On detecting and quantification of randomness for one-sided sequences

Nikolai Dokuchaev
Department of Mathematics & Statistics, Curtin University,
GPO Box U1987, Perth, 6845 Western Australia
January 22, 2016

Abstract

The paper studies discrete time processes and their predictability and randomness in deterministic pathwise setting, without using probabilistic assumptions on the ensemble. We suggest some frequency based criterions of predictability based on bandlimitness and some approaches to quantification of randomness based on frequency analysis of two-sided and one-sided sequences. We develop procedure allowing to represent an one-sided sequences as a sum of left-bandlimited and predictable sequences and a non-reducible noise.

Key words: causal approximation, sequences, one-sided sequences, randomness, bandlimitiness, prediction forecasting

1 Introduction

The paper studies discrete time processes and their predictability and randomness in deterministic pathwise setting, without using probabilistic assumptions on the ensemble.

Understanding of the pathwise randomness leads to many applications in Monte-Carlo methods, cryptography, and control systems. There are many classical works devoted to the concept of pathwise randomness and the problem of distinguishability of random sequences; see the references in [20, 14]. In particular, the approach from Borel (1909) [3], Mises (1919) [22], Church (1940) [6] was based on limits of the sampling proportions of zeros in the binary sequences and subsequences; Kolmogorov (1965) [18] and Loveland (1966) [21] developed a different concept of the algorithmic randomness and compressibility; Schnorr (1971) [23] suggested approach based on predicability and martingale properties. So far, the exiting theory is devoted to the problem of distinguishability of random sequences and does not consider the problem of quantification of the degree of randomness. This paper studies randomness in the sense of the pathwise predicability and attempts to develop
an approach for quantification and separation of the “noise” for the sequences that are deemed to be random. The estimation of the degree on randomness is a difficult problem, since the task of detecting the randomness is nontrivial itself.

We will investigate randomness and noise for the sequences in a more special setting originated from the linear filtering and prediction of stochastic processes rather than algorithmic randomness in the spirit of Downey (2004). We suggest exploring the following straightforward pathwise criterion: a class of sequences predictable in the sense of Lemma 1 is assumed to consist of non-random sequences.

For stationary discrete time processes, there is a criterion of predictability in the frequency domain setting given by the classical Szegö-Kolmogorov theorem; see [17, 25, 28] and recent literature reviews in [4, 24]. By this theorem, a stationary process is predictable if its spectral density is vanishing with a certain rate at a point the unit circle \( \{z \in \mathbb{C} : |z| = 1\} \). In particular, it holds if the spectral density vanishes on an arc of the unit circle, i.e., the process is bandlimited. There are many works devoted to smoothing in frequency domain and sampling; see, e.g., [1, 2, 15, 16, 27, 29] and the bibliography here. In [8, 10, 11, 12, 13], predictability was readressed in the deterministic setting for two-sided sequences for with Z-transform vanishing in a point on \( \mathbb{T} \), and some predictors were suggested. These results were based on frequency characteristics of the entire two-sided sequences, since the properties of the Z-transforms were used. Application of the two-sided Z-transform requires to select some past time at the middle of the time interval of the observations as the zero point for a model of the two-sided sequence; this could be inconvenient. In many applications, it is more convenient to represent data flow as one-sided sequences such that \( x(t) \) represents outdated observations with diminishing significance as \( t \to -\infty \). This leads to the analysis of the one-sided sequences directed backward to the past. However, the straightforward application of the one-sided Z-transform to the one-sided sequences does not generate Z-transform vanishing on a part of the unit circle even for a band-limited underlying sequence.

The paper suggests an extension of the of bandlimitiness on one-sided sequences and some approaches to quantification of randomness based on frequency analysis of two-sided and one-sided sequences. In addition, the paper suggests a procedure allowing to represent an one-sided sequences as a sum of left-bandlimited and predictable sequences and a non-reducible noise.

2 Definitions and background

We use notation \( \text{sinc}(x) = \sin(x)/x \) and \( \mathbb{T} = \{z \in \mathbb{C} : |z| = 1\} \), and we denote by \( \mathbb{Z} \) the set of all integers.

For a Hilbert space \( \mathcal{H} \), we denote by \( \langle \cdot, \cdot \rangle_\mathcal{H} \) the corresponding inner product. We denote by
Let \( \ell_r \) be the set of all sequences \( x = \{x(t)\}_{t \in \mathbb{Z}} \subset \mathbb{C} \), such that
\[
\|x\|_{\ell_r} = \left( \sum_{t=-\infty}^{\infty} |x(t)|^r \right)^{1/r} < +\infty \quad \text{for } r \in [1, \infty) \quad \text{or} \quad |x|_{\ell_\infty} = \sup_t |x(t)| < +\infty \quad \text{for } r = +\infty.
\]

For \( x \in \ell_1 \) or \( x \in \ell_2 \), we denote by \( X = \mathcal{Z}x \) the Z-transform
\[
X(z) = \sum_{t=-\infty}^{\infty} x(t) z^{-t}, \quad z \in \mathbb{T}.
\]

Respectively, the inverse Z-transform \( x = \mathcal{Z}^{-1}X \) is defined as
\[
x(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\epsilon t) e^{i\omega t} d\omega, \quad t = 0, \pm 1, \pm 2, \ldots
\]

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

Let \( \tau \in \mathbb{Z} \cup \{+\infty\} \) and \( \theta < \tau \); the case where \( \theta = -\infty \) is not excluded. We denote by \( \ell_2(\theta, \tau) \) the Hilbert space of complex valued sequences \( \{x(t)\}_{t=\theta}^{\tau} \) such that
\[
\|x\|_{\ell_2(\theta, \tau)} = \left( \sum_{t=\theta}^{\tau} |x(t)|^2 \right)^{1/2} < +\infty.
\]

Let \( \ell_2 = \ell_2(-\infty, +\infty) \) and \( \ell_2^- = \ell_2(-\infty, 0) \).

