In fact, being supported by $[-1, 1]$, this measure is unique (see, e.g., Feller (1971, Chapter 7)), so that the statistical problem of recovering the mixture density from estimates of the covariance function is relevant.

The above remark leads to a relatively natural method which can be summarized as follows: expand a function suitably related to $\varphi(x)$ on a well-chosen basis of polynomials; then use (1.6) and the sample covariances of the process X_t to estimate the first terms in the expansion of $\varphi(x)$.

Estimation of a density function via its expansion on a basis of orthogonal polynomials is a widely used method since the pioneering paper of Čencov (1962) (see also Viollaz (1980) and Hall (1982) among others). However it is important to notice that our situation is far from being standard, as the observations at hand *are not* independent data from a process having the marginal distribution density to be estimated. They form a sample of a process whose covariances are directly linked to the moments of the density $\varphi(x)$.

The paper is organized as follows. In Section 2 we construct the estimator of a mixture density based on the orthogonal Gegenbauer expansion. The convergence results are stated in Section 3. Section 4 provides a simulation study. Appendix A contains a lemma about the second moment of the sample covariance of X_t. Technical results about Gegenbauer polynomials are collected in Appendix B.

2 The estimators

2.1 Gegenbauer expansion of the mixture density

For $\alpha > -1$, consider the finite measure on $[-1, 1]$ having density $w^{(\alpha)}(x) = (1-x^2)^\alpha$. Let $L^2(w^{(\alpha)})$ be the space of functions which are square integrable with respect to this measure. Denote by $\left\{G_n^{(\alpha)},\ n = 0, 1, \ldots\right\}$ the orthonormal basis on $L^2(w^{(\alpha)})$ of Gegenbauer polynomials:

$$\int_{-1}^{1} G_j^{(\alpha)}(x) G_k^{(\alpha)}(x) w^{(\alpha)}(x)\ \mathrm{d}x = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

Any function h in $L^2(w^{(\alpha)})$ has an expansion

$$h(x) = \sum_{j=0}^{\infty} h_j G_j^{(\alpha)}(x) \quad \text{with} \quad h_j = \int_{-1}^{1} h(x) G_j^{(\alpha)}(x) w^{(\alpha)}(x)\ \mathrm{d}x,$$

which we will call the α-Gegenbauer expansion of h.

Denote $g_{n,j}^{(\alpha)}$ the coefficients of the polynomial $G_n^{(\alpha)}(x)$ (see (B.3) and (B.2) in Appendix B for an exact expression). Then

$$G_n^{(\alpha)}(x) = \sum_{j=0}^{n} g_{n,j}^{(\alpha)} x^j \tag{2.1}$$

yields

$$h_k = \int_{-1}^{1} h(x) G_k^{(\alpha)}(x) w^{(\alpha)}(x)\ \mathrm{d}x = \sum_{j=0}^{k} g_{k,j}^{(\alpha)} \int_{-1}^{1} h(x) x^j w^{(\alpha)}(x)\ \mathrm{d}x.$$

Consider now the function

$$\zeta(x) = \frac{\varphi(x)}{(1-x^2)^\alpha}.$$

Under the condition

$$\int_{-1}^{1} \frac{\varphi(x)^2}{(1-x^2)^\alpha}\ \mathrm{d}x < \infty, \tag{2.2}$$

ζ belongs to $L^2(w^{(\alpha)})$, and has an α-Gegenbauer expansion with coefficients

$$\zeta_k = \sum_{j=0}^{k} g_{k,j}^{(\alpha)} \int_{-1}^{1} \varphi(x) x^j\ \mathrm{d}x = \sum_{j=0}^{k} g_{k,j}^{(\alpha)} \left(\sigma(j) - \sigma(j+2)\right), \qquad k \geq 0, \tag{2.3}$$

where equality

$$\int_{-1}^{1} x^j \varphi(x)\ \mathrm{d}x = \sigma(j) - \sigma(j+2), \quad j \geq 0$$

follows from (1.6).

Formula (2.3) leads to the natural idea of replacing the covariances $\sigma(j)$ by their sample counterparts in order to estimate the mixture density via an estimate of $\zeta(x)$.

2.2 Estimation of the mixture density

Let $(X_1, \ldots, X_n)$ be a sample of the aggregated Gaussian process corresponding to the mixture density φ. Using the representation (2.3), a natural estimate of the density $\varphi(x)$ is

$$\widehat{\varphi}_n(x) = (1-x^2)^\alpha \widehat{\zeta}_n(x) = (1-x^2)^\alpha \sum_{k=0}^{K_n} \widehat{\zeta}_{n,k} G_k^{(\alpha)}(x), \tag{2.4}$$

where (K_n) is a non decreasing sequence which tends to infinity at a rate to be discussed and

$$\widehat{\zeta}_{n,k} = \sum_{j=0}^{k} g_{k,j}^{(\alpha)} (\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2)). \tag{2.5}$$

Here, $\widehat{\sigma}_n(j) = \frac{1}{n} \sum_{i=1}^{n-j} X_i X_{i+j}$ is the sample covariance of the zero mean process X_t.

3 Convergence results

3.1 Convergence in L^2 of the estimator

Theorem 1. *Let (X_t) be an aggregated process associated to a mixture density satisfying* (1.4)*, and suppose that $\alpha > -1$ is chosen so that condition* (2.2) *holds. If K_n satisfies*

$$K_n = [\gamma \log n] \quad \text{with} \quad 0 < \gamma < (2\log(1+\sqrt{2}))^{-1}, \tag{3.1}$$

then, for $\widehat{\varphi}_n$ defined in (2.4) and (2.5), it holds

$$\lim_{n\to\infty} \int_{-1}^{1} \frac{\mathbb{E}(\widehat{\varphi}_n(x) - \varphi(x))^2}{(1-x^2)^\alpha}\, \mathrm{d}x = 0. \tag{3.2}$$

Let us give two remarks about the assumptions of the theorem. Firstly, note that, under assumption (1.5), condition (2.2) is equivalent to $\alpha < 2\min(d_1, d_2) + 1$, so that this condition is satisfied for every α in $(-1, 1]$. More generally, if the function φ is bounded, condition (1.4) implies that (2.2) holds for all values of $\alpha \in (-1, 1]$. It is true, for example, in the case of densities used in Oppenheim and Viano (2003). Secondly, returning to the convergence (3.2), it is clear that the denominator $(1-x^2)^\alpha$ gives a better fit for the estimation at ± 1 for large values of α.

Proof of Theorem 1. Note that

$$\begin{aligned} v_n &:= \int_{-1}^{1} \frac{\mathbb{E}(\widehat{\varphi}_n(x) - \varphi(x))^2}{(1-x^2)^\alpha}\, \mathrm{d}x \\ &= \int_{-1}^{1} \mathbb{E}(\widehat{\zeta}_n(x) - \zeta(x))^2 w^{(\alpha)}(x)\, \mathrm{d}x \\ &= \int_{-1}^{1} \mathbb{E}\Big(\sum_{k=0}^{K_n} \widehat{\zeta}_{n,k} G_k^{(\alpha)}(x) - \sum_{k=0}^{\infty} \zeta_k G_k^{(\alpha)}(x)\Big)^2 w^{(\alpha)}(x)\, \mathrm{d}x. \end{aligned}$$

Set $R_n(x) = \sum_{k=K_n+1}^{\infty} \zeta_k G_k^{(\alpha)}(x)$. Then

$$\begin{aligned} v_n &= \int_{-1}^{1} \mathbb{E}\Big(\sum_{k=0}^{K_n} (\widehat{\zeta}_{n,k} - \zeta_k) G_k^{(\alpha)}(x) - R_n(x)\Big)^2 w^{(\alpha)}(x)\, \mathrm{d}x \\ &= \int_{-1}^{1} \mathbb{E}\Big(\sum_{k=0}^{K_n} (\widehat{\zeta}_{n,k} - \zeta_k) G_k^{(\alpha)}(x)\Big)^2 w^{(\alpha)}(x)\, \mathrm{d}x + \int_{-1}^{1} R_n^2(x) w^{(\alpha)}(x)\, \mathrm{d}x \\ &\quad -2\int_{-1}^{1} \mathbb{E}\Big(\sum_{k=0}^{K_n} (\widehat{\zeta}_{n,k} - \zeta_k) G_k^{(\alpha)}(x)\Big) R_n(x) w^{(\alpha)}(x)\, \mathrm{d}x. \end{aligned}$$

By the orthogonality of Gegenbauer polynomials, the last summand is zero and the first one is equal to

$$\int_{-1}^{1} \mathbb{E}\Big(\sum_{k,k'=0}^{K_n} (\widehat{\zeta}_{n,k} - \zeta_k)(\widehat{\zeta}_{n,k'} - \zeta_k') G_k^{(\alpha)}(x) G_{k'}^{(\alpha)}(x)\Big) w^{(\alpha)}(x)\, \mathrm{d}x = \sum_{k=0}^{K_n} \mathbb{E}(\widehat{\zeta}_{n,k} - \zeta_k)^2.$$

Also, noticing that $\int_{-1}^{1} R_n^2(x) w^{(\alpha)}(x)\, \mathrm{d}x = \sum_{k=K_n+1}^{\infty} \zeta_k^2$, we obtain

$$v_n = \sum_{k=0}^{K_n} \mathbb{E}(\widehat{\zeta}_{n,k} - \zeta_k)^2 + \sum_{k=K_n+1}^{\infty} \zeta_k^2, \tag{3.3}$$

an usual expression for the MISE, which involves a variance term and a bias term. Note that the bias term $\sum_{k=K_n+1}^{\infty} \zeta_k^2$ is exactly the one obtained in the standard case (see, e.g., Hall (1982)) and tends to zero as n tends to infinity.

