On predictors for band-limited and high-frequency time series*

Nikolai Dokuchaev

Department of Mathematics and Statistics, Curtin University,

GPO Box U1987, Perth, Western Australia, 6845

email N.Dokuchaev@curtin.edu.au. Tel.: 61 8 92663144.

1st version: December 1,2011. Revised version: March 15, 2012

Abstract

Pathwise predictability and predictors for discrete time processes are studied in deterministic setting. It is suggested to approximate convolution sums over future times by convolution sums over past time. It is shown that all band-limited processes are predictable in this sense, as well as high-frequency processes with zero energy at low frequencies. In addition, a process of mixed type still can be predicted if an ideal low-pass filter exists for this process.

Key words: prediction, spectral methods, z-transform, band-limited processes, low-pass filters, non-parametric forecast.

AMS 2000 classification : 42A38, 93E10, 42B30

PACS 2008 numbers: 02.30.Mv, 02.30.Nw, 02.30.Yy, 07.05.Mh, 07.05.Kf

1 Introduction

We study pathwise predictability of discrete time processes in deterministic setting. It is well known that certain restrictions on frequency distribution can ensure additional opportunities for prediction and interpolation of the processes. The classical result is Nyquist-Shannon-Kotelnikov interpolation theorem for the continuous time band-limited processes. It is also known that optimal prediction error for stationary Gaussian processes is zero for the case of degenerate spectral density. The related results can be found in Wainstein and Zubakov (1962), Knab (1981), Papoulis (1985), Marvasti (1986), Vaidyanathan (1987), Lyman *et al* (2000, 2001), Dokuchaev (2008,2010).

The present paper extends on discrete time setting the approach suggested for continuous time processes in Dokuchaev (2008). We study a special kind of predictors such that convolution sums over

* *Signal Processing* **92**, iss. 10, pp.2571-2575. In *Fast communications* section.

future are approximated by convolution sums over past times representing historical observations. We found some cases when this approximation can be made uniformly over a wide class of input processes, including all band-limited processes and high-frequency processes. For the processes of mixed type, we found that the similar predictability can be achieved when the model allows a low pass filter that acts as an ideal low-pass filter for this process. These results can be a useful addition to the existing theory of band-limited processes. The novelty is that we consider predictability of both high frequent and band-limited processes in a weak sense uniformly over classes of input processes. In addition, we suggest a new type of predictor. Its kernel is given explicitly in the frequency domain.

2 Definitions

Let $D = \{z \in \mathbf{C} : |z| \le 1\}$, $D^c = \mathbf{C} \setminus D$, $\mathbb{T} = \{z \in \mathbf{C} : \ |z| = 1\}$,

We denote by ℓ_r the set of all sequences $x = \{x(t)\}_{t=-\infty}^{\infty} \subset \mathbf{C}$ such that $\|x\|_{\ell_r} = \left(\sum_{t=-\infty}^{\infty} |x(t)|^r\right)^{1/r} < +\infty$ for $r \in [1, \infty)$, $\|x\|_{\ell_\infty} = \sup_t |x(t)| < +\infty$ for $r = +\infty$.

Let ℓ_r^+ be the set of all sequences $x \in \ell_r$ such that $x(t) = 0$ for $t = -1, -2, -3, ...$

For complex valued sequences $x \in \ell_1$ or $x \in \ell_2$, we denote by $X = \mathcal{Z}x$ the Z-transform

$$X(z) = \sum_{t=-\infty}^{\infty} x(t) z^{-t}, \quad z \in \mathbf{C}.$$

Respectively, the inverse $x = \mathcal{Z}^{-1}X$ is defined as

$$x(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X\left(e^{i\omega}\right) e^{i\omega t} d\omega, \quad t = 0, \pm 1, \pm 2, ...$$

If $x \in \ell_2$, then $X|_{\mathbb{T}}$ is defined as an element of $L_2(\mathbb{T})$.

Let H^r be the Hardy space of functions that are holomorphic on D^c including the point at infinity (see, e.g., Duren (1970)). Note that Z-transform defines bijection between the sequences from ℓ_2^+ and the restrictions (i.e. traces) of the functions from H^2 on $\mathbb{T}$.