For a set \( I \subset (-\pi, \pi) \), we denote \( I^c = (-\pi, \pi) \setminus I \).

Let \( J \) be the set of all \( I \subset (-\pi, \pi] \) such that the set \( \{e^{i\omega}\}_{\omega \in I} \) is a connected arc and \( I^c \neq \emptyset \).

For any \( I \in J \), we denote by \( \omega_I \) the middle point \( e^{i\omega_I} \) of the arc \( \{e^{i\omega}\}_{\omega \in I} \).

We denote by \( \mathcal{X}_{BL}(I) \) the set of all mappings \( X : \mathbb{T} \to \mathbb{C} \) such that \( X(e^{i\omega}) \in L_2(-\pi, \pi) \) and \( X(e^{i\omega}) = 0 \) for \( \omega \notin I \). We will call the corresponding processes \( x = \mathcal{Z}^{-1}X \) band-limited. Let \( \ell_2^{BL}(I) = \{x \in \ell_2 : X = \mathcal{Z}x \in \mathcal{X}_{BL}(I)\} \).

**Definition 1.** Assume that there exists \( I \in J \) such that \( x \in \ell_2^{BL}(I) \) represents the trace of a band-limited process \( x_{BL} \in \ell_2 \) with the spectrum on \( I \), i.e., \( x(t) = x_{BL}(t) \) for \( t \leq 0 \), and \( X_{BL} = \mathcal{Z}x_{BL} \in \mathcal{X}_{BL}(I) \). We call the process \( x \) **left band-limited** with the spectrum on \( I \).

Let \( \ell_2^{BL}(I) \) be the subset of \( \ell_2^- \) consisting of semi-infinite sequences \( \{x(t)\}_{t \leq 0} \), where \( x \in \mathcal{X} \) are such that \( x(t) = (\mathcal{Z}^{-1}X)(t) \) for \( t \leq 0 \) for some \( X \in \mathcal{X}_{BL}(I) \).

**Some predictability results for two-sided sequences**

Two-sided band-limited sequences are predicable in the following sense.

**Theorem 1.** (i) Let \( \mathcal{X} \subset \ell_2^{BL}(I) \) be a bounded set. Then, for any \( \varepsilon > 0 \), there exists a mapping \( \tilde{k}(\cdot) : \mathbb{Z} \to \mathbb{R} \) such that \( \sup_{t \in \mathbb{Z}} \|x(t) - \tilde{x}(t)\| \leq \varepsilon \) for all \( x \in \mathcal{X} \) for \( \tilde{x}(t) \overset{\Delta}{=} e^{i\omega t} \sum_{s=t-1}^{\infty} \tilde{k}(t-s) e^{-i\omega t} x(s) \).

(ii) Let \( J_1 \subset J \) be a set of \( I \) such that \( \sup_{t \in J_1} \text{mes}(I) < 2\pi \). Let \( \mathcal{X} \subset \bigcup_{I \in J_1} \ell_2^{BL}(I) \) be a bounded set in \( \ell_2 \). Then, for any \( \varepsilon > 0 \), there exists a mapping \( \tilde{k}(\cdot) : \mathbb{Z} \to \mathbb{R} \) such that \( \sup_{t \in \mathbb{Z}} \|x(t) - \tilde{x}(t)\| \leq \varepsilon \) for all \( x \in \mathcal{X} \) for \( \tilde{x}(t) \overset{\Delta}{=} e^{i\omega t} \sum_{s=t-1}^{\infty} \tilde{k}(t-s) e^{-i\omega t} x(s) \).
(iii) Let \( J_1 \) be the set of \( I \in \mathcal{J} \) such that \( \sup_{I \in J} \operatorname{mes}(I) < 2\pi \), and \( \mathcal{X} \subset \bigcup_{I \in J_1} \ell^BL_2(I) \) be a bounded set in \( \ell_2 \) such that \( \sum_{t \leq \tau} |x(t)|^2 \rightarrow 0 \) as \( \tau \rightarrow +\infty \) uniformly over \( x \in \mathcal{X} \). Then, for any \( \varepsilon > 0 \), there exists \( \tau < 0 \) and a mapping \( \hat{k}(\cdot): \mathbb{Z} \rightarrow \mathbb{R} \) such that \( \sup_{t \geq 1} \|x(t) - \hat{x}(t)\| \leq \varepsilon \) for all \( x \in \mathcal{X} \) for \( \hat{x}(t) \triangleq e^{i\omega t} \sum_{s=\tau}^{t-1} \hat{k}(t-s)e^{-i\omega s}x(s) \).