Let us study the variance term

$$\begin{aligned}
\mathbb{E}(\widehat{\zeta}_{n,k} - \zeta_k)^2 &= \mathbb{E}\Big(\sum_{j=0}^{k} g_{k,j}^{(\alpha)}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2) - \sigma(j) + \sigma(j+2))\Big)^2 \\
&\leq (k+1)\sum_{j=0}^{k} g_{k,j}^{(\alpha)2}\mathbb{E}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2) - \sigma(j) + \sigma(j+2))^2 \\
&= (k+1)\sum_{j=0}^{k} g_{k,j}^{(\alpha)2}\Big[\big(\mathbb{E}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2) - \sigma(j) + \sigma(j+2))\big)^2 \\
&\quad +\mathrm{var}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2))\Big].
\end{aligned}$$

As

$$\begin{aligned}
\big(\mathbb{E}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2) - \sigma(j) + \sigma(j+2))\big)^2 &= \Big(-\frac{j}{n}\,\sigma(j) + \frac{j+2}{n}\,\sigma(j+2)\Big)^2 \\
&\leq \frac{Cj^2\sigma^2(0)}{n^2},
\end{aligned}$$

it follows that

$$\mathbb{E}(\widehat{\zeta}_{n,k} - \zeta_k)^2 \leq Ck\left(\max_{0\leq j\leq k} g_{k,j}^{(\alpha)}\right)^2 \sum_{j=0}^{k}\left(\frac{j^2}{n^2} + \mathrm{var}\left(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2)\right)\right).$$

In these inequalities and below, C refers to a generic constant which may vary from one formula to another.

Lemma 5 gives an upper bound for the maximum of Gegenbauer coefficients, leading to

$$\mathbb{E}(\widehat{\zeta}_{n,k} - \zeta_k)^2 \leq Ck^{12}e^{2k\beta}\sum_{j=0}^{k}\left(\frac{j^2}{n^2} + \mathrm{var}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2))\right) \quad \text{with } \beta = \log(1+\sqrt{2}). \tag{3.4}$$

In order to estimate $\mathrm{var}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2))$, we can apply Lemma 4 from Appendix A. Indeed, from (1.7), the spectral density $f(\lambda)$ satisfies

$$\begin{aligned}
2\pi\big|1 - e^{-2i\lambda}\big|^2 f(\lambda) = \big|1 - e^{-2i\lambda}\big|^2 \int_{-1}^{1} \frac{\varphi(x)}{|1 - xe^{i\lambda}|^2}\,\mathrm{d}x &= \frac{\lambda^2}{\sin\lambda}\int_{\frac{-1-\cos\lambda}{\sin\lambda}}^{\frac{1-\cos\lambda}{\sin\lambda}} \frac{\varphi(t\sin\lambda + \cos\lambda)}{1+t^2}\,\mathrm{d}t \\
\leq \frac{\lambda^2}{\sin\lambda}\int_{\frac{-1-\cos\lambda}{\sin\lambda}}^{\frac{1-\cos\lambda}{\sin\lambda}} \varphi(t\sin\lambda + \cos\lambda)\,\mathrm{d}t &= \frac{\lambda^2}{\sin^2\lambda}\int_{-1}^{1}\varphi(x)\,\mathrm{d}x \leq C
\end{aligned}$$

by (1.4) and because $\sup_{\lambda\in[-\pi,\pi]} \lambda^2 \sin^{-2}\lambda \leq C$. Hence, from Lemma 4,

$$\mathrm{var}(\widehat{\sigma}_n(j) - \widehat{\sigma}_n(j+2)) \leq \frac{C}{n}. \tag{3.5}$$

So that, by (3.4),

$$\mathbb{E}(\widehat{\zeta}_{n,k} - \zeta_k)^2 \leq Ce^{2k\beta}\frac{k^{15}}{n}. \tag{3.6}$$

Finally, applying (3.1), we get

$$\sum_{j=0}^{K_n}\mathbb{E}(\widehat{\zeta}_{n,j} - \zeta_j)^2 \leq C\sum_{j=0}^{K_n} e^{2j\beta}\frac{j^{15}}{n} \leq Ce^{2K_n\beta}\frac{K_n^{15}}{n} \to 0 \tag{3.7}$$

as $n \to \infty$. □

Next theorem gives a rate of L^2-convergence for mixture densities regular enough. For the sake of simplicity, the result is given in the semi-parametric situation (1.5). It could be formulated in a more general setting as well.

Theorem 2. *Suppose that the mixture density has the form*

$$\varphi(x) = (1-x)^{d_1}(1+x)^{d_2}\psi(x), \quad d_1 > 0, d_2 > 0,$$

where $\psi(x)$ has a continuous derivative on $[-1,1]$ and does not vanish at ± 1. For every $\alpha > -1$ such that

$$\min(d_1, d_2) > \frac{\alpha}{2} + \frac{3}{4}$$

and $\widehat{\varphi}_n$ being defined in (2.4)–(2.5), with K_n satisfying (3.1), we have

$$\int_{-1}^{1} \frac{\mathbb{E}(\widehat{\varphi}_n(x) - \varphi(x))^2}{(1-x^2)^\alpha}\, \mathrm{d}x = O\left(\frac{1}{(\log n)^3}\right).$$

Proof. The proof is based on Proposition 6, whose conditions, as easily seen, are satisfied by the function

$$\zeta(x) = \frac{\varphi(x)}{(1-x^2)^\alpha}.$$

Hence, $\zeta_n = O(n^{-1})$, and the bias term in (3.3) is estimated by

$$\sum_{k=K_n+1}^{\infty} \zeta_k^2 = O\left(\frac{1}{K_n^3}\right) = O\left(\frac{1}{(\log n)^3}\right).$$

It remains to apply (3.7) which gives

$$\sum_{k=0}^{K_n} \mathbb{E}(\widehat{\zeta}_{n,k} - \zeta_k)^2 \leq C \frac{(\log n)^{15}}{n^{1-2\gamma \log(1+\sqrt{2})}}.$$

This completes the proof. □

3.2 Uniform convergence on compact sets

Next we study the uniform convergence of the mixture density estimator given by the semi-parametric form (1.5).

Theorem 3. *Let φ be a mixture density of the form*

$$\varphi(x) = (1-x)^{d_1}(1+x)^{d_2}\psi(x), \quad d_1 > 0, d_2 > 0,$$

where ψ is analytic on the open unit disk, continuous in $[-1,1]$, and does not vanish at ± 1. Denote $d = \min(d_1, d_2)$. Let us consider $\widehat{\varphi}_n(x)$, the estimator defined in (2.4) with K_n satisfying (3.1).

a) If $0 \leq \alpha < d + 1/2$, then for any $\delta \in [0, 1[$ we have

$$\mathbb{E}\left(\sup_{x\in[-\delta,\delta]} (\widehat{\varphi}_n(x) - \varphi(x))^2\right) \leq \frac{C(\delta)}{(\log n)^{2d-2\alpha+1}}. \tag{3.8}$$

b) If $5/2 \leq \alpha < d + 1/2$, the result holds with $\delta = 1$.