Definition 1 *Let $\mathcal{K}$ be the class of all functions $k \in \ell_\infty$ such that $k(t) = 0$ for $t > 0$ and $K = \mathcal{Z}k$ is*

$$K(z) = \frac{d(z)}{\delta(z)}, \tag{2.1}$$

where $d(\cdot)$ and $\delta(\cdot)$ are polynomials such that $\deg d < \deg \delta$, and if $\delta(z) = 0$ for $z \in \mathbf{C}$ then $|z| > 1$.

The class includes all kernels k representing the anti-causal linear constant-coefficient difference equations.

Definition 2 *Let $\widehat{\mathcal{K}}$ be the class of functions $\widehat{k} : \ell_\infty^+$ such that the function $\widehat{K}(\cdot) = \mathcal{Z}\widehat{k}$ belongs to $H^\infty \cap H^2$.*

It follows from the definitions that if $\widehat{k} \in \widehat{\mathcal{K}}$ then $\widehat{k}(t) = 0$ for $t < 0$.

We are going to study linear predictors in the form $\widehat{y}(t) = \sum_{s=-\infty}^{t} \widehat{k}(t-s)x(s)$ for the processes $y(t) = \sum_{s=t}^{+\infty} k(t-s)x(s)$, where $k \in \mathcal{K}$ and $\widehat{k} \in \widehat{K}$. The predictors use historical values of currently observable process $x(\cdot)$.

Definition 3 *Let $\mathcal{X} = \{x(\cdot)\}$ be a class of sequences from ℓ_∞, let $r \in [1, +\infty]$, and let $\widetilde{\mathcal{K}} \subset \mathcal{K}$ be a class of sequences.*

(i) We say that the class $\mathcal{X}$ is ℓ_r-predictable in the weak sense with respect to the class $\widetilde{\mathcal{K}}$ if, for any $k(\cdot) \in \widetilde{\mathcal{K}}$, there exists a sequence $\{\widehat{k}_m(\cdot)\}_{m=1}^{+\infty} = \{\widehat{k}_m(\cdot, \mathcal{X}, k)\}_{m=1}^{+\infty} \subset \widehat{\mathcal{K}}$ such that

$$\|y - \widehat{y}_m\|_{\ell_r} \to 0 \quad \text{as} \quad m \to +\infty \quad \forall x \in \mathcal{X},$$

where

$$y(t) = \sum_{s=t}^{+\infty} k(t-s)x(s), \qquad \widehat{y}_m(t) = \sum_{s=-\infty}^{t} \widehat{k}_m(t-s)x(s).$$

(ii) Let the set $\mathcal{Z}(\mathcal{X}) = \{X(e^{i\omega}) = \mathcal{Z}x|_{\mathbb{T}}, \quad x \in \mathcal{X}\}$ be provided with a norm $\|\cdot\|$. We say that the class $\mathcal{X}$ is ℓ_r-predictable in the weak sense with respect to the class $\widetilde{\mathcal{K}}$ uniformly with respect to the norm $\|\cdot\|$, if, for any $k(\cdot) \in \widetilde{\mathcal{K}}$ and $\varepsilon > 0$, there exists $\widehat{k}(\cdot) = \widehat{k}(\cdot, \mathcal{X}, k, \|\cdot\|, \varepsilon) \in \widehat{\mathcal{K}}$ such that

$$\|y - \widehat{y}\|_{\ell_r} \le \varepsilon \|X\| \quad \forall x \in \mathcal{X}, \quad X = \mathcal{Z}x.$$

Here $y(\cdot)$ is the same as above, $\widehat{y}(t) = \sum_{s=-\infty}^{t} \widehat{k}(t-s)x(s)$.

We call functions $\widehat{k}(\cdot)$ in Definition 3 predictors or predicting kernels.