Theorem \( \text{(iii)} \) states that some predicability based on finite sets of observations also can be achieved if we relax predicability requirement to cover times \( t \geq 1 \) only; this would be a weaker version of predicability comparing with the one described in Lemma \( \text{(ii)} \).

Some versions of this Theorem and some examples of predictable classes can be found in \( [9, 10] \).

In addition, it appears that the spectrum supporting sets \( I \) can be estimated from the set of observations \( \{x(s)\}_{s \leq \tau} \) for any \( \tau < 0 \). More precisely, the following theorem holds.

**Theorem 2.** Let \( \mathcal{X} \subset \ell_2 \) be a set such that each \( x \in \ell^BL_2(I) \) for some \( I = I(x) \in \mathcal{J} \), and that \( \nu \triangleq 2\pi - \sup_{x \in \mathcal{X}} \operatorname{mes}(I(x)) > 0 \). Let \( \tilde{\nu} = \nu/3 \). Then, for any \( \tau < 0 \), there exists a mapping \( F: \ell_2(-\infty, \tau) \rightarrow (-\pi, \pi) \) such that the following holds.

(i) \( \{e^{i(\omega + \pi)}, \omega \in I \} \subset \{e^{i\omega}, \omega \in I^c \} \) for \( \omega_I = F(x(t)|_{t \leq \tau}) \),

(ii) If \( x \in \mathcal{X} \) then \( x - e^{i\omega} \in I^c \), and \( \hat{I} \subset I \), where

\[
\hat{I} = \left\{ \omega \in (-\pi, \pi): e^{i\omega} \in \{e^{i(\omega + \pi)}, |\omega - \omega_I| \leq \tilde{\nu}\} \right\}.
\]

The set \( \hat{I} \) can be regarded as an estimate of \( I \) based on observations of \( \{x(t)\}_{t \leq \tau} \).

### 3 Quantification of randomness for two-sided sequences

Let us discuss first a straightforward approach where noise is associated with the high-frequency component. Consider a sequence \( x \in \ell_2 \) that does not feature predicability described in Lemma \( \text{(ii)} \). Let

\[
X = Z X, \quad \hat{X}_{BL} = (I_1 X)|_T, \quad y_{BL} = Z^{-1}\hat{X}
\]

for some given \( I \in \mathcal{J} \). Here \( I_1 \) is the indicator function, i.e., \( I_1(e^{i\omega}) = 1 \) if \( \omega \in I \) and \( I_1(e^{i\omega}) = 0 \) if \( \omega \in I^c = (-\pi, \pi)|I \). In many applications, it is acceptable to deem the process \( n_{BL}(t) = x(t) - y_{BL}(t) \) with \( I = I_0 = (-\Omega, \Omega) \), where \( \Omega \in (0, \pi) \), to be a noise accompanying the systematic movement \( \hat{x}(t) \). However, estimation of \( n_{BL}(t) \) will not help to quantify randomness of \( x \), since

\[
n_{BL} = x - y_{BL} = Z^{-1}(I_{I^c} Z x) \in \ell^BL_2(I^c),
\]
i.e., it is also a predictable band-limited process. In particular, this means that any two-sided sequence $x \in \ell_2$ can be represented as a sum

$$x = y_{BL} + n_{BL}, \quad \hat{x} \in \ell_2^{BL}(I), \quad y \in \ell_2^{BL}(I^c),$$

i.e., as a sum of two two-sided band-limited predictable in the sense of Lemma 1 sequences, and that can be done with any choice of $I$.

In addition, $\|x - y_{BL}\|_{\ell_2} \to 0$ as mes $(-\pi, \pi) \setminus I \to \pi$. This also does not lead to possibility to detect and quantify randomness.

Let us suggest a different approach. Let $X \subset \ell_2 \cap \ell_1$ be a class of processes. We assume that, for $x \in X$ and $X = X x$, for any $m > 0$, the functions $X (e^{i\omega})$ and $|X (e^{i\omega})|^{-1}$ can are differentiable in $\omega \in \mathbb{R}$ and that $\sup_{x \in X} \sup_{\omega \in [-\pi, \pi]} |dX (e^{i\omega}) / d\omega| < +\infty$. This can be without a loss of generality, if it for the purpose of the investigation of the predictability for $x$: it is sufficient to replace $x$ by some fast vanishing processes with the same predictability properties such $x(t)/(1 + |t|^m)$, $m \geq 2$.

Let us consider the case where $\min_{\omega \in [-\pi, \pi]} |X (e^{i\omega})| > 0$ for all $x \in X$ and $X = Z x$. We want to represent each $x \in X$ as

$$x = y_{BL} + n,$$

where $y_{BL}$ is a band-limited predictable process such that the class $\mathcal{Y} = \{y_{BL}\}_{x \in X}$, is predictable in the sense of Lemma 1. In this case, each $n = x - y_{BL}$ is a non-predictable (random) noise. We suggest some restrictions on the choice of $y_{BL}$:

(i) $\|X (e^{i\omega})\|_{L_d(-\pi, \pi)} = \|Y_{BL} (e^{i\omega})\|_{L_d(-\pi, \pi)} + \|N (e^{i\omega})\|_{L_d(-\pi, \pi)}, \quad d = 1, +\infty,$

where $Y_{BL} = Z y_{BL}$ and $N = Z n'$.