Proof. For all $0 \le \delta \le 1$ we have

$$\begin{aligned}
\sup_{x\in[-\delta,\delta]} (\widehat{\varphi}_n(x)-\varphi(x))^2 &= \sup_{x\in[-\delta,\delta]} (1-x^2)^{2\alpha}\left(\sum_{k=0}^{K_n}(\widehat{\zeta}_{n,k}-\zeta_k)G_k^{(\alpha)}(x) - \sum_{k=K_n+1}^{\infty}\zeta_k G_k^{(\alpha)}(x)\right)^2 \\
&\le 2\left(\sum_{k=0}^{K_n}|\widehat{\zeta}_{n,k}-\zeta_k|s_k\right)^2 + 2\sup_{x\in[-\delta,\delta]}(1-x^2)^{2\alpha}\left(\sum_{k=K_n+1}^{\infty}\zeta_k G_k^{(\alpha)}(x)\right)^2 \\
&=: 2J_{n,1}+2J_{n,2},
\end{aligned}$$

where $s_k := \sup_{x\in[-\delta,\delta]}(1-x^2)^{\alpha}|G_k^{(\alpha)}(x)|$.

By inequality (7.33.6) of Szegö (1967) adapted to the normalized Gegenbauer polynomials, $G_k^{(\alpha)}(x)$ satisfies

$$|G_k^{(\alpha)}(x)| \le \frac{C}{(\arccos x)^{\alpha+1/2}} \le \frac{C}{(1-x^2)^{\alpha/2+1/4}}, \quad -1<x<1. \tag{3.9}$$

Hence, $s_k \le C\sup_{x\in[-\delta,\delta]}(1-x^2)^{\alpha/2-1/4} \le C(\delta) < \infty$ in both cases a) and b). Combining this with (3.6), implies

$$\begin{aligned}
\mathbb{E}J_{n,1} &\le (K_n+1)\sum_{k=0}^{K_n}s_k^2\mathbb{E}(\widehat{\zeta}_{n,k}-\zeta_k)^2 \\
&\le C(\delta)K_n\sum_{k=0}^{K_n}e^{2k\beta}\frac{k^{15}}{n} \\
&\le C(\delta)(\log n)^{15}n^{2\gamma\beta-1}. \qquad (3.10)
\end{aligned}$$

Next we show

$$J_{n,2} \le C(\delta)(\log n)^{2\alpha-2d-1}. \tag{3.11}$$

The function $\zeta(x) = (1-x)^{d_1-\alpha}(1+x)^{d_2-\alpha}\psi(x)$ has the expansion

$$\zeta(x) = \sum_{k=0}^{\infty} z_k x^k \tag{3.12}$$

with $z_k = c_1k^{\alpha-d_1-1}+c_2(-1)^kk^{\alpha-d_2-1}+o(1) = O(k^{\alpha-d-1})$, where c_1 and c_2 are nonzero constants (see Leipus and Viano (2000) for the behavior of z_k). Therefore, since $\alpha < d+1/2$, we have $\sum_{k=0}^{\infty} z_k^2 < \infty$, i.e. the function $\zeta(x)$ belongs to the Hardy space H^2. By Proposition 7,

$$\begin{aligned}
\Big(\sum_{k=K_n+1}^{\infty}\zeta_k G_k^{(\alpha)}(x)\Big)^2 &\le C(1-|x|)^{-\alpha-5/2}\sum_{k=K_n+1}^{\infty}z_k^2 \\
&\le C(1-|x|)^{-\alpha-5/2}K_n^{2\alpha-2d-1} \qquad (3.13)
\end{aligned}$$

for all $x \in (-1,1)$. Hence,

$$J_{n,2} \le C\sup_{x\in[-\delta,\delta]}(1-|x|)^{\alpha-5/2}K_n^{2\alpha-2d-1},$$

where the supremum is finite even for $\delta = 1$ by assumption on α. Finally, (3.10) and (3.11) yield (3.8). □

4 Simulation study

In order to illustrate the behavior of estimates defined in (2.4) we present a short simulation study, discussing, in particular, the choice of number of terms K_n and the choice of the parameter α.

The following two families of densities are considered:

- a Beta-type distribution, denoted by $\mathcal{B}e(d_1, d_2)$,

$$\varphi(x) \propto (1-x)^{d_1}(1+x)^{d_2}\mathbf{1}_{[-1,1]}(x), \quad d_i > 0,\ i = 1, 2.$$

- a bimodal density $\mathcal{L}(d_1, d_2)$

$$\varphi(x) \propto x^2(1-x)^{d_1}(1+x)^{d_2}\mathbf{1}_{[-1,1]}(x), \quad d_i > 0,\ i = 1, 2. \tag{4.1}$$

In both cases, the density behaves as $(1-x)^{d_1}$ (respectively, $(1+x)^{d_2}$) when x goes to 1 (respectively, -1). By taking $d_1 = 0.25$ and $d_2 = 0.75$ (respectively, $d_1 = 1$ and $d_2 = 1.5$), we illustrate the behaviour of the estimate in a long (respectively, short) memory context. The simulated paths are obtained by aggregating $N = 5000$ independent AR(1) processes.

By construction, the estimator $\widehat{\varphi}_n$ is not necessarily positive and its integral is not equal to 1. We replace $\widehat{\varphi}_n$ by its positive part $\widehat{\varphi}_n^+$. Moreover, we renormalize $\widehat{\varphi}_n^+$ by its Riemann sum

$$\frac{1}{M}\sum_{i=-M}^{M}\widehat{\varphi}_n^+(i/M) \qquad \text{with} \quad M = 10^3.$$

We compare the estimated MISE for different values of K_n and α. The estimation of the MISE is obtained by a Monte Carlo procedure based on $N = 100$ independent replications. In all the figures, the dashed line corresponds to the true mixture density.

In Figures 1 and 2, we consider Beta distributions leading to short and long memory aggregated processes, respectively. The sample size is 5000. Both graphs on the left show that the MISE is optimal for $K_n = 2$ (note that the boundary value $[(2\log(1+\sqrt{2}))^{-1}\log n]$ in Theorem 1 equals 4) and that, for $K_n = 2$, the MISE is rather stable when α belongs to (0,1). Fixing $\alpha = 1/2$, graphs on the right represent the estimated densities. Clearly for $K_n = 2$ the behavior of the estimate is satisfying and it becomes really inefficient when $K_n \geq 4$.

Figures 3 and 4 concern the the bimodal density defined in (4.1). Apparently, two terms ($K_n = 2$) are not sufficient for a good estimation of the density. Increasing the number of observations up to $n = 10000$, we obtain an efficient estimator using $K_n = 4$.

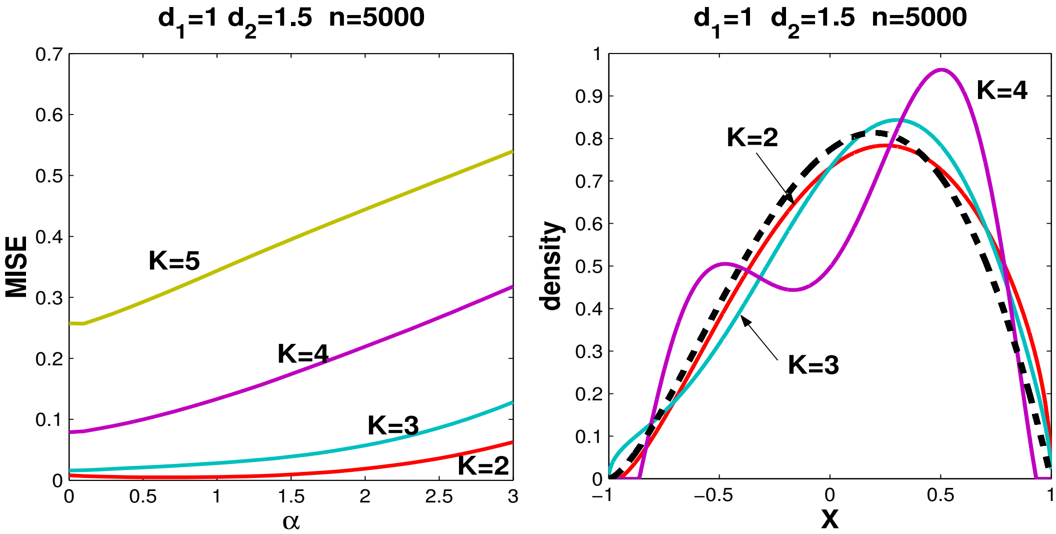

Figure 1: Representation of the estimated MISE versus α (left). Estimation of the mixture density in the case $\alpha = 1/2$ (right). The mixture distribution is $\mathcal{B}e(1, 1.5)$ and the sample size $n = 5000$.