3 The main result

Let $\Omega \in (0, \pi)$ be given, and let

$$\mathcal{X}_L = \{x(\cdot) \in \ell_2 : \ X\left(e^{i\omega}\right) = 0 \quad \text{if} \quad |\omega| > \Omega, \quad X = \mathcal{Z}x\},$$
$$\mathcal{X}_H = \{x(\cdot) \in \ell_2 : \ X\left(e^{i\omega}\right) = 0 \quad \text{if} \quad |\omega| < \Omega, \quad X = \mathcal{Z}x\}.$$

In particular, $\mathcal{X}_L$ is a class of band-limited processes, and $\mathcal{X}_H$ is a class of high-frequency processes.

3.1 Predictability of band-limited and high-frequency processes from L_2

Let $\mathcal{K}_0$ be the class of all functions $k \in \ell_\infty$ such that $k(t) = 0$ for $t > 0$ and that $K = \mathcal{Z}k$ can be represented as

$$K(z) = \frac{z+b}{z+a}, \tag{3.1}$$

for some real $a \in (-\infty, -1) \cup (1, +\infty)$ and $b \in \mathbf{R}$.

Theorem 1 *(i) The classes $\mathcal{X}_L$ and $\mathcal{X}_H$ are ℓ_2-predictable in the weak sense with respect to the class $\mathcal{K}_0$.*

(ii) The classes $\mathcal{X}_L$ and $\mathcal{X}_H$ are ℓ_∞-predictable in the weak sense with respect to the class $\mathcal{K}_0$ uniformly with respect to the norm $\|X(e^{i\omega})\|_{L_2(-\pi,\pi)}$.

(iii) For any $q > 2$, the classes $\mathcal{X}_L$ and $\mathcal{X}_H$ are ℓ_2-predictable in the weak sense with respect to the class $\mathcal{K}_0$ uniformly with respect to the norm $\|X(e^{i\omega})\|_{L_q(-\pi,\pi)}$.

The question arises how to find the predicting kernels. In the proof of Theorem 1, a possible choice of the kernels is given explicitly via Z-transforms.

4 On a model with ideal low pass-pass filter

Corollary 1 *Assume a model with a process $x(\cdot)$ such that it is possible to decompose it as $x(t) = x_L(t) + x_H(t)$, where $x_L(\cdot) \in \mathcal{X}_L$ and $x_H(\cdot) \in \mathcal{X}_H$. Then this observer would be able to predict (approximately, in the sense of weak predictability with respect to the class $\mathcal{K}_0$) the values of $y(t) = \sum_{s=t}^{+\infty} k(t-s)x(s)$ for $k(\cdot) \in \mathcal{K}$ by predicting the processes $y_L(t) = \sum_{s=t}^{+\infty} k(t-s)x_L(s)$ and $y_H(t) = \sum_{s=t}^{+\infty} k(t-s)x_H(s)$ separately. More precisely, the process $\widehat{y}(t) = \widehat{y}_L(t) + \widehat{y}_H(t)$ is the prediction of $y(t)$, where $y_L(t) = \sum_{-\infty}^{t} \widehat{k}_L(t-s)x_L(s)$ and $y_H(t) = \sum_{-\infty}^{t} \widehat{k}_H(t-s)x_H(s)$, and where $\widehat{k}_L(\cdot)$ and $\widehat{k}_H(\cdot)$ are predicting kernels which existence for the processes $x_L(\cdot)$ and $x_H(\cdot)$ is established above.*

Let $\chi_L\left(e^{i\omega}\right) = \mathbb{I}_{\{|\omega|\le\Omega\}}$ and $\chi_H\left(e^{i\omega}\right) = 1 - \chi_L\left(e^{i\omega}\right) = \mathbb{I}_{\{|\omega|>\Omega\}}$, where $\omega \in \mathbf{R}$; $\mathbb{I}$ denote the indicator function.