(ii) $n$ does not allow a similar representation $n = y'_{BL} + n'$, with a non-random (predictable) non-zero $y'_{BL}$ such that

$$\|N (e^{i\omega})\|_{L_d(-\pi, \pi)} = \|Y' (e^{i\omega})\|_{L_d(-\pi, \pi)} + \|N' (e^{i\omega})\|_{L_d(-\pi, \pi)}, \quad d = 1, +\infty,$

where $Y'_{BL} = Z y'_{BL}$ and $N' = Z n'$.

It appears that $n$ featuring these properties exists in some case exists and can be derived explicitly from $X$. Let us show this.

Let $\omega_I \in [-\pi, \pi]$ be such that $|X (e^{i\omega_I})| = \min_{\omega \in [-\pi, \pi]} |X (e^{i\omega})|$, and let

$$\gamma (e^{i\omega}) = \frac{|X (e^{i\omega_I})|}{|X (e^{i\omega})|}, \quad Y (e^{i\omega}) = [1 - \gamma (e^{i\omega})] X (e^{i\omega}), \quad N (e^{i\omega}) = \gamma (e^{i\omega}) X (e^{i\omega}).$$
It can be seen that
\[ X = Y + N, \quad Y(e^{i\omega}) = 0, \quad |N(e^{i\omega})| \equiv |X(e^{i\omega})| = \text{const}, \]
and (3.2) holds with \( d = 1 \) and \( d = \infty \). By continuity of \( X(e^{i\omega}) \) and \( |X(e^{i\omega})|^{-1} \), the function \( Y(e^{i\omega}) \) is also continuous \( \omega \).

If \( Y(e^{i\omega}) \) vanishes fast enough when \( \omega \to \omega_I \) (see [9]), then \( y = Z^{-1}Y \) is predicable; in this case, the set \( \{n\}_{x \in X} \) can be considered as the set of pathwise noises; therefore, this gives a quantification \( n \) as a norm of \( n \) or \( N \), such as
\[ \|N(e^{i\omega})\|_{L_1(-\pi,\pi)} = \min_{\omega \in [-\pi,\pi]} |X(e^{i\omega})|. \quad (3.4) \]
However, this assumption on \( X \) would be too restrictive if we require that (3.3) leads to \( Y(e^{i\omega}) \) that vanishes fast enough when \( \omega \to \omega_I \) such that \( y_{BL} \) is predicable. To overcome this, we suggest to replace (3.3) by
\[
\begin{align*}
\gamma_\varepsilon(e^{i\omega}) &= 1, \quad \text{if } |e^{i\omega} - e^{i\omega_I}| \leq \varepsilon, \\
\gamma_\varepsilon(e^{i\omega}) &= \frac{|X(e^{i\omega_I})|}{|X(e^{i\omega})|}, \quad \text{if } |e^{i\omega} - e^{i\omega_I}| > \varepsilon, \\
Y_\varepsilon(e^{i\omega}) &= [1 - \gamma_\varepsilon(e^{i\omega})]X(e^{i\omega}), \quad N_\varepsilon(e^{i\omega}) = \gamma_\varepsilon(e^{i\omega})X(e^{i\omega}),
\end{align*}
\]
(3.5)
where \( \varepsilon \to 0 \). In this case,
\[
x = y_\varepsilon + n_\varepsilon, \quad y_\varepsilon = Z^{-1}Y_\varepsilon \in L^2_e(I_\varepsilon), \quad n_\varepsilon = Z^{-1}N_\varepsilon,
\]
(3.6)
where \( I_\varepsilon = \{\omega : |e^{i\omega} - e^{i\omega_I}| \leq \varepsilon\} \),
\[
\begin{align*}
|N_\varepsilon(e^{i\omega})| &= X(e^{i\omega}), & & \text{if } \omega \in I_\varepsilon, \\
|N_\varepsilon(e^{i\omega})| &= |X(e^{i\omega_I})| = \text{const}, & & \text{if } \omega \notin I_\varepsilon,
\end{align*}
\]
(3.7)
We regard \( n_\varepsilon \) as approximation of the noise as \( \varepsilon \to 0^+ \).

To justify this description of the noise, we have to show that the set of band-limited processes \( \{y_\varepsilon\} \) in (3.5)-(3.6) is predicable in some sense. Theorem (1)(i)-(ii) does not ensure predicability of this set, since it assumes that the values \( \omega_I \) are known. This would require to know \( I_\varepsilon \), which is inconsistent for predicability. However, Theorem (2) ensures sufficient estimation of \( I_\varepsilon \) based on observations of \( \{x(t)\}_{t \leq \tau} \). This allows to satisfy conditions of Theorem (1)(iii). Therefore, the set of band-limited processes \( \{y_\varepsilon\} \) in (3.5)-(3.6) is predicable in the sense of Theorem (1)(iii). This predicability covers times \( t \geq 1 \) only; it is a weaker version of predicability comparing with the one described in Lemma (1)(i)-(ii). In other words, we use the set \( \{x(s)\}_{\tau \leq s \leq t} \) for prediction of \( x(t) \), and set set \( \{x(s)\}_{s \leq \tau} \) for estimation of \( \tilde{\omega}(I) \approx \omega(I) \).
Since a norm for \(N_\varepsilon\) is approaching the norm for \(N\), the norm of \(N\) can be used for quantification of the randomness of the two-sided sequences. To estimate the normalized size of the noise, one may use the value

\[
\frac{\|N(e^{i\omega})\|_{L_1(-\pi, \pi)}}{\|X(e^{i\omega})\|_{L_1(-\pi, \pi)}} = \min_{\omega \in [-\pi, \pi]} |X(e^{i\omega})|.
\] (3.8)

The process \(y_\varepsilon = Z^{-1}Y_\varepsilon\) can also be interpreted as an output of a smoothing filter.

