

On the Convergence of Finite Order Approximations of Stationary Time Series

Symantak DATTA GUPTA¹, Ravi R. MAZUMDAR² and Peter W. GLYNN³

¹ Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
e-mail: sdattagu@engmail.uwaterloo.ca

² Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada
e-mail: mazum@ece.uwaterloo.ca

³ Department of Management Science and Engineering, Stanford University, Stanford, Ca 94305, USA e-mail:
glynn@stanford.edu

September 20, 2011

Abstract

The approximation of a stationary time-series by finite order autoregressive (AR) and moving averages (MA) is a problem that occurs in many applications. In this paper we study asymptotic behavior of the spectral density of finite order approximations of wide sense stationary time series. It is shown that when the on the spectral density is non-vanishing in $[-\pi, \pi]$ and the covariance is summable, the spectral density of the approximating autoregressive sequence converges at the origin. Under additional mild conditions on the coefficients of the Wold decomposition it is also shown that the spectral densities of both moving average and autoregressive approximations converge in L_2 as the order of approximation increases.

Keywords: Wide sense stationary time series, autoregressive estimate, moving average estimate, spectral density, Wold decomposition, time average variance constant

Short-title: Convergence of spectral density

1 Introduction

The linear estimation of a time series from a finite number of past observations is a problem encountered in many fields of application. This paper deals with the spectral properties of finite order linear approximations for a *regular* zero-mean, real-valued wide sense stationary (WSS) sequence with finite second moment. From the Wold decomposition ([10, 24]), it follows that such a process can be expressed as an infinite weighted sum of unit variance, uncorrelated random variables called the *innovation process*. This is a moving average (MA) type model for the process. By a one-step application of the projection theorem, on the other hand, one obtains an infinite order autoregressive model of the process, wherein the process is expressed as a weighted sum of all its past values, plus the current value of the innovation process.

The problem of approximating an infinite-order process and its spectral density using a finite order model has had a long history of research ([9]) and a detailed survey of the key results in this area is available in [16]. A number of problems have been studied related to choosing the optimal order for the approximation for parsimonious modeling as in the work of Akaike through the well known Akaike Information Criterion (AIC, [3]) and the Final Prediction Error criterion ([1], [2]). Related issues are the estimation of the spectral density through autoregressive (AR) models in the work of Parzen ([17, 18, 19]) or that of Priestley ([6]) which compares the performance of autoregressive models to that of window based spectral estimators. Among the papers that address the asymptotic behaviour of finite-order estimates as the model order goes to infinity, the results of [5] are noteworthy. In the works of Pourahmadi ([22, 11, 23]), the nature and rates of convergence of the autoregressive coefficients have been discussed for univariate and multivariate stochastic processes. Similar results on the rate of convergence have been considered in [13] and [21].

While there are a number of works concerned with the optimal model order; few results are available on the convergence of the spectral density of the approximating finite order autoregressive process. The main motivation of this paper is to study the asymptotic behavior of the spectral density of finite order approximation models, as the order approaches infinity and to obtain conditions under which the spectral density of the approximation converges to the true spectral density.

The value of the spectral density at the origin plays a special role in the invariance principles for stationary ergodic sequences. Let $\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k$ where $\{X_k\}$ is a WSS ergodic process. Let the value of the spectral density be finite at the origin and denote it by Γ^2 . This quantity is called the Time Average Variance Constant (TAVC) of the process ([25]) and according to the central limit theorem due to Ibragimov and Linnik ([14]),

$$\sqrt{n}(\bar{X}_n - \mu) \implies N(0, \Gamma^2)$$

where \implies denotes convergence in distribution.

Γ^2 plays an important role in steady-state simulations, where the objective is to compute the limit $\lim_{n \rightarrow \infty} \bar{X}_n$ when it exists ([4]). One way of estimating Γ^2 is by windowed estimates based on finite order autoregressive approximations of the observed stationary process. Instead of directly estimating the moments, one obtains approximations for the spectral density at the origin.

In this paper it is shown that when the spectral density of the process is strictly non-vanishing in $(-\frac{1}{2}, \frac{1}{2}]$ and its covariance sequence is in ℓ_1 , the spectral density of the approximating autoregressive sequence converges at the origin, as the order of approximation goes to infinity. It is further shown that under the additional condition that the coefficients of the Wold decomposition of the original process belong to ℓ_1 , spectral densities of both the moving average and the autoregressive estimates converge in L_2 .

2 Preliminaries

Consider a zero-mean, regular, discrete time, real-valued WSS stochastic process $\{X_n\}_{n \in \mathbb{Z}}$ defined on a probability space $(\Omega, \mathfrak{F}, \mathbb{P})$. Let $L_2(\mathbb{P})$ denote the Hilbert space of random variables with finite second moment with the inner product $\mathbb{E}[\cdot, \cdot]$ defined thereon. Two zero mean¹ random variables $X, Y \in L_2(\mathbb{P})$ are said to be orthogonal if $\mathbb{E}[XY] = 0$. Let $R_k = \mathbb{E}[X_{n+k}X_n]$ denote the covariance sequence.

The spectral density $S(\lambda)$ is defined as the Fourier transform of the covariance sequence.

$$S(\lambda) = \sum_{k \in \mathbb{Z}} R_k e^{-2\pi i \lambda k}, \quad \lambda \in \left(-\frac{1}{2}, \frac{1}{2}\right]$$

The covariance sequence can be recovered from the spectral density using the inverse Fourier Transform.

$$R_k = \int_{-\frac{1}{2}}^{\frac{1}{2}} S(\lambda) e^{2\pi i \lambda k} d\lambda, \quad k \in \mathbb{Z}$$

¹In this paper it is always assumed that the underlying random variables are of zero mean

For $n \in \mathbb{Z}$ define the space $H_n =$ linear span of $\{X_n, X_{n-1}, X_{n-2}, \dots\}$, i.e., the Hilbert subspace generated by the closure of the linear combinations of X_n and all its past values at time n .

Let Y be a random variable defined on $L_2(\mathbb{P})$. Define $\overline{\mathbb{E}}[Y|H_n]$ as the projection of Y onto the space H_n with respect to the inner-product $\mathbb{E}[\cdot, \cdot]$. Then $\overline{\mathbb{E}}[Y|H_n]$ is the minimum mean squared error (MMSE) linear estimate of Y given H_n .