The assumptions of Corollary 1 mean that there are a low-pass filter and a high-pass filter with the transfer functions χ_L and χ_H respectively, with $x(\cdot)$ as the input, i.e., that the values $x_L(s)$ and $x_H(s)$ for $s \le t$ are available at time t, where

$$x_L(\cdot) = \mathcal{Z}^{-1}X_L, \quad X_L\left(e^{i\omega}\right) = \chi_L\left(e^{i\omega}\right) X\left(e^{i\omega}\right),$$
$$x_H(\cdot) = \mathcal{Z}^{-1}X_H, \quad X_H\left(e^{i\omega}\right) = \chi_H\left(e^{i\omega}\right) X\left(e^{i\omega}\right),$$

and where $X = \mathcal{Z}x$. It follows that the predictability in the weak sense with respect to the class $\mathcal{K}_0$ is possible for any process $x(\cdot)$ that can be decomposed without error on a band limited process and a high-frequency process, i.e., when there is a low-pass filters which behave as an ideal filter for this process. (Since $x_H(t) = x(t) - x_L(t)$, existence of the low pass filter implies existence of the high pass filter). On the other hand, Corollary 1 implies that the existence of ideal low-pass filters is impossible for general processes, since they cannot be predictable in the sense of Definition 3.

Clearly, processes $x(\cdot) \in \mathcal{X}_L \cup \mathcal{X}_H$ are automatically covered by Corollary 1, i.e., the existence of the filters is not required for this case. For instance, we have immediately that $x_L(\cdot) = x(\cdot)$ and $x_H(\cdot) \equiv 0$ for band-limited processes.

5 Proofs

It suffices to present a set of predicting kernels $\widehat{k}$ with the desired properties. We will use a version of the construction introduced in Dokuchaev (2008) for continuous time setting. This construction is very straightforward and does not use the advanced theory of H^p-spaces.

Let $\mathcal{K}_1$ be the class of all functions $k \in \mathcal{K}_0$ such that $K = \mathcal{Z}k$ can be represented as

$$K(z) = \frac{1}{z+a}, \tag{5.1}$$

for some real $a \in (-\infty, -1) \cup (1, +\infty)$.

If $k \in \mathcal{K}_0$, then $K = \mathcal{Z}k$ can be represented as

$$K(z) = \frac{z+b}{z+a} = \frac{z+a+b-a}{z+a} = 1 + \frac{c}{z+a},$$

with $a \in (-\infty) \cup (1, +\infty)$, $b \in \mathbf{R}$, and $c = b - a$. It follows that the process $y(t)$ for $k \in \mathcal{K}_0$ can be represented as $y(t) = x(t) + c\sum_{s=t}^{+\infty} k_1(t-s)x(s)$, where $k_1 \in \mathcal{K}_1$. Therefore, it suffices to prove theorem for $k \in \mathcal{K}_1$ only.

Let $k(\cdot) \in \mathcal{K}_1$ and $K\left(e^{i\omega}\right) = \mathcal{Z}k$ be defined by (5.1) for some for $a \in (-\infty, -1) \cup (1, +\infty)$.

Let $G = (-\Omega, \Omega)$, and let

$$\alpha = -\frac{1 + a\cos(\Omega)}{a + \cos(\Omega)}. \tag{5.2}$$

Let us show that $\alpha = f(a) \in (-1, 1)$. Clearly, the function

$$f(a) = \frac{1 + a\cos(\Omega)}{a + \cos(\Omega)}$$

is such that $f'(a) < 0$ for all a such that $|a| \geq 1$, $f(-1) = -1$, $f(1) = 1$, and $f(\pm\infty) = \cos(\Omega)$. These properties imply that $\alpha = f(a) \in (-1, 1)$.

Further, we have that $1 + \alpha a + (a + \alpha)\cos(\Omega) = 0$, and

$$\begin{aligned} &\text{sign}\,(a + \alpha)(1 + \alpha a + (a + \alpha)\cos(\omega)) > 0, \quad \omega \in G, \\ &\text{sign}\,(a + \alpha)(1 + \alpha a + (a + \alpha)\cos(\omega)) < 0, \quad \omega \in (-\Omega, \Omega)\backslash G. \end{aligned} \tag{5.3}$$

Set

$$V(z) = 1 - \exp\left(\gamma \text{sign}\,(a + \alpha)\frac{z + a}{z + \alpha}\right), \quad \widehat{K}(z) = V(z)K(z), \quad \gamma \in \mathbf{R}. \tag{5.4}$$

Lemma 1 *(i) $V(z) \in H^\infty$ and $\widehat{K}(z) = K(z)V(z) \in H^\infty \cap H^2$.*

(ii) If $\gamma < 0$ and $\omega \in [-\Omega, \Omega]$, then $|V(e^{i\omega})| \le 2$. If $\gamma > 0$ and $\omega \in [-\pi, \pi]\backslash(-\Omega, \Omega)$, then $|V(e^{i\omega})| \le 2$.