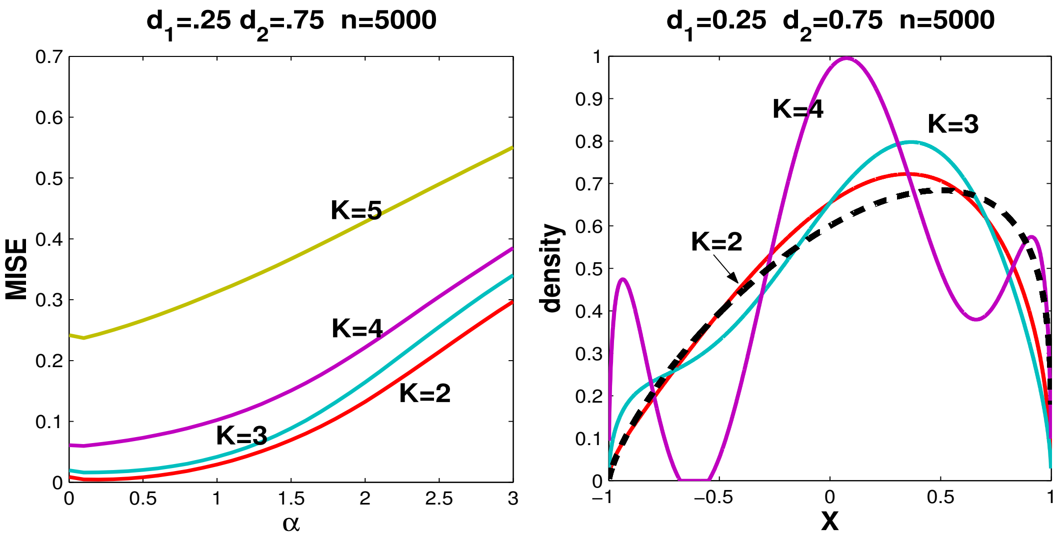

Figure 2: Representation of the estimated MISE versus α (left). Estimation of the mixture density in the case $\alpha = 1/2$ (right). The mixture distribution is $\mathcal{B}e(0.25, 0.75)$ and the sample size $= n = 5000$.

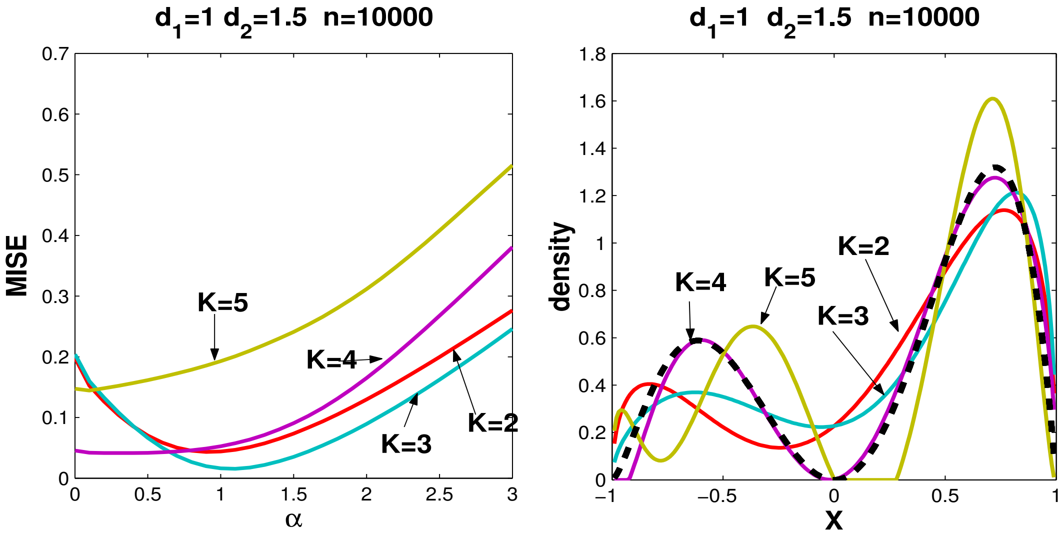

Figure 3: Representation of the estimated MISE versus α (left). Estimation of the mixture density in the case $\alpha = 1/2$ (right). The mixture distribution is $\mathcal{L}(1, 1.5)$ and the sample size $n = 10000$.

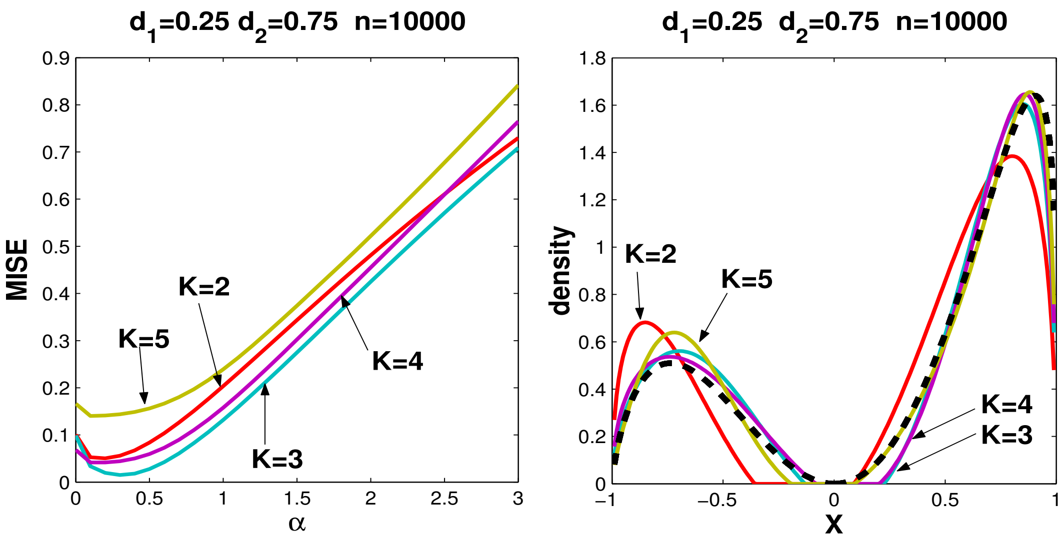

Figure 4: Representation of the estimated MISE versus α (left). Estimation of the mixture density in the case $\alpha = 1/2$ (right). The mixture distribution is $\mathcal{L}(0.25, 0.75)$ and the sample size $n = 10000$.

A Proof of inequality (3.5)

Lemma 4. *Let X_t be a stationary Gaussian process with spectral density f such that the function $\left|1-e^{-2i\lambda}\right|^2 f(\lambda)$ is bounded. Then, for $\widehat{\sigma}_n(k) = n^{-1}\sum_{i=1}^{n-k} X_i X_{i+k}$,*

$$\operatorname{var}(\widehat{\sigma}_n(k) - \widehat{\sigma}_n(k+2)) \le \frac{C}{n}, \tag{A.1}$$

where the constant C does not depend on k.

Proof. Denoting $D_k = X_k - X_{k+2}$, we have

$$\operatorname{var}\Big(\widehat{\sigma}_n(k) - \widehat{\sigma}_n(k+2)\Big) = \frac{1}{n^2}\operatorname{var}\Bigg(\sum_{j=1}^{n-k} X_j X_{j+k} - \sum_{j=1}^{n-k-2} X_j X_{j+k+2}\Bigg)$$

$$= \frac{1}{n^2}\operatorname{var}\Bigg(\sum_{j=1}^{n-k-2} X_j D_{j+k} + X_{n-k-1}X_{n-1} + X_{n-k}X_n\Bigg)$$

$$\le \frac{C}{n^2}\Bigg(\operatorname{var}\Bigg(\sum_{j=1}^{n-k-2} X_j D_{j+k}\Bigg) + \operatorname{var}(X_{n-k-1}X_{n-1}) + \operatorname{var}(X_{n-k}X_n)\Bigg)$$

$$\le \frac{C}{n^2}\Bigg(\operatorname{var}\Bigg(\sum_{j=1}^{n-k-2} X_j D_{j+k}\Bigg) + 1\Bigg).$$