Define $\{\nu_n\}_{n \in \mathbb{Z}}$ as the innovation process associated with $\{X_n\}_{n \in \mathbb{Z}}$, i.e.,

$$\nu_n = X_n - \overline{\mathbb{E}}[X_n|H_{n-1}]$$

with $\mathbb{E}[\nu_n] = 0$ and $\mathbb{E}[\nu_n \nu_k] = \sigma^2 \delta_{n-k}$, where δ_k denotes the Kronecker delta. Without loss of generality it is assumed that $\sigma^2 = 1$. By construction, ν_n is orthogonal to H_{k-1} , for all $k \leq n$. The sequence $\{\nu_k\}_{k \in \mathbb{Z}, k \leq n}$ spans the subspace H_n and constitutes an orthogonal basis in $L_2(\mathbb{P})$ for the latter. Note that $\overline{\mathbb{E}}[X_n|H_{n-1}]$ corresponds to the MMSE linear estimate of X_n given the space H_{n-1} and it can therefore be written as a linear combination of all the past values of X_n as follows:

$$\overline{\mathbb{E}}[X_n|H_{n-1}] = \sum_{k=1}^{\infty} b_k X_{n-k} \quad (2.1)$$

where the b_k s minimize the mean squared error. The corresponding mean squared error is

$$\begin{aligned} \mathbb{E}[(X_n - \sum_{k=1}^{\infty} b_k X_{n-k})^2] &= \mathbb{E}[(X_n - \overline{\mathbb{E}}[X_n|H_{n-1}])^2] \\ &= \mathbb{E}[\nu_n^2] = 1 \text{ for all } n \end{aligned} \quad (2.2)$$

$\overline{\mathbb{E}}[X_n|H_{n-1}]$ is thus an infinite-order autoregressive estimate of X_n . On the other hand, being a wide sense stationary process, X_n may be expressed in terms of its Wold decomposition as follows ([10, 24]).

$$X_n = \sum_{k=0}^{\infty} a_k \nu_{n-k} \text{ for all } n \in \mathbb{Z} \quad (2.3)$$

$$\text{with } \sum_{k=0}^{\infty} |a_k|^2 < \infty \text{ and } a_0 = 1.$$

This gives an infinite order moving average representation of X_n . The coefficients $\{b_k\}$ and $\{a_k\}$ are related to each other as follows. For each k ,

$$b_k = \sum_{j=1}^k a_j b_{k-j} \quad (2.4)$$

From (2.3) and the properties of the innovations sequence it is seen that:

$$R_k = \sum_{n=0}^{\infty} a_n a_{n-k}$$

Therefore, if the spectral density is finite at the origin then:

$$\sum_{k \in \mathbb{Z}} R_k = \sum_{k \in \mathbb{Z}} \sum_{n=0}^{\infty} a_n a_{n-k} = \left(\sum_{k=0}^{\infty} a_k \right)^2 < \infty$$

Let p be a positive integer, and define the space $H_n^p =$ linear span of $\{X_n, X_{n-1}, X_{n-2}, \dots, X_{n-p+1}\}$, i.e., the space of all linear combinations of the p most recent values of the sequence at time n , including X_n ; and all

their limits when they exist. H_n^p is then a closed Hilbert subspace of H_n , for all p and n . Clearly, $H_n^\infty = H_n$, by definition.

Define $\bar{X}_{n,p}$ as the MMSE autoregressive approximation of X_n of order p . Then $\bar{X}_{n,p}$ is a linear combination of $\{X_{n-1}, X_{n-2}, \dots, X_{n-p}\}$ and is given by

$$\bar{X}_{n,p} = \mathbb{E}[X_n | H_{n-1}^p] = \sum_{k=1}^p b_{k,p} X_{n-k} \quad (2.5)$$

where the coefficients $b_{k,p}$ minimize the error $\mathbb{E}[(X_n - \sum_{k=1}^p b_{k,p} X_{n-k})^2]$. The coefficients $b_{k,p}$ can be obtained as solutions to the modified Yule-Walker equations ([20, 15]) given by:

$$\mathbf{B}_p = \mathbf{R}_p^{-1} \mathbf{r}_p$$

where

$$\mathbf{R}_p = \begin{pmatrix} R_0 & R_1 & \cdots & R_{p-1} \\ R_1 & R_0 & \cdots & R_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ R_{p-1} & R_{p-2} & \cdots & R_0 \end{pmatrix}$$

$$\mathbf{r}_p = [R_1 \ R_2 \ \dots \ R_p]$$

and

$$\mathbf{B}_p = [b_{1,p} \ \dots \ b_{p,p}]$$

3 Moving average approximations of regular stationary sequences

Consider a moving average approximation of X_n of order p , constructed using the innovation sequence $\{\nu_n\}$. Note by assumption, $\text{Var}[\nu_n] = \sigma_\nu^2 = 1$ for all $n \in \mathbb{N}$. We first state some known facts without proof.

F 3.1. *The best p th order moving average approximation of X_n is given by $\hat{X}_{n,p} = \sum_{k=0}^p a_k \nu_{n-k}$.*

The next result is related to the mean square convergence of the p -th order approximation that readily follows from the properties of the innovation process and the Wold decomposition.