(iii) If $\omega \in (-\Omega, \Omega)$, then $V(e^{i\omega}) \to 1$ as $\gamma \to -\infty$. If $\omega \in [-\pi, \pi]\backslash[-\Omega, \Omega]$, then $V(e^{i\omega}) \to 1$ as $\gamma \to +\infty$.

(iv) For any $\varepsilon \in (0, \Omega)$, $V(e^{i\omega}) \to 1$ as $\gamma \to -\infty$ uniformly in $\omega \in [-\Omega + \varepsilon, \Omega - \varepsilon]$ as $\gamma \to -\infty$, and $V(e^{i\omega}) \to 1$ as $\gamma \to +\infty$ uniformly in $\omega \in [-\pi, \pi]\backslash(-\Omega + \varepsilon, \Omega - \varepsilon)$.

Proof of Lemma 1. Clearly, $V \in H^\infty$, and $(z + a)^{-1}V(z) \in H^2 \cap H^\infty$, since the pole of $(z + a)^{-1}$ is being compensated by multiplying with V. It follows that $K(z)V(z) \in H^2 \cap H^\infty$. Then statement (i) follows.

Further, for $\omega \in \mathbf{R}$,

$$\frac{e^{i\omega} + a}{e^{i\omega} + \alpha} = \frac{(e^{i\omega} + a)(e^{-i\omega} + \alpha)}{|e^{i\omega} + \alpha|^2} = \frac{1 + a\alpha + ae^{-i\omega} + \alpha e^{i\omega}}{|e^{i\omega} + \alpha|^2}.$$

Hence

$$\text{Re}\,\frac{e^{i\omega} + a}{e^{i\omega} + \alpha} = \frac{1 + a\alpha + (a + \alpha)\cos(\omega)}{|e^{i\omega} + \alpha|^2}.$$

Then statements (ii)-(iv) follow from (5.3). This completes the proof of Lemma 1. □

Proof of Theorem 1. For $x(\cdot) \in \ell_2$, let $X = \mathcal{Z}x$, $k = \mathcal{Z}^{-1}K$, $\widehat{k} = \mathcal{Z}^{-1}\widehat{K}$,

$$y(t) = \sum_{s=t}^{\infty} k(t - s)x(s), \quad \widehat{y}(t) = \sum_{s=-\infty}^{t} \widehat{k}(t - s)x(s).$$

Let $Y = \mathcal{Z}y$, let V and $\widehat{K}$ be as defined above, and let $\widehat{Y} = \widehat{K}X$.

Let us consider the cases of $\mathcal{X}_L$ and $\mathcal{X}_H$ simultaneously. For the case of the class $\mathcal{X}_L$, consider $\gamma < 0$ and assume that $\gamma \to -\infty$. Set $\Gamma = [-\Omega, \Omega]$ for this case. For the case of the class $\mathcal{X}_H$, consider $\gamma > 0$ and $\gamma \to +\infty$. Set $\Gamma = [-\pi, -\Omega] \cup [\Omega, +\pi]$ for this case.

Let $x(\cdot) \in \mathcal{X}_L$ or $x(\cdot) \in \mathcal{X}_H$. In both cases, Lemma 1 gives that $|V\left(e^{i\omega}\right)| \le 2$ for all $\omega \in \Gamma$. If $\gamma \to -\infty$ or $\gamma \to +\infty$ respectively for $\mathcal{X}_L$ or $\mathcal{X}_H$ cases, then $V\left(e^{i\omega}\right) \to 1$ for a.e. $\omega \in \Gamma$, i.e., for a.e. ω such that $X\left(e^{i\omega}\right) \neq 0$.