As the process X_t is Gaussian, it follows

$$W_{n,k} := \frac{1}{n^2}\operatorname{var}\Bigg(\sum_{j=1}^{n-k-2} X_j D_{j+k}\Bigg) = \frac{1}{n^2}\sum_{j,\ell=1}^{n-k-2} \operatorname{cov}(X_j D_{j+k}, X_\ell D_{\ell+k})$$

$$= \frac{1}{n^2}\Bigg(\sum_{j,\ell=1}^{n-k-2} \mathbb{E}(X_j X_\ell)\mathbb{E}(D_{j+k}D_{\ell+k}) + \sum_{j,\ell=1}^{n-k-2} \mathbb{E}(X_j D_{k+\ell})\mathbb{E}(X_\ell D_{j+k})\Bigg).$$

Setting $\sigma(h) = \operatorname{cov}(X_h, X_0)$, the expectations can be rewritten as follows

$$\begin{aligned}
\mathbb{E}(D_{j+k}D_{\ell+k}) &= 2\sigma(j-\ell) - \sigma(j-\ell-2) - \sigma(j-\ell+2),\\
\mathbb{E}(X_j D_{\ell+k}) &= \sigma(j-\ell-k) - \sigma(j-\ell-k-2),\\
\mathbb{E}(X_\ell D_{j+k}) &= \sigma(\ell-j-k) - \sigma(\ell-j-k-2).
\end{aligned}$$

Then, using the spectral representation,

$$\begin{aligned}
\operatorname{cov}(X_j D_{j+k}, X_\ell D_{\ell+k}) &= \int e^{-i(j-\ell)\lambda} f(\lambda)\,\mathrm{d}\lambda \int e^{-i(j-\ell)\nu}(2 - e^{-2i\nu} - e^{2i\nu}) f(\nu)\,\mathrm{d}\nu\\
&\quad + \int e^{-i(j-\ell-k)\lambda}(1-e^{2i\lambda}) f(\lambda)\,\mathrm{d}\lambda \int e^{-i(\ell-j-k)\nu}(1-e^{2i\nu}) f(\nu)\,\mathrm{d}\nu\\
&= \int\int e^{-i(j-\ell)(\lambda-\nu)}\left|1-e^{2i\lambda}\right|^2 f(\lambda) f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu\\
&\quad + \int\int e^{-i(j-\ell)(\lambda-\nu)} e^{ik(\lambda+\nu)}(1-e^{2i\lambda})(1-e^{2i\nu}) f(\lambda) f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu,
\end{aligned}$$

where the integrals are all on $[-\pi,\pi]$. Therefore,

$$\begin{aligned}
W_{n,k} &= \frac{1}{n^2}\int\int\left|\sum_{j=1}^{n-k-2} e^{-ij(\lambda-\nu)}\right|^2 \left|1-e^{2i\lambda}\right|^2 f(\lambda)f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu \\
&\quad+\frac{1}{n^2}\int\int\left|\sum_{j=1}^{n-k-2} e^{-ij(\lambda-\nu)}\right|^2 e^{ik(\lambda+\nu)}(1-e^{2i\lambda})(1-e^{2i\nu})f(\lambda)f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu \\
&= \frac{1}{n}\int\int F_n(\lambda-\nu)\left|1-e^{2i\lambda}\right|^2 f(\lambda)f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu \\
&\quad+\frac{1}{n}\int\int F_n(\lambda-\nu)e^{ik(\lambda+\nu)}(1-e^{2i\lambda})(1-e^{2i\nu})f(\lambda)f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu,
\end{aligned}$$

where

$$F_n(x)=\frac{1}{n}\left|\sum_{j=1}^{n} e^{-ijx}\right|^2$$

is proportional to the Fejer kernel, and satisfies $\int_{-\pi}^{\pi} F_n(x)\,\mathrm{d}x = 2\pi$. Write

$$nW_{n,k} = \int f(\nu)\int F_n(\lambda-\nu)\left|1-e^{2i\lambda}\right|^2 f(\lambda)\,\mathrm{d}\lambda\,\mathrm{d}\nu \tag{A.2}$$

$$+\int\int F_n(\lambda-\nu)e^{ik(\lambda+\nu)}(1-e^{2i\lambda})(1-e^{2i\nu})f(\lambda)f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu. \tag{A.3}$$

To prove that $W_{n,k}\le C/n$, it suffices to prove that both integrals (A.2) and (A.3) are finite.

Concerning the integral (A.2), we have

$$\int F_n(\nu-\lambda)\left|1-e^{2i\lambda}\right|^2 f(\lambda)\,\mathrm{d}\lambda \le 2\pi C$$

since the function $\left|1-e^{2i\lambda}\right|^2 f(\lambda)$ is bounded. Hence,

$$\int f(\nu)\int F_n(\nu-\lambda)\left|1-e^{2i\lambda}\right|^2 f(\lambda)\,\mathrm{d}\lambda\,\mathrm{d}\nu \le 2\pi C||f||_1.$$

To verify the finiteness of the integral (A.3), note that the function $\tilde f(\lambda)=\left|1-e^{-2i\lambda}\right|f(\lambda)$ is in L^2 since f is integrable and $\left|1-e^{2i\lambda}\right|^2 f(\lambda)$ is bounded. Thus,

$$\begin{aligned}
&\left|\int\int F_n(\lambda-\nu)e^{ik(\nu+\lambda)}(1-e^{2i\lambda})(1-e^{2i\nu})f(\lambda)f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu\right| \\
&\le \int\int F_n(\lambda-\nu)\tilde f(\lambda)\tilde f(\nu)\,\mathrm{d}\lambda\,\mathrm{d}\nu \\
&= \int F_n(u)\int \tilde f(\nu+u)\tilde f(\nu)\,\mathrm{d}u\,\mathrm{d}\nu \le 2\pi||\tilde f||_2^2.
\end{aligned}$$

This completes the proof.

□

B The Gegenbauer coefficients and Gegenbauer expansion

Here we establish some useful properties of the Gegenbauer polynomials $G_n^{(\alpha)}(x)$.

Note, that for the sake of shortness, all the lemmas and propositions below are proved in the case $\alpha\neq -1/2$. Indeed, the function $\Gamma(\alpha+1/2)$ appearing in the proofs is not defined for this value. However, obvious modifications show that our results hold also for $\alpha=-1/2$.

B.1 The maximum of Gegenbauer coefficients

Lemma 5. *Consider the coefficients $g_{n,j}^{(\alpha)}$ of the Gegenbauer polynomial $G_n^{(\alpha)}$ as defined in* (2.1). *For any n,*

$$\pi_n := \max_{0\leq j\leq n} |g_{n,j}^{(\alpha)}| \leq C n^{11/2} e^{n\beta} \quad \textit{with} \quad \beta = \log(1+\sqrt{2}).$$

Proof. Hereafter we adopt the notation $C_n^{(\alpha)}(x)$ of Abramovitz and Stegun (1965) (see formula 22.2.3) for the nonnormalized Gegenbauer polynomial corresponding to the weight function $(1-x^2)^{\alpha-1/2}$. The normalized polynomial $G_n^{(\alpha)}(x)$ is linked to $C_n^{(\alpha+1/2)}(x)$ by equality

$$G_n^{(\alpha)}(x) = g_n^{-1/2} C_n^{(\alpha+1/2)}(x), \tag{B.1}$$

where

$$g_n = \frac{\pi}{2^{2\alpha}} \frac{\Gamma(n+2\alpha+1)}{\Gamma^2(\alpha+1/2)\Gamma(n+1)(n+\alpha+1/2)}. \tag{B.2}$$

Then (see, e.g., formula 22.3.4 in Abramovitz and Stegun (1965)),

$$G_n^{(\alpha)}(x) = \frac{{g_n}^{-1/2}}{\Gamma(\alpha+1/2)} \sum_{m=0}^{\lfloor n/2\rfloor} (-1)^m \frac{\Gamma(\alpha+1/2+n-m)}{\Gamma(m+1)\Gamma(n-2m+1)} \, (2x)^{n-2m}. \tag{B.3}$$

Consequently, the coefficients of the Gegenbauer polynomial $G_n^{(\alpha)}$ are equal to

$$g_{n,n-2m}^{(\alpha)} = (-1)^m \frac{(g_n)^{-1/2}}{\Gamma(\alpha+1/2)} \frac{2^{n-2m}\Gamma(\alpha+1/2+n-m)}{\Gamma(m+1)\Gamma(n-2m+1)} \qquad \text{for} \quad 0 \leq m \leq [n/2]\,. \tag{B.4}$$

Hence,

$$\left| \frac{g_{n,n-2(m+1)}^{(\alpha)}}{g_{n,n-2m}^{(\alpha)}} \right| = \frac{1}{4} \frac{(n-2m)(n-2m-1)}{(\alpha-1/2+n-m)(m+1)} =: R(m). \tag{B.5}$$