F 3.2. *As $p \rightarrow \infty$, $\hat{X}_{n,p}$ converges to X_n in quadratic mean.*

3.1 Convergence of the Spectral Density in L_2

For all functions $F : (-\frac{1}{2}, \frac{1}{2}] \rightarrow \mathbb{C}$ such that $\int_{-\frac{1}{2}}^{\frac{1}{2}} |F(\lambda)|^2 d\lambda < \infty$, define $\|\cdot\|$ to be the L_2 norm as follows:

$$\|F(\lambda)\| = \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} |F(\lambda)|^2 d\lambda \right|^{\frac{1}{2}}$$

Proposition 3.1. *Under the condition $\sum_{k=0}^{\infty} |a_k| < \infty$ where $\{a_k\}$ correspond to the coefficients in the Wold decomposition, the spectral density of $\hat{X}_{n,p}$ converges uniformly in L_2 to that of X_n as $p \rightarrow \infty$.*

Proof. Let $\sum_{k=0}^{\infty} |a_k| = S$. It follows from the properties of the innovation process, that the covariance sequence of $\{X_n\}$ is $\{R_k\}$ where

$$R_k = \sum_{l=0}^{\infty} a_l a_{l-k}$$

The spectral density $S_X(\lambda)$ is given by

$$\begin{aligned} S_X(\lambda) &= \sum_{k \in \mathbb{Z}} R_k e^{-2\pi i k \lambda} \\ &= \sum_{k \in \mathbb{Z}} \sum_{l=0}^{\infty} a_l a_{l-k} e^{-2\pi i k \lambda} \\ &= A_0^\infty(\lambda) A_0^\infty(-\lambda) \end{aligned}$$

where for $M, N \in \mathbb{N}$

$$A_M^N(\lambda) = \sum_{l=M}^N a_l e^{-2\pi i l \lambda}$$

and

$$A_M^\infty(\lambda) = \sum_{l=M}^{\infty} a_l e^{-2\pi i l \lambda}$$

Similarly, the spectral density of $\hat{X}_{n,p}$ is given by

$$S_{\hat{X}_p}(\lambda) = A_0^p(\lambda) A_0^p(-\lambda)$$

Now consider

$$\begin{aligned} & \|S_X(\lambda) - S_{\hat{X}_p}(\lambda)\| \\ &= \|A_0^\infty(\lambda) A_0^\infty(-\lambda) - A_0^p(\lambda) A_0^p(-\lambda)\| \\ &= \|(A_0^p(\lambda) A_{p+1}^\infty(-\lambda) + A_{p+1}^\infty(\lambda) A_0^\infty(-\lambda))\| \\ &\leq \|A_0^p(\lambda) A_{p+1}^\infty(-\lambda)\| + \|A_{p+1}^\infty(\lambda) A_0^\infty(-\lambda)\| \\ &= \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} I_1(p, \lambda) d\lambda \right|^{\frac{1}{2}} + \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} I_2(p, \lambda) d\lambda \right|^{\frac{1}{2}} \end{aligned} \tag{3.1}$$

where

$$\begin{aligned} I_1(p, \lambda) &= |A_0^p(\lambda) A_{p+1}^\infty(-\lambda)|^2 \\ I_2(p, \lambda) &= |A_{p+1}^\infty(\lambda) A_0^\infty(-\lambda)|^2 \end{aligned}$$

Then,

$$\begin{aligned} I_1(p, \lambda) &= |A_0^p(\lambda)|^2 |A_{p+1}^\infty(-\lambda)|^2 \\ &\leq \left(\sum_{k=0}^p |a_k| \right)^2 \left(\sum_{k=p+1}^{\infty} |a_k| \right)^2 \\ &\leq S^2 \left(\sum_{k=p+1}^{\infty} |a_k| \right)^2 \end{aligned}$$

and similarly,

$$I_2(p, \lambda) \leq S^2 \left(\sum_{k=p+1}^{\infty} |a_k| \right)^2$$

Since the sum $\sum_{k=0}^{\infty} |a_k|$ is non-decreasing and converges to $S < \infty$, for any given $\epsilon > 0$, there exists a positive integer p such that $\sum_{k=p+1}^{\infty} |a_k| < \frac{\epsilon}{2S}$, so that

$$I_1(p, \lambda) < \frac{\epsilon^2}{4} \quad (3.2)$$

$$I_2(p, \lambda) < \frac{\epsilon^2}{4} \quad (3.3)$$

Combining (3.2) and (3.3) in (3.1) yields

$$\|S_X(\lambda) - S_{\hat{X}_p}(\lambda)\| < \left(\left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\epsilon^2}{4} d\lambda \right|^{\frac{1}{2}} + \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{\epsilon^2}{4} d\lambda \right|^{\frac{1}{2}} \right) = \epsilon$$

Thus, for any $\epsilon > 0$, there exists a positive integer p such that $\|S_X(\lambda) - S_{\hat{X}_p}(\lambda)\| < \epsilon$ for all λ . Therefore, $S_{\hat{X}_p}(\lambda)$ converges uniformly in L_2 to $S_X(\lambda)$ as $p \rightarrow \infty$. \square

4 Finite AR approximations of regular stationary sequences

We begin by looking at the asymptotic behaviour of the autoregressive approximation given by (2.5). Define $\nu_{n,p}$ as the error in estimation corresponding to the AR- p approximation of $\{X_n\}$, i.e.,

$$\nu_{n,p} = X_n - \mathbb{E}[X_n | H_{n-1}^p]$$

Note that for all $p \in \mathbb{N}$

$$\begin{aligned} \overline{\mathbb{E}}[\nu_{n,p} | H_{n-1}] &= \overline{\mathbb{E}}[(X_n - \overline{\mathbb{E}}[X_n | H_{n-1}^p]) | H_{n-1}] \\ &= \overline{\mathbb{E}}[X_n | H_{n-1}] - \overline{\mathbb{E}}[X_n | H_{n-1}^p] \\ &= (X_n - \overline{\mathbb{E}}[X_n | H_{n-1}^p]) \\ &\quad - (X_n - \overline{\mathbb{E}}[X_n | H_{n-1}]) \\ &= \nu_{n,p} - \nu_n \end{aligned} \quad (4.1)$$

Let $\overline{\mathbb{E}}[\nu_{n,p} | H_{n-1}] = \epsilon_{n,p}$. Then

$$\begin{aligned} \epsilon_{n,p} &= \nu_{n,p} - \nu_n \text{ and} \\ \mathbb{E}[\epsilon_{n,p}] &= \mathbb{E}[\nu_{n,p} - \nu_n] = 0 \end{aligned}$$

As X_n is a WSS sequence in n and $\nu_{n,p}$ is constructed as a linear combination of $X_n, X_{n-1}, \dots, X_{n-p}$ whose coefficients do not depend on n , it follows that $\nu_{n,p}$ is also a WSS sequence in n . The variance of $\nu_{n,p}$, then, is only a function of p . Let $\text{Var}[\nu_{n,p}] = \sigma_p^2$. By (4.1), and the fact that ν_n is orthogonal to the subspace H_{n-1} (and hence to $\epsilon_{n,p}$) we obtain:

$$\begin{aligned} \sigma_p^2 &= \text{Var}[\nu_n] + \text{Var}[\epsilon_{n,p}] \\ &= 1 + \text{Var}[\epsilon_{n,p}] \\ &= 1 + \mathbb{E}[\epsilon_{n,p}^2] \\ &\geq 1 \text{ for all } p \in \mathbb{N} \end{aligned}$$