Let us prove (i). Since $K\left(e^{i\omega}\right) \in L_\infty(-\pi,\pi)$, $\widehat{K}\left(e^{i\omega}\right) \in L_\infty(-\pi,\pi)$, and $X\left(e^{i\omega}\right) \in L_2(-\pi,\pi)$, we have that $Y\left(e^{i\omega}\right) = K\left(e^{i\omega}\right)X\left(e^{i\omega}\right) \in L_2(-\pi,\pi)$ and $\widehat{Y}\left(e^{i\omega}\right) = \widehat{K}\left(e^{i\omega}\right)X\left(e^{i\omega}\right) \in L_2(-\pi,\pi)$. By Lemma 1, it follows that

$$\widehat{Y}\left(e^{i\omega}\right) \to Y\left(e^{i\omega}\right) \quad \text{for a.e.} \quad \omega \in \mathbf{R}, \tag{5.5}$$

as $\gamma \to -\infty$ or $\gamma \to +\infty$ respectively for $\mathcal{X}_L$ or $\mathcal{X}_H$ cases. We have that

$$|\widehat{K}\left(e^{i\omega}\right) - K\left(e^{i\omega}\right)| \le |V\left(e^{i\omega}\right) - 1||K\left(e^{i\omega}\right)| \le 2|K_m\left(e^{i\omega}\right)|, \quad \omega \in \Gamma, \tag{5.6}$$

$$|\widehat{Y}\left(e^{i\omega}\right) - Y\left(e^{i\omega}\right)| \le 2|Y\left(e^{i\omega}\right)| = 2|K\left(e^{i\omega}\right)||X\left(e^{i\omega}\right)|, \quad \omega \in \Gamma. \tag{5.7}$$

By (5.5),(5.7), and by Lebesque Dominance Theorem, it follows that

$$\|\widehat{Y}\left(e^{i\omega}\right) - Y\left(e^{i\omega}\right)\|_{L_2(-\pi,\pi)} \to 0, \quad \text{i.e.,} \quad \|\widehat{y} - y\|_{L_2(-\pi,\pi)} \to 0 \tag{5.8}$$

as $\gamma \to -\infty$ or $\gamma \to +\infty$ respectively for $\mathcal{X}_L$ or $\mathcal{X}_H$ cases, where $\widehat{y} = \mathcal{Z}^{-1}\widehat{Y}$.

Let us prove (ii)-(iii). Take $d = 1$ for (ii) and take $d = 2$ for (iii). If $X\left(e^{i\omega}\right) \in L_\nu(-\pi,\pi)$ for $\nu > d$, then Hölder inequality gives

$$\|\widehat{Y}\left(e^{i\omega}\right) - Y\left(e^{i\omega}\right)\|_{L_d(-\pi,\pi)} \le \|\widehat{K}\left(e^{i\omega}\right) - K\left(e^{i\omega}\right)\|_{L_\mu(\Gamma)}\|X\left(e^{i\omega}\right)\|_{L_\nu(\Gamma)}, \tag{5.9}$$

where μ is such that $1/\mu + 1/\nu = 1/d$. By (5.6) and by Lebesque Dominance Theorem again, it follows that

$$\|\widehat{K}\left(e^{i\omega}\right) - K\left(e^{i\omega}\right)\|_{L_\mu(\Gamma)} \to 0 \quad \forall \mu \in [1, +\infty), \tag{5.10}$$

as $\gamma \to -\infty$ or $\gamma \to +\infty$ respectively for $\mathcal{X}_L$ or $\mathcal{X}_H$ cases. Then, by (5.9)-(5.10), it follows that the predicting kernels $\widehat{k}(\cdot) = \widehat{k}(\cdot,\gamma) = \mathcal{Z}^{-1}\widehat{K}$ are such as required in statements (ii)–(iii). This completes the proof of Theorem 1. $\square$

Corollary 1 follows immediately from Theorem 1.

6 On the prediction error generated by a high-frequency noise

Let us estimate the prediction error for the case when predictor (5.4) designed for a band-limited process is applied to a process with a small high-frequency noise.

Let $\Omega \in (0,\pi)$ and $\nu \in [0,1)$ be given. Let us consider a process $x(\cdot) \in \ell_\infty$ such that $|X(i\omega)| \le 1$ for $\omega \in G$ and $|X(i\omega)| \le \nu$ for $\omega \in [-\pi,\pi]\backslash G$, where $X = \mathcal{Z}x$ and $G = (-\Omega,\Omega)$.