In some numerical experiments, we calculated value (3.8) for some historical financial time series (returns of stock prices). The value of (3.8) was found to be in the range of 0.05-0.6 for different stocks.

4 Separating the noise for one-sided sequences

Unfortunately, representation (3.1) does not lead toward a solution of the predicability problem, since it would require to know the entire sequence \(\{x(t)\}_{t=-\infty}^{+\infty}\) to calculate \(X = Zx\).

On the other hand, it is natural to use one-sided sequences interpreted as available past observations for predictability problems. For this, we have to use the notion of left bandlimitness for one-sided sequences. We will use a modification of representation (3.1) that was stated for two-sided sequences.

For this, we have to of representation to two-sided sequences. We suggest to replace the "ideal" projections \(\hat{x}_{BL} = Z^{-1}(I_{I}Zx) \in \ell_2\) for \(x \in \ell_2\) and \(y_{BL} = x - \hat{x}_{BL} = Z^{-1}(I_{I}^{c}Zx)\) by their "optimal" one-sided substitutes.

Uniqueness of the extrapolation for left band-limited processes

**Lemma 1.** For any \(I \in \mathcal{J}\) and any \(x \in \ell_2^{-LBL}(I)\), there exists an unique \(x_{BL} \in \ell_2^{BL}(I)\) such that \(x(t) = x_{BL}(t)\) for \(t \leq 0\).

By Lemma 1, the future values \(x_{BL}(t)|_{t>0}\) of a band-limited process \(x_{BL}\), are uniquely defined by the trace \(x_{BL}(t)|_{t\leq0}\). This statement represent a reformulation in the deterministic setting of the classical Szegö-Kolmogorov Theorem for stationary Gaussian processes [18, 25, 26, 28].

Existence of optimal band-limited approximation

Let \(x \in \ell_2^-\) be a semi-infinite one-sided sequence representing available historical data, and let \(I \in \mathcal{J}\).
Theorem 3. There exists a unique optimal solution $\hat{x}$ of the minimization problem

$$\text{Minimize } \sum_{t=-\infty}^{0} |\hat{x}(t) - x(t)|^2 \text{ over } \hat{x} \in \ell_2^{-BL}(I). \quad (4.1)$$

For the optimal solution of problem (4.1), we will use the notation $\hat{x} = P_Ix$.

By Lemma 1, there exists a unique band-limited process $x_{BL} \in \ell_2^{BL}(I)$ such that $\hat{x}(t)|_{t \leq 0} = x_{BL}(t)|_{t \leq 0}$. This offers a natural way to extrapolate a left band-limited solution $\hat{x} \in \ell_{2}^-$ of problem (4.1) on the future times $t > 0$.

The optimal solution

Let $i \in J$ be given, and let $\text{mes } (I) = 2\Omega$ for some $\Omega \in (0, \pi)$.

Let $I_0 = (-\Omega, \Omega)$, i.e., $\omega_{I_0} = 0$.

For $\omega \in [-\pi, \pi)$, let the operator $p_\omega : \ell_2^- \to \ell_2^-$ be defined as $\bar{x}(t) = e^{i\omega t} \bar{x}(t)$ for $\bar{x} = p_\omega x$.

Let the operator $Q : \ell_2 \to \ell_{2}^{-BL}(I_0)$ be defined as $\hat{x} = Qy = \mathcal{Z}^{-1}\hat{X}$, where

$$\hat{X}(e^{i\omega}) = \sum_{k \in \mathbb{Z}} y_k e^{ik\omega\pi/\Omega} \mathbb{1}_{\{|\omega| \leq \Omega\}}, \quad (4.2)$$

for the corresponding $y = \{y_k\} \in \ell_2$. Similarly to the classical sinc representation, we obtain that

$$\hat{x}(t) = \frac{1}{2\pi} \int_{-\Omega}^{\Omega} \left( \sum_{k \in \mathbb{Z}} y_k e^{ik\omega\pi/\Omega} \right) e^{i\omega t} d\omega = \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} y_k \int_{-\Omega}^{\Omega} e^{ik\omega\pi/\Omega + i\omega t} d\omega$$

$$= \frac{1}{2\pi} \sum_{k \in \mathbb{Z}} y_k \frac{e^{ik\pi + i\Omega t} - e^{-ik\pi - i\Omega t}}{ik\pi/\Omega + it} = \frac{\Omega}{\pi} \sum_{k \in \mathbb{Z}_N} y_k \text{sinc}(k\pi + \Omega t) = (Qy)(t). \quad (4.3)$$