Note that for any $q, p \in \mathbb{N}$ such that $q > p$, $H_{n-1}^p \subset H_{n-1}^q$. It follows then, that $\overline{\mathbb{E}}[X_n | H_{n-1}^q]$ is at least as good a linear estimate of X_n as $\overline{\mathbb{E}}[X_n | H_{n-1}^p]$, in terms of mean squared error. Therefore,

$$\mathbb{E}[(X_n - \overline{\mathbb{E}}[X_n | H_{n-1}^q])^2] \leq \mathbb{E}[(X_n - \overline{\mathbb{E}}[X_n | H_{n-1}^p])^2]$$

and hence

$$\sigma_q^2 \leq \sigma_p^2$$

Hence, the sequence σ_p^2 is a non-increasing sequence in p , bounded below by 1 and must therefore has a limit as $p \rightarrow \infty$ that is bounded from below by 1. It can be shown that the limit is in fact equal to 1.

Remark 4.1. As $p \rightarrow \infty$, $\overline{\mathbb{E}}[X_n|H_{n-1}^p] \rightarrow \overline{\mathbb{E}}[X_n|H_{n-1}]$ and $\nu_{n,p} \rightarrow \nu_n$ in quadratic mean.

The proof is a direct application of [10, Lemma 3.1(b)].

4.1 Convergence of the Spectral Density at the Origin

We now study conditions under which the spectral density of the finite order AR approximation converges at the origin. As mentioned $S(0)$ is referred to as the Time Average Variance Constant (TAVC) and plays an important role in simulations. Throughout the following it is assumed that the spectral density $S(\lambda)$ is non-vanishing in $(-\frac{1}{2}, \frac{1}{2}]$ i.e. there exists $\varepsilon > 0$ such that $S(\lambda) > \varepsilon$ for all $\lambda \in (-\frac{1}{2}, \frac{1}{2}]$.

We start with the following lemma on the pointwise convergence of the coefficients $b_{k,p}$ ([12, Proposition 3.1]).

Lemma 4.1. As $p \rightarrow \infty$, $b_{k,p} \rightarrow b_k$ for each $k \in \mathbb{N}$.

Proof. For each $p \in \mathbb{N}$ and $k \in \{1, \dots, p\}$ define $\tilde{b}_{0,p} = 1$ and $\tilde{b}_{k,p} = -b_{k,p}$. For any p , ν_n^p is given by

$$\nu_{n,p} = \sum_{k=0}^p \tilde{b}_{k,p} X_{n-k}$$

The above can be written in matrix form as $\nu = B\mathbf{X}$ where $\nu = [\nu_{n,0} \cdots \nu_{n,p}]^T$, $\mathbf{X} = [X_n \cdots X_{n-p}]^T$ and B is a lower triangular matrix whose first column is $[\tilde{b}_{0,p} \cdots \tilde{b}_{p,p}]$. The matrix B is invertible with inverse A which satisfies $\mathbf{X} = A\nu$. The inverse A is lower triangular with first column $[a_{0,p} \cdots a_{p,p}]$ and the elements $a_{i,j}$ satisfy for all $n \in \mathbb{N}$

$$X_n = \sum_{k=0}^p a_{k,p} \nu_{n-k,p-k} \quad (4.2)$$

and for $p > k$

$$\tilde{b}_{k,p} = - \sum_{j=1}^k a_{j,p} \tilde{b}_{k-j,p-j} \quad (4.3)$$

By definition, $\nu_{n,n}$ and $\nu_{m,m}$ are orthogonal to each other for $n \neq m$ and hence (4.2) provides an orthogonal decomposition of X_n . Therefore,

$$\mathbb{E}[X_n \nu_{n-k,p-k}] = a_{k,p} \sigma_{p-k}^2 \quad (4.4)$$

From (2.3), on the other hand, we have

$$\mathbb{E}[X_n \nu_{n-k}] = a_k \quad (4.5)$$

However, by remark 4.1, $\nu_{n,p} \rightarrow \nu_n$ in quadratic mean. It then follows from (4.4) and (4.5) that for all $k \in \mathbb{N}$

$$\begin{aligned} & \lim_{p \rightarrow \infty} |a_{k,p} \sigma_{p-k}^2 - a_k| \\ &= \lim_{p \rightarrow \infty} |\mathbb{E}[X_n \nu_{n-k,p-k}] - \mathbb{E}[X_n \nu_{n-k}]| \\ &\leq \lim_{p \rightarrow \infty} |\mathbb{E}[X_n^2] \mathbb{E}[(\nu_{n-k,p-k} - \nu_{n-k})^2]|^{\frac{1}{2}} = 0 \end{aligned} \quad (4.6)$$

and hence,

$$\lim_{p \rightarrow \infty} a_{k,p} = a_k \quad (4.7)$$

Finally, to show the pointwise convergence of $\tilde{b}_{k,p}$ (and therefore that of $b_{k,p}$) first observe that $\tilde{b}_{0,p} = b_0 = 1$ holds for all p . Let

$$\lim_{p \rightarrow \infty} \tilde{b}_{j,p} = -b_j$$

for all $j \leq k$. Then, using the recursive relation given by (4.3) and comparing with (2.4), one obtains

$$\lim_{p \rightarrow \infty} \tilde{b}_{k+1,p} = -b_{k+1}$$

Therefore, by the principle of mathematical induction, as $p \rightarrow \infty$, $\tilde{b}_{k,p} \rightarrow -b_k$; i.e., $b_{k,p} \rightarrow b_k$ for each $k \in \mathbb{N}$. \square

Next, we present a key result on the summability of the autoregressive coefficients known as Baxter's inequality [5, Theorem 2.2]).