Assume that predictor (5.4) is constructed under the hypothesis that $\nu = 0$ (i.e, that $x(\cdot)$ is a band-limited processes from $\mathcal{X}_L$), for some $a \in \mathbf{R}\backslash[-1,1]$. For an arbitrarily small $\varepsilon > 0$, we can find $\gamma = \gamma(\varepsilon)$ such that if the hypothesis that $\nu = 0$ is correct, then

$$\|\widehat{y} - y\|_{\ell_\infty} \le \frac{\varepsilon}{2\pi}, \tag{6.1}$$

where $y(\cdot)$ and $\widehat{y}(\cdot)$ are such as in Definition 3.

Let us estimate the prediction error for the case when $\nu > 0$. We have that

$$\|\widehat{y} - y\|_{\ell_\infty} \le \frac{1}{2\pi}\|\widehat{Y}\left(e^{i\omega}\right) - Y\left(e^{i\omega}\right)\|_{L_1(-\pi,\pi)},$$

where $Y = \mathcal{Z}y$ and $\widehat{Y} = \mathcal{Z}\widehat{y}$. Let $\Omega_1 = \Omega - \varepsilon/4$ and $G_1 = (-\Omega_1, \Omega_1)$. By the assumptions on X, we have that

$$\|\widehat{Y}\left(e^{i\omega}\right) - Y\left(e^{i\omega}\right)\|_{L_1(-\pi,\pi)} \le I_1 + I_2 + \nu I_3,$$

where

$$I_1 = \kappa\int_{G_1} e^{\gamma\psi(\omega)}d\omega, \quad I_2 = \kappa\int_{G\backslash G_1} e^{\gamma\psi(\omega)}d\omega, \quad I_3 = \kappa\int_{(-\pi,\pi)\backslash G} e^{\gamma\psi(\omega)}d\omega,$$

and where $\kappa = \max_\omega ||K(e^{i\omega})|$,

$$\psi(\omega) = \text{sign}\,(a+\alpha)\text{Re}\,\frac{e^{i\omega}+a}{e^{i\omega}+\alpha}.$$

Note that $\psi(\omega) > 0$ for $\omega \in G$. Let $\psi_0 = \min_{\omega\in G_1}\psi(\omega)$, and let $\gamma = -\log(2\kappa/\varepsilon)/\psi_0$. Then $I_1 \le \varepsilon/2$. Further, $I_2 \le \kappa\,\text{mes}\,(G\backslash G_1) = \varepsilon/2$. Therefore, (6.1) holds if $\nu = 0$.

The value $I_1 + I_2$ represents the forecast error when $\nu = 0$; this error can be done arbitrarily small with γ selected as above when $\varepsilon \to 0$.

Let us estimate I_3. Clearly, $|\psi(\omega)| \le 1 + \left|\frac{a-\alpha}{e^{i\omega}+\alpha}\right| \le \mu$, where $\mu = 1 + |a-\alpha|/(1-\alpha)$. Hence

$$I_3 \le \kappa\int_{(-\pi,\pi)\backslash G} e^{\gamma\mu}d\omega = \kappa\int_{(-\pi,\pi)\backslash G} e^{\frac{\log(2\kappa/\varepsilon)}{\psi_0}\mu}d\omega = 2\kappa(\pi-\Omega)e^{\frac{\log(2\kappa/\varepsilon)\mu}{\psi_0}}.$$

Hence

$$\nu I_3 \le 2\kappa\nu(\pi-\Omega)\left(\frac{2\kappa}{\varepsilon}\right)^{\frac{\mu}{\psi_0}}.$$

The value νI_3 represents the additional error caused by the presence of unexpected high-frequency noise (when $\nu > 0$). It can be seen that if $\varepsilon \to 0$ than this error is increasing as a polynomial of ε^{-1} with the rate depending on α (defined by Ω and a). If $\Omega \to \pi$ then $|\alpha| \to 1$ and $\mu \to +\infty$, and, for a given ε, the error is increasing exponentially in μ.