It follows that the $Q : \ell_2 \to \ell_{2}^{-BL}(I_0)$ is actually defined as

$$\hat{x}(t) = (Qy)(t) = \frac{\Omega}{\pi} \sum_{k \in \mathbb{Z}} y_k \text{sinc}(k\pi + \Omega t).$$

Consider the operator $Q^* : \ell_{2}^{-BL}(I_0) \to \ell_2$ being adjoint to the operator $Q : \ell_2 \to \ell_{2}^{-BL}(I_0)$, i.e., such that

$$(Q^*x)_k = \frac{\Omega}{\pi} \sum_{i \in T} \text{sinc}(k\pi + \Omega t)x(t). \quad (4.4)$$

Consider the linear bounded non-negatively defined Hermitian operator $R : \ell_2 \to \ell_2$ defined as

$$R = Q^*Q.$$

Theorem 4. (i) The operator $R : \ell_2 \to \ell_2$ has a bounded inverse operator $R^{-1} : \ell_2 \to \ell_2$. 

8
(ii) Problem (4.1) has a unique solution
\[ \hat{x} = P_I x = p_{\omega I} Q R^{-1} Q^* p_{-\omega I} x. \] (4.5)

**Remark 1.** It can be noted that \( \hat{x} = p_{\omega I} Q Q^+ p_{-\omega I} x \), where \( Q^+ = R^{-1} Q^* : \ell_2 \rightarrow \ell_2 \) is a Moore–Penrose pseudoinverse of the operator \( Q : \ell_2 \rightarrow \ell_2 \).

Let us elaborate equation (4.5). The optimal process \( \hat{x} \) can be expressed as
\[ \hat{x}(t) = e^{i\omega t} \frac{\Omega}{\pi} \sum_{k \in \mathbb{Z}} \hat{y}_k \text{sinc}(k\pi + \Omega t). \]

Here \( \hat{y} = \{\hat{y}_k\}_{k \in \mathbb{Z}} \) is defined as
\[ \hat{y} = R^{-1} Q p_{-\omega I} x. \] (4.6)

The operator \( R \) can be represented via a matrix \( R = \{R_{km}\} \), where \( k, m \in \mathbb{Z} \). In this setting, \((Ry)_k = \sum_{k=-\infty}^{\infty} R_{km} y_m\), and the components of the matrix \( R \) are defined as
\[ R_{km} = \frac{\Omega^2}{\pi^2} \sum_{j=-\infty}^{0} \text{sinc}(m\pi + \Omega j) \text{sinc}(k\pi + \Omega j). \]

Respectively, the components of the vector \( Q^* x = \{(Q^* x)_k\}_{k \in \mathbb{Z}} \) are defined as
\[ (Q^* x)_k = \frac{\Omega}{\pi} \sum_{j=-\infty}^{0} \text{sinc}(k\pi + \Omega j) x(j). \] (4.7)

### 4.1 A multi-step procedure for one-sided sequences

Unfortunately, the approach described in Section 3 does not lead toward a solution of the predictability problem, since it would require to know the entire sequence \( \{x(t)\}_{t=\infty}^{\infty} \) to calculate \( X = \mathcal{Z} x \) and quantitative characteristics suggested in Section 3.

On the other hand, it is natural to use one-sided sequences interpreted as available past observations for predictability problems. For this, we have to use the notion of left bandlimitness for one-sided sequences. We will use a modification of representation (3.1) that was stated for two-sided sequences.

For this, we have to of representation to two-sided sequences. We suggest to replace the "ideal" projections \( \hat{x}_{BL} = \mathcal{Z}^{-1}(\mathbb{I}_{\mathcal{L}} Z x) \in \ell_2 \) for \( x \in \ell + 2 \) and \( y_{BL} = x - \hat{x}_{BL} = \mathcal{Z}^{-1}(\mathbb{I}_{\mathcal{L}} Z x) \) by their "optimal" one-sided substitutes \( \hat{x} = P_I x \in \ell_2 \) and \( \hat{y} = P_I (x - \hat{x}) \in \ell_2 \); this substitution is optimal on \( \{t \leq 0\} \) in the sense of optimization problem (4.1). Unfortunately, it may happen that
\[ \hat{y} = P_I (x - \hat{x}) \notin B(I^c). \]

We suggest the following multi-step procedure that to deal with this complication.

Assume that we observe a semi-infinite one-sided sequence \( \{x(t)\}_{t \leq 0} \in \ell_2 \).
Lemma 2. For any $I \in \mathcal{J}$ and $x \in \ell_2^-$, the following holds:

(i) $\|x\| \geq \|x - P_I x\|$, and

(ii) $\|x\| = \|x - P_I x\|$ if and only if $P_I x = 0$.

Let us consider some similar definitions for two-sided sequences.

Consider a sequence of sets $\{I_k\}_{k=0,1,2,\ldots} \subset \mathcal{J}$, with the corresponding middle points $\omega_k \in I_k$. Further, let us consider the following set of sequences from $\ell_2^-$:

- $x_0 = x, \quad \hat{x}_0 = P_{I_0} x_0, \quad y_0 = x_0 - \hat{x}_0, \quad \hat{y}_0 = P_{I_0^c} y_0, \quad x_1 = y_0 - \hat{y}_0$.