Let the function $R(z)$ be defined by (B.5). The roots z_-, z_+ of the equation $|R(z)| = 1$, are obtained by solving

$$8z^2 - 2z(4n-4+2\alpha) + n^2 - 5n + 2 - 4\alpha = 0,$$

and equal

$$z_\pm = \frac{n-1}{2} + \frac{\alpha}{4} \pm \frac{\sqrt{2}}{4}\sqrt{\Delta_n} \quad \text{with} \quad \Delta_n = n^2 + n(1+2\alpha) + (\alpha^2+2\alpha)/2.$$

A straightforward verification shows that for sufficiently large n and all α the following bounds are true:

$$\frac{n}{2}\left(1-\frac{\sqrt{2}}{2}\right) - \frac{\sqrt{2}-1}{4}\,\alpha - 1 \leq z_- \leq \frac{n}{2}\left(1-\frac{\sqrt{2}}{2}\right) - \frac{\sqrt{2}-1}{4}\,\alpha =: z^*. \tag{B.6}$$

Since z_- is the only root satisfying $0 \leq z_- \leq [n/2]$ and it holds

$$|R(z)| \geq 1 \quad \text{for} \quad z \leq z_-; \quad |R(z)| \leq 1 \quad \text{for} \quad z_- \leq z \leq [n/2], \tag{B.7}$$

(B.6)–(B.7) imply that

$$\pi_n = \max_{0\leq m\leq [n/2]} |g_{n,n-2m}^{(\alpha)}| = \max(|g_{n,n-2m^*}^{(\alpha)}|, |g_{n,n-2(m^*+1)}^{(\alpha)}|),$$

where m^* is the integer number such that $m^* \leq z_- \leq m^* + 1$. It is clear that m^* and $m^* + 1$ are in $[z^* - 2, z^* + 1]$ and hence

$$\begin{aligned}
\pi_n &\leq \frac{g_n^{-1/2}}{\Gamma(\alpha+1/2)} \frac{2^{n-2z^*+4}\Gamma(\alpha+1/2+n-z^*+2)}{\Gamma(z^*-1)\Gamma(n-2z^*-1)} \\
&\leq Cg_n^{-1/2} \frac{2^{n\sqrt{2}/2}\Gamma\big(n(2+\sqrt{2})/4)+\alpha+1/2+\alpha(\sqrt{2}-1)/4+2\big)}{\Gamma\big(n(2-\sqrt{2})/4-\alpha(\sqrt{2}-1)/4-1\big)\Gamma\big(n\sqrt{2}/2+\alpha(\sqrt{2}-1)/2-1\big)} \\
&=: Cg_n^{-1/2}2^{n\sqrt{2}/2} \frac{\Gamma(a_1 n+b_1)}{\Gamma(a_2 n+b_2)\Gamma(a_3 n+b_3)} \qquad \text{(B.8)}
\end{aligned}$$

with $a_1 - a_2 - a_3 = 0$. The relation

$$\Gamma(an+b) \sim \sqrt{2\pi}e^{-an}(an)^{an+b-1/2} \qquad (a > 0)$$

then implies

$$\frac{\Gamma(a_1 n+b_1)}{\Gamma(a_2 n+b_2)\Gamma(a_3 n+b_3)} \sim \frac{1}{\sqrt{2\pi}} \frac{a_1^{b_1-1/2}}{a_2^{b_2-1/2}a_3^{b_3-1/2}} \left(\frac{a_1^{a_1}}{a_2^{a_2}a_3^{a_3}}\right)^n n^{b_1-b_2-b_3+1/2}. \qquad \text{(B.9)}$$

Since

$$\frac{a_1^{a_1}}{a_2^{a_2}a_3^{a_3}} = (1+\sqrt{2})2^{-\sqrt{2}/2} \quad \text{and} \quad b_1 - b_2 - b_3 + 1/2 = \alpha + 5,$$

(B.8)–(B.9), together with relation

$$g_n \sim \frac{\pi}{2^{2\alpha}\Gamma(\alpha+1/2)}\, n^{2\alpha-1}, \qquad \text{(B.10)}$$

imply that, for n large enough,

$$\pi_n \leq Cn^{\alpha+5}(1+\sqrt{2})^n.$$

This concludes the proof. □

B.2 Estimation of the coefficients in a Gegenbauer expansion

Next proposition establishes the bound for coefficient in the Gegenbauer expansions of a function having one derivative.

Proposition 6. *Let h be a function in $L^2(w^{(\alpha)})$, $\alpha > -1$, such that $(1-x^2)^\alpha h(x)$ is of bounded variation on $(-1, 1[$. Assume that h is differentiable on $(-1, 1)$, such that $(1-x^2)^{\alpha/2+1/4}h'(x)$ is of bounded variation on (-1,1) and*

$$\int_{-1}^{1} (1-x^2)^{\alpha/2-3/4}\, |h'(x)| \;\mathrm{d}x < \infty. \qquad \text{(B.11)}$$

Then

$$h_n = \int_{-1}^{1} h(x)G_n^{(\alpha)}(x)w^{(\alpha)}(x)\;\mathrm{d}x = O\left(\frac{1}{n^2}\right). \qquad \text{(B.12)}$$

Proof. The proof follows the same arguments as the proof of Theorem 1 in Hall (1982), where the rate of L^2-convergence of the Legendre expansion of a function depending on the number of existing derivatives was established. The difference lies in the need to extend some properties of Legendre polynomials used in Hall (1982) to the Gegenbauer polynomials.

Let $C_n^{(\alpha+1/2)}(x)$ be the nonnormalized Gegenbauer polynomial (see (B.1)–(B.2)). From Rodrigues' formula

$$(1-x^2)^\alpha C_n^{(\alpha+1/2)}(x) = A_n \frac{d^n}{\mathrm{d}x^n}\left((1-x^2)^{\alpha+n}\right) \quad \text{with } A_n = (-1)^n \frac{\Gamma(\alpha+1)\Gamma(n+2\alpha+1)}{2^n n!\Gamma(2\alpha+1)\Gamma(\alpha+n+1)}$$

it is easy to deduce that

$$(2n+2\alpha+1)C_n^{(\alpha+1/2)}(x)w^{(\alpha)}(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left[U_n^{(\alpha)}(x)w^{(\alpha)}(x)\right], \tag{B.13}$$

where

$$U_n^{(\alpha)}(x) = \frac{n+1}{n+2\alpha+1}\, C_{n+1}^{(\alpha+1/2)}(x) - \frac{n+2\alpha}{n}\, C_{n-1}^{(\alpha+1/2)}(x). \tag{B.14}$$

Equality (B.13) implies

$$(2n+2\alpha+1)\int_{-1}^{1} h(x)C_n^{(\alpha+1/2)}(x)w^{(\alpha)}(x)\ \mathrm{d}x = \int_{-1}^{1} h(x)\frac{\mathrm{d}}{\mathrm{d}x}\left[U_n^{(\alpha)}(x)w^{(\alpha)}(x)\right]\ \mathrm{d}x.$$

As $C_n^{(\alpha+1/2)}(1) = (-1)^n C_n^{(\alpha+1/2)}(-1) = \binom{n+2\alpha}{n}$, it follows that $U_n^{(\alpha)}(1) = U_n^{(\alpha)}(-1) = 0$. By assumption of the proposition, the function $h(x)w^{(\alpha)}(x)$ is bounded on [-1,1], so that $h(x)U_n^{(\alpha)}(x)w^{(\alpha)}(x)$ vanishes at ± 1. Hence, integrating by parts we see that

$$\int_{-1}^{1} h(x)C_n^{(\alpha+1/2)}(x)w^{(\alpha)}(x)\ \mathrm{d}x = \frac{1}{2n+2\alpha+1}\int_{-1}^{1} h'(x)U_n^{(\alpha)}(x)w^{(\alpha)}(x)\ \mathrm{d}x. \tag{B.15}$$

Now, from Theorem 8.21.11 in Szegö (1967), we obtain for $\theta \in (0,\pi)$

$$C_n^{(\alpha+1/2)}(\cos\theta) = C_1(\alpha)\frac{\Gamma(n+2\alpha+1)}{\Gamma(n+\alpha+3/2)}\frac{\cos((n+\alpha+1/2)\theta - \pi(\alpha+1/2)/2)}{(\sin\theta)^{\alpha+1/2}} + R(n,\theta), \tag{B.16}$$

where

$$|R(n,\theta)| \le C_2(\alpha)\frac{\Gamma(n+2\alpha+1)}{\Gamma(n+\alpha+5/2)}\frac{1}{(\sin\theta)^{\alpha+3/2}},$$