Let $\{X_n\}$ be a WSS process with spectral density function $S_X(\lambda) > 0$ and let $\bar{X}_{n,p}$ be the p -th order MMSE linear predictor of X_n , defined by 2.5 and let σ_p^2 be the corresponding mean squared error. Let $\{b_k\}$ be the limits of the coefficients $\{b_{k,p}\}$ and let $\sigma^2 > 0$ be the limit of σ_p^2 as $p \rightarrow \infty$ (in our case $\sigma^2 = 1$ by remark 4.1). Define the sequence $\{u_{k,p}\}$ as $u_{k,p} = -\frac{b_{k,p}}{\sigma_p^2}$ and let $\{U_k\}$ be the limit of $\{u_{k,p}\}$. Then, the theorem is stated as follows.

Theorem 4.1. Baxter's Inequality: *If $S_X(\lambda)$ is a positive continuous function whose Fourier coefficients have γ moments, then there exists an integer $N > 0$ and a constant $c > 0$, both depending only on $S_X(\lambda)$ such that for all $p \geq N$,*

$$\sum_{k=1}^p (2^\gamma + k^\gamma) |u_{k,p} - U_k| \leq c \sum_{k=p+1}^{\infty} (2^\gamma + k^\gamma) |U_k|$$

Note that the Fourier coefficients of the spectral density are the elements of the covariance sequence $\{R_k\}$.

The above theorem can be used to establish the following lemma on the convergence of the coefficients $\{b_{k,p}\}$ as $p \rightarrow \infty$.

Lemma 4.2. *When the spectral density of $\{X_n\}$ is strictly positive in $\lambda \in (-\frac{1}{2}, \frac{1}{2}]$, and the covariance sequence is in ℓ_1 , i.e.,*

$$\sum_{k \in \mathbb{Z}} |R_k| < \infty$$

then

$$\lim_{p \rightarrow \infty} \sum_{k=1}^p |b_{k,p} - b_k| = 0$$

Proof. The proof is a simple application of Baxter's inequality. Note that when the covariance sequence is in ℓ_1 , the spectral density is continuous. Pointwise convergence of the $b_{k,p}$ to b_k for each k follows from lemma 4.1. Moreover, summability of the covariance sequence also imply that the sequence has a finite 0-th moment ($\gamma = 0$ in Theorem 4.1). It then follows from Theorem 4.1 that under the assumption that the spectral density of $\{X_n\}$ is strictly positive in $\lambda \in (-\frac{1}{2}, \frac{1}{2}]$, there exists a positive integer N and a constant $c > 0$ such that

$$\sum_{k=1}^p \left| \frac{b_{k,p}}{\sigma_p^2} - b_k \right| \leq c \sum_{k=p+1}^{\infty} |b_k| \quad (4.8)$$

for all $p > N$. Since the covariance sequence has been assumed to be in ℓ_1 , the sequence of autoregressive coefficients of the original process are also in ℓ_1 ([7]), i.e.,

$$\sum_{k=1}^{\infty} |b_k| < \infty$$

Then, for any $\epsilon > 0$, there exists an integer N_0 such that

$$\sum_{k=p+1}^{\infty} |b_k| < \frac{\epsilon}{c}$$

for all $p > N_0$. Define $N^* = \max\{N, N_0\}$. Then for all $p > N^*$,

$$\sum_{k=1}^p \left| \frac{b_{k,p}}{\sigma_p^2} - b_k \right| < \epsilon$$

Therefore,

$$\lim_{p \rightarrow \infty} \sum_{k=1}^p \left| \frac{b_{k,p}}{\sigma_p^2} - b_k \right| = 0 \quad (4.9)$$

It follows from the triangle inequality that

$$\begin{aligned} \sum_{k=1}^p \left| \frac{b_{k,p} - b_k}{\sigma_p^2} \right| &\leq \sum_{k=1}^p \left| \frac{b_{k,p}}{\sigma_p^2} - b_k \right| + \sum_{k=1}^p \left| \frac{b_k}{\sigma_p^2} - b_k \right| \\ &\leq \sum_{k=1}^p \left| \frac{b_{k,p}}{\sigma_p^2} - b_k \right| + \frac{\sigma_p^2 - 1}{\sigma_p^2} \sum_{k=1}^p |b_k| \end{aligned}$$

Therefore, as $p \rightarrow \infty$,

$$\begin{aligned} \lim_{p \rightarrow \infty} \sum_{k=1}^p \left| \frac{b_{k,p} - b_k}{\sigma_p^2} \right| &\leq \lim_{p \rightarrow \infty} \sum_{k=1}^p \left| \frac{b_{k,p}}{\sigma_p^2} - b_k \right| \\ &\quad + \lim_{p \rightarrow \infty} \frac{\sigma_p^2 - 1}{\sigma_p^2} \sum_{k=1}^p |b_k| \end{aligned}$$

By (4.9), the first limit on the right hand side is zero and by remark 4.1,

$$\lim_{p \rightarrow \infty} \sigma_p^2 = 1$$

Therefore,

$$\lim_{p \rightarrow \infty} \sum_{k=1}^p |b_{k,p} - b_k| = 0$$

□

Proposition 4.1. *Let $S_X(\lambda) > 0$ for $\lambda \in (-\frac{1}{2}, \frac{1}{2}]$ and let $\sum_{k \in \mathbb{Z}} |R_k| < \infty$. Then, as $p \rightarrow \infty$, the spectral density of $\overline{\mathbb{E}}[X_n | H_{n-1}^p] = \overline{X}_{n,p}$ converges to that of $\overline{\mathbb{E}}[X_n | H_{n-1}]$ at the origin.*

Proof. Let $\{\bar{R}_{k,p}\}_{k \in \mathbb{Z}}$, $\{\bar{R}_k\}_{k \in \mathbb{Z}}$ be the covariance sequences of $\bar{\mathbb{E}}[X_n|H_{n-1}^p]$, $\bar{\mathbb{E}}[X_n|H_{n-1}]$ respectively. If $S_{\bar{X}_p}(\lambda)$, $S_{\bar{X}}(\lambda)$ be the spectral densities of the two processes respectively, then

$$S_{\bar{X}_p}(\lambda) = \sum_{k \in \mathbb{Z}} \bar{R}_{k,p} e^{-2\pi i \lambda k}$$

$$S_{\bar{X}}(\lambda) = \sum_{k \in \mathbb{Z}} \bar{R}_k e^{-2\pi i \lambda k}$$