- For $k \geq 1$, set
  \[ \hat{x}_k = P_{I_k} x_k, \quad y_k = x_k - \hat{x}_k, \quad \hat{y}_k = P_{I_k^c} y_k, \quad x_{k+1} = y_k - \hat{y}_k. \]

Stopping upon arriving at a predicable process

If there exists $k \geq 0$ such that $y_k = 0$ then

\[ x = \hat{x}_0 + y_0 = \hat{x}_0 + \hat{y}_0 + x_1 + y_1 + \ldots = \hat{x}_0 + \hat{y}_0 + \hat{x}_1 + \hat{y}_1 + \ldots + \hat{x}_k. \] \hspace{1cm} (4.8)

This means that $x$ is a finite sum of left band-limited processes. These processes were calculated by the observer, and, in this sense, each of them can be deemed to be observed, with known (pre-selected) $I_k$; in particular, $x$ can be predicted without error. Similarly, if there exists $k \geq 0$ that $x_{k+1} = 0$, then

\[ x = \hat{x}_0 + y_0 = \hat{x}_0 + \hat{y}_0 + x_1 + y_1 + \ldots = \hat{x}_0 + \hat{y}_0 + \hat{x}_1 + \hat{y}_1 + \ldots + \hat{y}_k. \] \hspace{1cm} (4.9)

This means that $x$ again is a finite sum of observed left band-limited processes. Again, $x$ can be predicted without error.

The norms $\|\eta_k\|_{\ell_2^-}$ and $\|\bar{\eta}_k\|_{\ell_2^-}$ can be used for quantification of the randomness of one-sided semi-infinite sequences.

The case of never stopping procedure

It may happen that, for any $N > 0$, there exists $k \geq N$ such that either $\|y_k\|_{\ell_2^-} + \|x_k\|_{\ell_2^-} > 0$. In this, the randomness can be quantified as

\[ \max \left( \limsup_{k \to +\infty} \|x_k\|_{\ell_2^-}, \limsup_{k \to +\infty} \|y_k\|_{\ell_2^-} \right). \]
Arrival at a non-reducible noise

A process \( x \in \ell_2^+ \) is either left band-limited or not band-limited. Therefore, some processes cannot be represented as a finite sum of left bandlimited processes such as \( (4.8) \) or \( (4.9) \) with a finite \( k \). In this case, the procedure will not be stopped according to the rule described above. It could be beneficial to stop procedure using the following rule.

Let

\[
\delta_k \overset{\Delta}{=} \|x_k\|_{\ell_2^+} - \|x_k - \hat{x}_k\|_{\ell_2^+}, \quad \tilde{\delta}_k \overset{\Delta}{=} \|y_k\|_{\ell_2^-} - \|y_k - \hat{y}_k\|_{\ell_2^-},
\]

i.e., \( \delta_k = \|x_k\|_{\ell_2^+} - \|y_k\|_{\ell_2^-}, \quad \tilde{\delta}_k = \|y_k\|_{\ell_2^-} - \|x_{k+1}\|_{\ell_2^-} \)

\[
\|x_k\|_{\ell_2^+} = \|y_k\|_{\ell_2^-} + \delta_k = \|x_{k+1}\|_{\ell_2^-} + \tilde{\delta}_k, \quad k = 0, 1, ...
\]

\[
\|y_k\|_{\ell_2^-} = \|x_{k+1}\|_{\ell_2^-} + \tilde{\delta}_k = \|y_{k+1}\|_{\ell_2^-} + \delta_k + \tilde{\delta}_k, \quad k = 0, 1, ...
\]

By Lemma 2 it follows that \( \delta_k \geq 0 \) and \( \tilde{\delta}_k \geq 0 \) for all \( k \), i.e.,

\[
\|x_k\|_{\ell_2^+} \geq \|y_k\|_{\ell_2^-} \geq \|x_{k+1}\|_{\ell_2^-}, \quad k = 0, 1, ...
\]

It may happen that \( \delta_k = 0 \), i.e., \( \|x_k\|_{\ell_2^+} = \|y_k\|_{\ell_2^-} \). To save the resources, the procedure should be stopped when this occurs, since further steps will not improve the result. On this step, \( x \) is presented as

\[
x = \hat{x}_0 + \hat{y}_0 + \hat{x}_1 + \hat{y}_1 + ... + \hat{x}_k + y_k = x_p^{(k)} + \eta_k.
\]

where \( x_p^{(k)} = \hat{x}_0 + \hat{y}_0 + \hat{x}_1 + \hat{y}_1 + ... + \hat{x}_k \) is a predictable process since it is a finite sum of observed left band-limited processes, and \( \eta_k = y_k \) is a noise. We are unable to reduce the norm of this noise further given the selected set \( \{I_k\} \). Therefore, we call \( y_k \) a non-reducible noise.