7 Concluding remarks

- By (5.2), $\alpha \to \pm 1$ as $\Omega \to \pi$, and the predictor suggested above loses its feasibility as $\Omega \to \pi$. (In particular, $\|\widehat{k}\|_{\ell_\infty} \to +\infty$).

- If $k(\cdot)$ is a real valued function, then $\widehat{k}$ is also real valued. It follows from the fact that $K(\bar{z}) = \overline{K(z)}$, and, therefore, $K\left(e^{-i\omega}\right) = \overline{K\left(e^{i\omega}\right)}$.

- A similar approach can be applied to the case when $X(z)$ vanishes on some connected set $I \subset \mathbb{T}$. In this case, the classes K_0 and K_1 have to be replaced by similar classes with complex $a \in D^c$. For real valued kernels, it could be meaningful to include the functions K represented by the sums of two simple fractions, to ensure that the process $\mathcal{Z}^{-1}k$ is real (i.e, that $\overline{K\left(e^{i\omega}\right)} = K\left(e^{-i\omega}\right)$).

- The predictors obtained above require the past values of $x(s)$ for all $s \in (-\infty, t]$. In practice, $\sum_{s=-\infty}^{t} \widehat{k}(t-s)x(s)$ can be approximated by $\sum_{s=-M}^{t} \widehat{k}(t-s)x(s)$ for large enough $M > 0$. In addition, the corresponding transfer functions can be approximated by rational fraction polynomials.

- The system for the suggested predictors is stable, since the corresponding transfer functions have poles in the domain $\{|z| < 1\}$ only. However, the suggested predictors are not robust. For instance, if the predictor is designed for the class $\mathcal{X}_L$ and it is applied for a process $x(\cdot) \notin \mathcal{X}_L$ with small non-zero energy at the frequencies outside $[-\Omega, \Omega]$, then the error generated by the presence of this energy is increasing if $\gamma \to \infty$.

- The results of this paper can be applied to discrete time stationary random Gaussian processes. In particular, assume that the spectral density of the underlying process $x(t)$ vanishes outside the interval $[-\Omega, \Omega] \subset (-\pi, \pi)$. It is known that the minimal (optimal) predicting error is zero in this case. The sequence of the predictors constructed above represents a sequence of suboptimal predictors leading to vanishing prediction error.

Acknowledgments

This work was supported by ARC grant of Australia DP120100928 to the author. In addition, the author thanks Prof. Augusto Ferrante for useful discussion and advice on spectral analysis of time series.

References

[1] Dokuchaev, N. (2008). The predictability of band-limited, high-frequency, and mixed processes in the presence of ideal low-pass filters. *Journal of Physics A: Mathematical and Theoretical* **41** No 38, 382002 (7pp).

[2] Dokuchaev, N. (2010). Predictability on finite horizon for processes with exponential decrease of energy on higher frequencies. *Signal Processing* **90** Iss. 2, 696–701.

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

[4] Knab J.J. (1979). Interpolation of band-limited functions using the approximate prolate series. *IEEE Transactions on Information Theory* **25**(6), 717–720.

[5] Lyman R.J, Edmonson W.W., McCullough S., and Rao M. (2000). The predictability of continuous-time, bandlimited processes. *IEEE Transactions on Signal Processing* **48**, Iss. 2, 311–316.

[6] Lyman R.J and Edmonson W.W. (2001). Linear prediction of bandlimited processes with flat spectral densities. *IEEE Transactions on Signal Processing* **49**, Iss. 7, 1564–1569.

[7] Marvasti F. (1986). Comments on "A note on the predictability of band-limited processes." *Proceedings of the IEEE*, **74**(11), 1596.

[8] Papoulis A. (1985). A note on the predictability of band-limited processes. *Proceedings of the IEEE*, **73**(8), 1332–1333.

[9] Vaidyanathan P.P. (1987). On predicting a band-limited signal based on past sample values. *Proceedings of the IEEE*, **75**(8), 1125–1127.

[10] Wainstein L.A. and Zubakov V.D. (1962). *Extraction of Signals from Noise.* Englewood Cliffs, NJ: Prentice-Hall.