At the origin, i.e., at $\lambda = 0$, the above spectral densities are given by

$$S_{\bar{X}_p}(0) = \sum_{k \in \mathbb{Z}} \bar{R}_{k,p}$$

$$S_{\bar{X}}(0) = \sum_{k \in \mathbb{Z}} \bar{R}_k$$

For some p and k , $\bar{R}_{k,p}$ may be obtained from (2.5) as

$$\begin{aligned} \bar{R}_{k,p} &= \mathbb{E}[\bar{X}_{n,p} \bar{X}_{n-k,p}] \\ &= \mathbb{E} \left[\sum_{j=1}^p b_{j,p} X_{n-j} \sum_{l=1}^p b_{l,p} X_{n-k-l} \right] \\ &= \sum_{j=1}^p b_{j,p}^2 R_k + \sum_{t=1}^{p-1} \sum_{j=1}^{p-t} b_{j,p} b_{j+t,p} (R_{k-t} + R_{k+t}) \end{aligned} \tag{4.10}$$

Summing the above over all $k \in \mathbb{Z}$ gives

$$\begin{aligned} \sum_{k \in \mathbb{Z}} \bar{R}_{k,p} &= \left(\sum_{j=1}^p b_{j,p}^2 + 2 \sum_{t=1}^{p-1} \sum_{j=1}^{p-t} b_{j,p} b_{j+t,p} \right) \sum_{k \in \mathbb{Z}} R_k \\ &= \left(\sum_{j=1}^p b_{j,p} \right)^2 \sum_{k \in \mathbb{Z}} R_k \end{aligned} \tag{4.11}$$

From where it follows that

$$\lim_{p \rightarrow \infty} \sum_{k \in \mathbb{Z}} \bar{R}_{k,p} = \lim_{p \rightarrow \infty} \left(\sum_{j=1}^p b_{j,p} \right)^2 \sum_{k \in \mathbb{Z}} R_k \tag{4.12}$$

Proceeding similarly, using the expression in (2.1), one obtains the following expression for $\sum_{k \in \mathbb{Z}} \bar{R}_k$:

$$\sum_{k \in \mathbb{Z}} \bar{R}_k = \left(\sum_{j=1}^{\infty} b_j \right)^2 \sum_{k \in \mathbb{Z}} R_k \tag{4.13}$$

Recall that by lemma 4.2, under the stated conditions,

$$\lim_{p \rightarrow \infty} \sum_{k=1}^p |b_{k,p} - b_k| = 0 \quad (4.14)$$

Clearly, then,

$$\lim_{p \rightarrow \infty} \sum_{k=1}^p (b_{k,p} - b_k) = 0 \quad (4.15)$$

i.e.,

$$\lim_{p \rightarrow \infty} \left(\sum_{k=1}^p b_{k,p} \right)^2 = \left(\sum_{k=1}^{\infty} b_k \right)^2 \quad (4.16)$$

Combining (4.16) with (4.12), and comparing with (4.13), we obtain

$$\begin{aligned} \lim_{p \rightarrow \infty} \sum_{k \in \mathbb{Z}} \bar{R}_{k,p} &= \left(\sum_{k=1}^{\infty} b_k \right)^2 \sum_{k \in \mathbb{Z}} R_k \\ &= \sum_{k \in \mathbb{Z}} \bar{R}_k \end{aligned} \quad (4.17)$$

This completes the proof. □

Remark 4.2. *Combining the above result with the fact that $\nu_{n,p}$ converges to ν_n in quadratic mean it readily follows that the spectral density of $\{\bar{X}_{n,p} + \nu_{n,p}\}$ converges to the spectral density of $\{X_n\}$ at the origin, i.e., at $\lambda = 0$.*

4.2 Convergence of the Spectral Density in L_2

We now present a sufficient condition for the L_2 convergence of the spectral density of the autoregressive approximation as $p \rightarrow \infty$.

Proposition 4.2. *Let $S_X(\lambda) > 0$ for $\lambda \in (-\frac{1}{2}, \frac{1}{2}]$ and let*

$$\sum_{k \in \mathbb{Z}} |a_k| < \infty \quad (4.18)$$

Then as $p \rightarrow \infty$, $S_{\bar{X}_p}(\lambda)$ converges to $S_{\bar{X}}(\lambda)$ in L_2 .

Proof. We begin by noting that when the sequence $\{a_k\}$ is in ℓ_1 , (as stated above), both $\{R_k\}$ and $\{b_k\}$ are also in ℓ_1 and therefore the conditions of Lemma 4.2 are satisfied.

Refer to equation (4.10) for an expansion of $\bar{R}_{k,p}$ for each k , for a given p :

$$\bar{R}_{k,p} = \sum_{j=1}^p b_{j,p}^2 R_k + \sum_{t=1}^{p-1} \sum_{j=1}^{p-t} b_{j,p} b_{j+t,p} (R_{k-t} + R_{k+t})$$

Consider the WSS process given by $\sum_{j=1}^p b_j X_{n-k}$ and let $\{R_{k,p}\}$ be its covariance sequence. Proceeding as in the case of (4.10), we can obtain a similar expression for $R_{k,p}$ as follows.

$$R_{k,p} = \sum_{j=1}^p b_j^2 R_k + \sum_{t=1}^{p-1} \sum_{j=1}^{p-t} b_j b_{j+t} (R_{k-t} + R_{k+t}) \quad (4.19)$$

so that for all $k \in \mathbb{Z}$

$$\begin{aligned}
& |R_{k,p} - \overline{R}_{k,p}| \\
&= \left| \sum_{j=1}^p (b_j^2 - b_{j,p}^2) R_k \right. \\
&\quad \left. + \sum_{t=1}^{p-1} \sum_{j=1}^{p-t} (b_j b_{j+t} - b_{j,p} b_{j+t,p}) (R_{k-t} + R_{k+t}) \right| \\
&\leq \sum_{j=1}^p |(b_j^2 - b_{j,p}^2)| |R_k| \\
&\quad + \sum_{t=1}^{p-1} \sum_{j=1}^{p-t} |(b_j b_{j+t} - b_{j,p} b_{j+t,p})| (|R_{k-t}| + |R_{k+t}|)
\end{aligned} \tag{4.20}$$