$C_1(\alpha)$ and $C_2(\alpha)$ being positive constants which depend only on α. This allows to estimate the integral in the right-hand side of (B.15). Indeed, since $h'(\cos\theta)|\sin\theta|^{1/2-\alpha}$ is of bounded variation on $[0,\pi]$, Riemann-Lebesgue's lemma (see, e.g., Whittaker and Watson (1927, Chapter 9.41)) leads to

$$\int_0^\pi h'(\cos\theta)\cos((n+\alpha+1/2)\theta - \pi(\alpha+1/2)/2)(\sin\theta)^{\alpha+1/2}\ \mathrm{d}\theta = O(n^{-1}). \tag{B.17}$$

Since $\frac{\Gamma(n+2\alpha+1)}{\Gamma(n+\alpha+3/2)} = O(n^{\alpha-1/2})$, it follows from (B.16)–(B.17) that

$$\begin{aligned}\int_{-1}^{1} h'(x)C_n^{(\alpha+1/2)}(x)w^{(\alpha)}(x)\ \mathrm{d}x &= \int_0^\pi h'(\cos\theta)C_n^{(\alpha+1/2)}(\cos\theta)(\sin\theta)^{2\alpha+1}\ \mathrm{d}\theta \\ &= O(n^{\alpha-3/2}) + \int_0^\pi h'(\cos\theta)R(n,\theta)(\sin\theta)^{2\alpha+1}\ \mathrm{d}\theta.\end{aligned}$$

Then, $\frac{\Gamma(n+2\alpha+1)}{\Gamma(n+\alpha+5/2)} = O(n^{\alpha-3/2})$ and, from (B.11), $\int_0^\pi |h'(\cos\theta)|(\sin\theta)^{\alpha-1/2}\ \mathrm{d}\theta < \infty$. Consequently,

$$\left|\int_0^\pi h'(\cos\theta)R(n,\theta)(\sin\theta)^{2\alpha+1}\ \mathrm{d}\theta\right| \le C_3(\alpha)n^{\alpha-3/2}\int_0^\pi |h'(\cos\theta)|(\sin\theta)^{\alpha-1/2}\ \mathrm{d}\theta = O(n^{\alpha-3/2}).$$

The same estimation holds if $C_n^{(\alpha+1/2)}(x)$ is replaced by $C_{n+1}^{(\alpha+1/2)}(x)$ or by $C_{n-1}^{(\alpha+1/2)}(x)$ so that using (B.14) and (B.15) leads to

$$\int_{-1}^{1} h(x)C_n^{(\alpha+1/2)}(x)w^{(\alpha)}(x)\ \mathrm{d}x = O(n^{\alpha-5/2}).$$

Finally, to obtain (B.12), use (B.1) and (B.10). □

B.3 Uniform convergence of the Gegenbauer expansion

To our knowledge, there are only few studies about uniform convergence in the expansion of functions on a basis of orthogonal polynomials. Sansone (1959, p. 202, Picone's theorem) considered Legendre expansions (i.e., the case Gegenbauer expansion with $\alpha = 0$), proving the uniform convergence for functions having bounded variation on $[-1, 1]$.

In the proposition below, we establish that, for functions regular enough, the remainder of their Gegenbauer expansion converges uniformly to zero on every compact subset of $(-1, 1)$.

Let $h(z) = \sum_{n=0}^{\infty} a_n z^n$ be a function from the Hardy space H^2. For all $\alpha \geq 0$, the function h belongs to $L^2(w^{(\alpha)})$:

$$\begin{aligned}
\int_{-1}^{1} h^2(x) w^{(\alpha)}(x)\,\mathrm{d}x &\leq \int_{-1}^{1} h^2(x)\,\mathrm{d}x \\
&= \int_{-1}^{1} \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} a_k a_j x^{j+k}\,\mathrm{d}x \\
&\leq 2\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} |a_k||a_j| \int_{0}^{1} x^{j+k}\,\mathrm{d}x \\
&= 2\sum_{k=0}^{\infty}\sum_{j=0}^{\infty} \frac{|a_k||a_j|}{k+j+1} \leq 2\pi \sum_{k=1}^{\infty} a_k^2, \qquad \text{(B.18)}
\end{aligned}$$

where (B.18) is Hilbert's inequality (see Duren (1970, p. 48)). Hence, for $\alpha \geq 0$, the function h admits an α-Gegenbauer expansion whose partial sums we denote

$$S_n(x) = \sum_{j=1}^{n} h_j G_j^{(\alpha)}(x). \qquad \text{(B.19)}$$

Proposition 7. *Let $h(z) = \sum_{n=0}^{\infty} a_n z^n$ be a function from the Hardy space H^2. For all $x \in (-1, 1)$, it holds*

$$|h(x) - S_n(x)| \leq C(1-|x|)^{-\alpha/2-5/4} \left(\sum_{j=n+1}^{\infty} a_j^2 \right)^{1/2} \qquad (\alpha \geq 0). \qquad \text{(B.20)}$$

Remark 1. Note that the assumption that h is from H^2 in the above proposition is neither necessary nor sufficient for h to be of bounded variation on [-1,1]. Boundedness of variation together with $h \in H^2$, however, are satisfied by the function $h(x) = (1-x^2)^d \quad (d > 0)$. In this case, the rate of convergence of the Legendre remainder given by Picone's theorem is $n^{-1/2}$, while Proposition 7 uses $a_j \sim j^{-d-1}$ and hence gives $|h(x) - S_n(x)| \leq C(x) n^{-d-1/2}$, which is a more precise rate.

Proof of Proposition 7. Denote the $L^2(w^{(\alpha)})$-norm of the remainder term in the Gegenbauer expansion by

$$\delta_n(h) = \left(\sum_{j=n+1}^{\infty} h_j^2 \right)^{1/2}. \qquad \text{(B.21)}$$

To prove (B.20) we will show that for any $x \in (-1, 1)$ we have

$$|h(x) - S_n(x)| \leq \frac{C}{(1-x^2)^{\alpha/2+1/4}}\, \delta_{n-1}(h_x), \qquad \text{(B.22)}$$

where $h_x(t) = \frac{h(x)-h(t)}{x-t}$ and

$$\delta_{n-1}(h_x) \leq \frac{\sqrt{2\pi}}{1-|x|} \left(\sum_{j=n+1}^{\infty} a_j^2 \right)^{1/2}. \qquad \text{(B.23)}$$

First consider (B.22). The difference $h(x) - S_n(x)$ can be rewritten as

$$h(x) - S_n(x) = \int_{-1}^{1} (h(x) - h(y)) \sum_{j=0}^{n} G_j^{(\alpha)}(x) G_j^{(\alpha)}(y)\, w^{(\alpha)}(y)\, \mathrm{d}y. \tag{B.24}$$

By Christoffel-Darboux's formula (see Szegö (1967)),

$$\sum_{j=0}^{n} G_j^{(\alpha)}(x) G_j^{(\alpha)}(y) = \frac{g_{n,n}^{(\alpha)}}{g_{n+1,n+1}^{(\alpha)}} \frac{G_{n+1}^{(\alpha)}(x) G_n^{(\alpha)}(y) - G_n^{(\alpha)}(x) G_{n+1}^{(\alpha)}(y)}{x - y}, \tag{B.25}$$

where the $g_{n,n}^{(\alpha)}$ are given by (2.1). From (B.4), taking $m = 0$, we obtain

$$\frac{g_{n,n}^{(\alpha)}}{g_{n+1,n+1}^{(\alpha)}} = \frac{1}{2} \sqrt{\frac{(n+1)(n+2\alpha+1)}{(n+\alpha+1/2)(n+\alpha+3/2)}},$$

which is a bounded sequence. Therefore, combining (B.24) and (B.25), we obtain

$$\begin{aligned}
|h(x) - S_n^{(\alpha}(x)| &\leq C \left| \int_{-1}^{1} h_x(y) (G_{n+1}^{(\alpha)}(x) G_n^{(\alpha)}(y) - G_n^{(\alpha)}(x) G_{n+1}^{(\alpha)}(y)) w^{(\alpha)}(y)\, \mathrm{d}y \right| \\
&\leq C |G_{n+1}^{(\alpha)}(x) \int_{-1}^{1} h_x(y) G_n^{(\alpha)}(y) w^{(\alpha)}(y)\, \mathrm{d}y| \\
&\quad + C |G_n^{(\alpha)}(x) \int_{-1}^{1} h_x(y) G_{n+1}^{(\alpha)}(y) w^{(\alpha)}(y)\, \mathrm{d}y| \\
&= C \left(|G_{n+1}^{(\alpha)}(x) h_{x,n}| + |G_n^{(\alpha)}(x) h_{x,n+1}| \right).
\end{aligned} \tag{B.26}$$

In (B.26), the $h_{x,n}$ are the coefficients in the α-Gegenbauer expansion

$$h_x(y) = \sum_{n=0}^{\infty} h_{x,n} G_n^{(\alpha)}(y).$$

Inequalities (3.9) and $|h_{x,n}| + |h_{x,n+1}| \leq \sqrt{2}\delta_{n-1}(h_x)$ imply (B.22).