Summing over all $k \in \mathbb{Z}$

$$\begin{aligned}
& \sum_{k \in \mathbb{Z}} |R_{k,p} - \overline{R}_{k,p}| \\
&\leq \left(\sum_{j=1}^p |b_j^2 - b_{j,p}^2| \right. \\
&\quad \left. + 2 \sum_{t=1}^{p-1} \sum_{j=1}^{p-t} |b_j b_{j+t} - b_{j,p} b_{j+t,p}| \right) \sum_{k \in \mathbb{Z}} |R_k| \\
&= \left(\sum_{j=1}^p \sum_{i=1}^p |b_i b_j - b_{i,p} b_{j,p}| \right) \sum_{k \in \mathbb{Z}} |R_k| \\
&\leq \left(\sum_{j=1}^p \sum_{i=1}^p |b_i| |b_j - b_{j,p}| + \sum_{j=1}^p \sum_{i=1}^p |b_{j,p}| |b_i - b_{i,p}| \right) \sum_{k \in \mathbb{Z}} |R_k| \\
&= \left(\sum_{i=1}^p |b_i| + \sum_{i=1}^p |b_{i,p}| \right) \left(\sum_{i=1}^p |b_i - b_{i,p}| \right) \sum_{k \in \mathbb{Z}} |R_k|
\end{aligned} \tag{4.21}$$

As the covariance sequence has been assumed to be in ℓ_1 , the sequence $\{b_k\}$ is also in ℓ_1 ([7]) and by Lemma (4.2), as $p \rightarrow \infty$ the second term goes to 0. Finally, by the same lemma,

$$\lim_{p \rightarrow \infty} \sum_{i=1}^p |b_{i,p}| \leq \lim_{p \rightarrow \infty} \sum_{i=1}^p |b_i|$$

Therefore,

$$\lim_{p \rightarrow \infty} \sum_{k \in \mathbb{Z}} |R_{k,p} - \overline{R}_{k,p}| = 0 \tag{4.22}$$

Now note that

$$\begin{aligned}
R_{k,p} &= \mathbb{E}[(X_n - \nu_n - \sum_{i=p+1}^{\infty} b_i X_{n-i})(X_{n+k} \\
&\quad - \nu_{n+k} - \sum_{j=p+1}^{\infty} b_j X_{n+k-j})] \\
&= \bar{R}_k - \sum_{j=p+1}^{\infty} b_j R_{k-j} - \sum_{i=p+1}^{\infty} b_i R_{k+i} \\
&\quad + \sum_{j=p+1}^{\infty} b_j \mathbb{E}[X_{n+k-j} \nu_n] \\
&\quad + \sum_{i=p+1}^{\infty} b_i \mathbb{E}[X_{n-i} \nu_{n+k}] \\
&\quad + \mathbb{E} \left[\left(\sum_{i=p+1}^{\infty} b_i X_{n-i} \right) \left(\sum_{j=p+1}^{\infty} b_j X_{n+k-j} \right) \right] \\
&= \bar{R}_k - \sum_{j=p+1}^{\infty} b_j R_{k-j} - \sum_{i=p+1}^{\infty} b_i R_{k+i} \\
&\quad + \sum_{j=p+1}^{\infty} b_j a_{|k|-j} + \left(\sum_{i=p+1}^{\infty} b_i^2 \right) R_k \\
&\quad + \sum_{t=1}^{\infty} \sum_{i=p+1}^{\infty} b_i b_{i+t} (R_{k+t} + R_{k-t})
\end{aligned} \tag{4.23}$$

Therefore,

$$\begin{aligned}
&|R_{k,p} - \bar{R}_k| \\
&\leq \sum_{j=p+1}^{\infty} |b_j| |R_{k-j}| + \sum_{i=p+1}^{\infty} |b_i| |R_{k+i}| \\
&\quad + \sum_{j=p+1}^{\infty} |b_j| |a_{|k|-j}| + \left(\sum_{i=p+1}^{\infty} |b_i|^2 \right) |R_k| \\
&\quad + \sum_{t=1}^{\infty} \sum_{i=p+1}^{\infty} |b_i| |b_{i+t}| (|R_{k+t}| + |R_{k-t}|)
\end{aligned}$$

Summing over all $k \in \mathbb{Z}$,

$$\begin{aligned}
& \sum_{k \in \mathbb{Z}} |R_{k,p} - \bar{R}_k| \\
& \leq \left(\sum_{j=p+1}^{\infty} |b_j| \right) \sum_{k \in \mathbb{Z}} |R_{k-j}| + \left(\sum_{i=p+1}^{\infty} |b_i| \right) \sum_{k \in \mathbb{Z}} |R_{k+i}| \\
& + \left(\sum_{j=p+1}^{\infty} |b_j| \right) \sum_{k \in \mathbb{Z}} |a_{|k|-j}| + \left(\sum_{i=p+1}^{\infty} |b_i|^2 \right) \sum_{k \in \mathbb{Z}} |R_k| \\
& + \left(\sum_{t=1}^{\infty} \sum_{i=p+1}^{\infty} |b_i| |b_{i+t}| \right) \left(\sum_{k \in \mathbb{Z}} |R_{k+t}| + \sum_{k \in \mathbb{Z}} |R_{k-t}| \right) \\
& = \left(\sum_{j=p+1}^{\infty} |b_j| \right) \left(2 \sum_{k \in \mathbb{Z}} |R_k| + \sum_{k=0}^{\infty} |a_k| \right) \\
& + \left(\sum_{i=p+1}^{\infty} |b_i| \right)^2 \sum_{k \in \mathbb{Z}} |R_k|
\end{aligned} \tag{4.24}$$

As $p \rightarrow \infty$, each term on the right hand side of the above inequality goes to zero, because the covariance sequence and the sequences $\{b_k\}$ and $\{a_k\}$ are in ℓ_1 . Therefore,

$$\lim_{p \rightarrow \infty} \sum_{k \in \mathbb{Z}} |R_{k,p} - \bar{R}_k| = 0 \tag{4.25}$$

Combining the results of (4.22) and (4.25) we obtain

$$\lim_{p \rightarrow \infty} \sum_{k \in \mathbb{Z}} |\bar{R}_{k,p} - \bar{R}_k| = 0 \tag{4.26}$$

Finally,

$$\begin{aligned}
& \lim_{p \rightarrow \infty} \|S_{\bar{X}_p}(\lambda) - S_{\bar{X}}(\lambda)\| \\
& = \lim_{p \rightarrow \infty} \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \left| \sum_{k \in \mathbb{Z}} (\bar{R}_{k,p} - \bar{R}_k) e^{-2\pi i \lambda k} \right|^2 d\lambda \right|^{\frac{1}{2}} \\
& \leq \lim_{p \rightarrow \infty} \left| \int_{-\frac{1}{2}}^{\frac{1}{2}} \sum_{k \in \mathbb{Z}} |\bar{R}_{k,p} - \bar{R}_k|^2 d\lambda \right|^{\frac{1}{2}} \\
& \leq \lim_{p \rightarrow \infty} \left(\sum_{k \in \mathbb{Z}} |\bar{R}_{k,p} - \bar{R}_k| \right) \\
& = 0
\end{aligned} \tag{4.27}$$

where (4.26) is used for the last equality. This completes the proof. \square

Remark 4.3. Combining the above result with the fact that $\nu_{n,p}$ converges to ν_n in quadratic mean it readily follows that the spectral density of $\{\bar{X}_{n,p} + \nu_{n,p}\}$ converges to the spectral density of $\{X_n\}$ in L_2 .

5 Conclusion

In this paper it has been shown that the spectral density of a finite autoregressive approximation of a WSS process converges at the origin when the spectral density is strictly positive. Thus, any unbiased spectral estimator derived from a finite autoregressive approximation will converge to the spectrum of the original process at the origin. This would enable easy approximation of the TAVC, which is important in the context of steady-state simulation.

Furthermore, it has been shown that the spectral density of both the moving average and the autoregressive type approximations converge in L_2 when the sequence of the Wold expansion coefficients is in ℓ_1 . For a zero mean wide sense stationary processes having an infinite order moving average representation, the condition $\sum_{0 \leq k < \infty} |a_k| < \infty$ is met when the covariance R_k tends to zero at an exponential rate as $k \rightarrow \infty$ asymptotically; i.e., there exist constants $C \in \mathbb{R}$, $\alpha \in (0, 1)$ such that $|R_k| \sim C\alpha^{|k|}$ ([8]). A more trivial example is that of a process which has a finite order moving average representation. In such cases, the spectral density of the original process can be approximated over $\lambda \in [-\frac{1}{2}, \frac{1}{2}]$ from that of the finite order estimate.

Acknowledgment

The research of SDG was supported in part by the Natural Sciences and Engineering Council of Canada (NSERC) through the Strategic Grant program.

References

- [1] H. Akaike. Fitting autoregressive models for prediction. *Ann. Inst. Statist. Math.*, 21:243–247, 1969.
- [2] H. Akaike. Statistical predictor identification. *Ann. Inst. Statist. Math.*, 22:203–207, 1970.
- [3] H. Akaike. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, 19(6):716 – 723, 1974.
- [4] S. Asmussen and P.W. Glynn. *Stochastic Simulation*. Springer, 2007.
- [5] G. Baxter. An asymptotic result for the finite predictor. *Math. Scand.*, 10:137–144, 1962.
- [6] N. Beamish and M. B. Priestley. A study of autoregressive and window spectral estimation. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 30(1):41–58, 1981.
- [7] R. J. Bhansali. The criterion autoregressive transfer function of Parzen. *Journal of Time Series Analysis*, 7(2):79–104, 1986.
- [8] R.J. Bhansali and P. S. Kokoszka. Prediction of long memory time series: An overview. *Estadistica*, 53:41–96, 2001.
- [9] P.M.T. Broersen. *Automatic Autocorrelation and Spectral Analysis*. Springer, 2006.
- [10] P. Caines. *Linear Stochastic Systems*. John Wiley & Sons, 1988.
- [11] R. Cheng and M. Pourahmadi. Baxter’s inequality and convergence of finite predictors of multivariate stochastic processes. *Probability Theory and Related Fields*, 95:115–124, 1993.
- [12] S. Degerine. Partial autocorrelation function for a scalar stationary discrete-time process. In *Alternative Approaches to Time Series Analysis, Proceedings of the 3rd Franco-Belgian Meeting of Statisticians*, pages 79–94, 1982.

- [13] D. F. Findley. Convergence of finite multistep predictors from incorrect models and its role in model selection. *Note di Matematica*, XI:145–155, 1991.
- [14] I.A. Ibragimov and J.V. Linnik. *Independent and stationary sequences of random variables*. Groningen, Wolters-Noordhoff, 1971.
- [15] T. Kailath, A. H. Sayed, and B. Hassibi. *Linear Estimation*. Prentice Hall, 2000.
- [16] S.M. Kay and Jr. Marple, S.L. Spectrum analysis; a modern perspective. *Proceedings of the IEEE*, 69(11):1380 – 1419, 1981.
- [17] E. Parzen. Multiple time series modelling. *Multivariate Analysis II*, pages 398–410, 1969.
- [18] E. Parzen. Solutions to the stationary time series modeling and prediction problem. In *Decision and Control including the 13th Symposium on Adaptive Processes, 1974 IEEE Conference on*, volume 13, pages 468 –473, 1974.
- [19] E. Parzen. Some recent advances in time series modeling. *Automatic Control, IEEE Transactions on*, 19(6):723 – 730, 1974.
- [20] B. Picinbono. *Random Signals and Systems*. Prentice Hall, 1993.
- [21] D. S. Poskitt. A note on autoregressive modeling. *Econometric Theory*, 10(5):884–899, 1994.
- [22] M. Pourahmadi. On the convergence of finite linear predictors of stationary processes. *Journal of Multivariate Analysis*, 30(2):167–180, 1989.
- [23] M. Pourahmadi. *Foundations of Time Series Analysis and Prediction Theory*. John Wiley, 2001.
- [24] A. N. Shiryaev. *Probability*. Springer, 1996.
- [25] W. B. Wu. Recursive estimation of time-average variance constants. *Annals of Applied Probability*, 19(4):1529–1552, 2009.