To prove (B.23), rewrite

$$\begin{aligned}
h_x(t) = \sum_{j=1}^{\infty} a_j \frac{x^j - t^j}{x - t} &= \sum_{j=1}^{n} a_j \frac{x^j - t^j}{x - t} + \sum_{j=n+1}^{\infty} a_j \frac{x^j - t^j}{x - t} \\
&=: P_x^{(n-1)}(t) + R_x^{(n)}(t).
\end{aligned} \tag{B.27}$$

Let $\mathcal{P}_n$ be the class of all polynomials of order smaller than $n+1$. Obviously, $P_x^{(n-1)}(t)$ belongs to $\mathcal{P}_{n-1}$. Next, for every function $g \in L^2(w^{(\alpha)})$ we have

$$\delta_n(g) = \min_{P \in \mathcal{P}_{n-1}} \| g - P \|_{L^2(w^{(\alpha)})},$$

hence

$$\begin{aligned}
\delta_{n-1}(h_x) &= \delta_{n-1}(R_x^{(n)}) \leq \delta_0(R_x^{(n)}) \\
&\leq \| R_x^{(n)} \|_{L^2(w^{(\alpha)})} \leq \| R_x^{(n)} \|_{L^2}.
\end{aligned}$$

Applying Lemma 8 below to the function $R^{(n)}(t) = \sum_{j=n+1}^{\infty} a_j t^j$, we obtain

$$\| R_x^{(n)} \|_{L^2} \leq \frac{\sqrt{2\pi}}{1 - |x|} \| R^{(n)} \|_{H^2} = \frac{\sqrt{2\pi}}{1 - |x|} \left(\sum_{j=n+1}^{\infty} a_j^2 \right)^{1/2}$$

and (B.23) follows. □

Lemma 8. *Let h be a function in H^2. For any $x \in (-1,1)$, the function $h_x(t) = \frac{h(x)-h(t)}{x-t}$ satisfies the following inequality*

$$\| h_x \|_{L^2} \leq \frac{\sqrt{2\pi}}{1-|x|} \| h \|_{H^2} \tag{B.28}$$

Proof. Let $h(x) = \sum_{j=0}^{\infty} a_j x^j$ be the ordinary Taylor expansion of the function $h(x)$. The proof of (B.28) consists in two steps.

Step 1. We prove that

$$\| h_x \|_{L^2}^2 = (a_0, a_1, \dots) A^\top(x) H A(x) (a_0, a_1, \dots)^\top, \tag{B.29}$$

where $A(x) = \big(A_{i,j}(x)\big)_{i,j=0,1,\dots}$ is an upper triangular matrix whose components are

$$A_{j,k}(x) = \begin{cases} x^{k-j} & \text{if } j < k, \\ 0 & \text{otherwise,} \end{cases}$$

and where H is a Hankel matrix of the form

$$H_{j,k} = \begin{cases} 0 & \text{for } k+j \text{ odd,} \\ \frac{2}{j+k+1} & \text{for } k+j \text{ even.} \end{cases}$$

(B.29) implies that

$$\| h_x \|_{L^2}^2 \leq 2\pi \, \| A^\top(x) A(x) \| \sum_{j=0}^{\infty} a_j^2. \tag{B.30}$$

Step 2. We show that

$$\| A(x) \| \leq (1-|x|)^{-1}. \tag{B.31}$$

We start with Step 1. The sequence $A(x)(a_0, a_1, \dots)^\top = (b_1, b_2, \dots)^\top$ corresponds to the coefficients in the following expansion of $h_x(t)$:

$$h_x(t) = \sum_{j=1}^{\infty} b_j t^j.$$

We have also

$$(b_1, b_2, \dots) H (b_1, b_2, \dots)^\top = \int_{-1}^{1} h_x^2(t) \ \mathrm{d}t.$$

Indeed, by Fubini theorem,

$$\begin{aligned} \int_{-1}^{1} h_x^2(t) \ \mathrm{d}t &= \int_{-1}^{1} \sum_{k,j=1}^{\infty} b_k b_j t^{j+k} \ \mathrm{d}t = \sum_{k,j=1}^{\infty} b_k b_j \int_{-1}^{1} t^{j+k} \ \mathrm{d}t \\ &= 2 \sum_{\substack{k,j=1 \\ k+j \text{ even}}}^{\infty} \frac{b_k b_j}{k+j+1} = (b_1, b_2, \dots) H (b_1, b_2, \dots)^\top. \end{aligned}$$

This proves (B.29).

To complete the proof of the lemma, it suffices to execute Step 2. Let $k = \{k_i, i \geq 0\}$ and $\ell = \{\ell_i, i \geq 0\}$ be two sequences in l^2. We have

$$\begin{aligned} |(k_0, k_1, \dots) A(x) (\ell_0, \ell_1, \dots)^\top| &= |\sum_{k=0}^{\infty} \sum_{j=0}^{k} x^{k-j} \bar{k}_j \ell_k| \leq \sum_{u=0}^{\infty} \sum_{j=0}^{\infty} |x|^u |\bar{k}_j| |\ell_{j+u}| \\ &= \sum_{u=0}^{\infty} |x|^u \sum_{j=0}^{\infty} |\bar{k}_j| |\ell_{j+u}| \leq \sum_{u=0}^{\infty} |x|^u \, \| k \|_{l^2} \| \ell \|_{l^2} \\ &= \frac{1}{1-|x|} \ \| k \|_{l^2} \| \ell \|_{l^2}. \end{aligned}$$

Therefore, (B.31) holds. □

Acknowledgement. We would like to thank Bernhard Beckermann for his collaboration in proving Proposition 7.

References

Abramovitz, M. and Stegun, I. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Publications, New York.

Čencov, N. N. (1962). Evaluation of an unknown distribution density from observations. *Soviet Math.*, **3**, 1559–1562.

Dacunha-Castelle, D. and Oppenheim, G. (2001). Mixtures, aggregations and long-memory. Technical report, Université de Paris Sud.

Duren, P. (1970). *Theory of H^p Spaces.* Academic Press, New York.

Feller, W. (1971). *An Introduction to Probability Theory and its Applications. Vol II. 2nd ed.* Wiley Series in Probability and Mathematical Statistics, New York.

Granger, C. W. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics*, **14**, 227–238.

Hall, P. (1982). Comparison of two orthogonal series methods of estimating a density and its derivatives on an interval. *Journal of Multivariate Analysis*, **12**, 432–449.

Haubrich, J. G. and Lo, A. W. (2001). The sources and nature of long-term dependence in the business cycle. *Economic Review (Federal Reserve Bank of Cleveland, Q II)*, **37**, 15–30.

Leipus, R. and Viano, M.-C. (2000). Modelling long-memory time series with finite or infinite variance: a general approach. *Journal of Time Series Analysis*, **21**, 61–74.

Lippi, M. and Zaffaroni, P. (1998). Aggregation of simple linear dynamics: exact asymptotic results. Preprint.

Oppenheim, G. and Viano, M.-C. (2003). Aggregation of random parameters ornstein-uhlenbeck or AR processes: some convergence results. *Journal of Time Series Analysis.* Forthcoming.

Robinson, P. (1978). Statistical inference for a random coefficient autoregressive model. *Scand. J. Statistics, Theory Appl.*, **5**, 169–172.

Sansone, G. (1959). *Orthogonal Functions.* Interscience Publisher, Inc., New-York.

Szegö, G. (1967). *Orthogonal Polynomials.* American Mathematical Society, New York.

Viollaz, A. J. (1980). Asymptotic distribution of L_2 norms of the deviations of density function estimates. *Annals of Statistics* **8**, 322–346.

Whittaker, E. and Watson, G. (1927). *A Course on Modern Analysis.* Cambridge University Press, Cambridge.