Similarly, it may happen that \( \tilde{\delta}_k = 0 \) and \( \delta_k > 0 \), i.e., \( \|y_k\|_{\ell_2^-} = \|x_{k+1}\|_{\ell_2^-} \). Again, the procedure should be stopped when this occurs, since further steps will not improve the result. This means that the procedure have to stop on the step where \( x \) is presented as

\[
x = \hat{x}_0 + \hat{y}_0 + \hat{x}_1 + \hat{y}_1 + ... + \hat{y}_k + x_{k+1} = y_{\text{BL}}^{(k)} + \bar{\eta}_k.
\]

Here \( y_{\text{BL}}^{(k)} = \hat{x}_0 + \hat{y}_0 + \hat{x}_1 + \hat{y}_1 + ... + \hat{y}_k \) is a predictable process again, and \( \bar{\eta}_k = x_{k+1} \) is a non-reducible noise again.
5 Proofs

The proofs of Theorem 1, Lemma 1, and Theorems 3-4 for the case where \( I_0 = (-\Omega, \Omega) \), i.e., \( \omega_I = 0 \), can be found in [9]. Let us extend these proofs on case where \( \omega_I \neq 0 \).

Let us observe that \( x \in \ell_2^{-,BL}(I) \) and \( X = \mathcal{X}x \in \mathcal{X}_{BL}(I) \) if and only if \( x_0 \triangleq p_{-\omega_I}x \in \ell_2^{-,BL}(I_0) \) and \( X_0 \triangleq Zx_0 \in \mathcal{X}(I_0) \). In this case, \( x = p_{\omega_I}x_0 \), and

\[
X(e^{i\omega}) = \sum_{t=-\infty}^{\infty} x(t)e^{-i\omega t} = \sum_{t=-\infty}^{\infty} x_0(t)e^{-i\omega t}e^{-i\omega t} = X_0 \left( e^{i(\omega_I - \omega)t} \right), \quad \omega \in [0, 2\pi).
\]

Then the proof of Theorem 1(i)-(ii) and Lemma 1 follows.

Proof of Theorem 2 (iii) follows from the robustness of the predictor used in [9] with respect to truncation of inputs from \( \ell_2 \). □

Further, we have that

\[
\| \hat{x} - x \|_{\ell_2^-} = \| p_{-\omega_I} \hat{x} - p_{-\omega_I} x \|_{\ell_2^-} \quad \text{for any} \quad \hat{x}, x \in \ell_2
\]

Hence the problem

\[
\text{Minimize} \quad \| p_{-\omega_I} \hat{x} - p_{-\omega_I} x \|_{\ell_2^-} \quad \text{over} \quad \hat{x} \in \ell_2^{-,BL}(I) \quad (5.1)
\]

has the same sets of solution as problem (4.1). Therefore, there is a bijection between the sets of optimal solutions for problem (4.1) and for the problem

\[
\text{Minimize} \quad \| \hat{y} - y \|_{\ell_2^-} \quad \text{over} \quad \hat{y} \in \ell_2^{-,BL}(I_0),
\]

where \( y = p_{-\omega_I}x \). This bijection has the form \( \hat{y} = p_{-\omega_I} \hat{x} \). Therefore, the proof for \( \omega_I \neq 0 \) follows from the proof for \( \omega_I = 0 \) from [9]. Then the proof of Theorem 3 and Theorem 4 follows. □

Proof of Theorem 2 It is easy to see that there exists a finite set \( \{ \tilde{I}_k \}_{k=1}^M \subset \mathcal{J} \), \( M < +\infty \), such that \( \text{mes}(I_k) \leq \nu/3 \), that \( \cup \tilde{I}_k = (0, 2\pi] \) and that intersections of two different \( I_k \) contain no more than one element, \( k = 1, ..., M \).

Let \( P_I \) be operators such as defined in Section 4 with rather technical adjustment: we assume that the set of times \( \{ t \leq 0 \} \) in Theorem 3 is replaced by the \( \{ t \leq 0 \} \), and that \( \ell_2^- \) replaced by \( \ell_2(-\infty, \tau) \). As is shown in Theorem 4 the values \( d_k \triangleq \| P_{\tilde{I}_k}x - x \|_{\ell_2(-\infty, \tau)} \) for \( k = 1, ..., M \) can be found based on observations of \( \{ x(t) \}_{t \leq \tau} \). By the assumptions on \( x \), there exists \( m \) such that \( d_m = 0 \). The set \( \tilde{I} = \tilde{I}_m \) is such as required. □

Proof of Lemma 2 Statement (i) follows from the choice of \( P_Ix \) as a solution of optimization problem (4.1). To prove statement (ii), it suffices to show that if \( \| x \| = \| x - P_Ix \| \) then \( P_Ix = 0 \). If \( \| x \| = \| x - P_Ix \| \) then \( \| x - 0 \| = \| x - P_Ix \| \). Hence both sequences \( 0_{\ell_2^-} \) and \( P_Ix \) are solutions of problem (4.1). We proved that the solution is unique, hence \( \| P_Ix \| = 0_{\ell_2^-} \). This completes the proof. □
6 Possible applications and future development

The approach suggested in this paper allows many modifications. We outline below some possible straightforward modifications as well as more challenging problems and possible applications that we leave for the future research.

(i) It would be interesting to investigate sensitivity of the prediction results and of the set \( \{I_k\} \) and an optimal choice of the set \( \{I_k\} \) such as

\[
\text{Maximize } \delta_k + \bar{\delta}_k \text{ over } I \in \mathcal{J}
\]

for \( k = 1, 2, \ldots \).

(ii) It could be interesting to try another basis in \( L_2(I_0) \) for expansion in (4.2).

(iii) Optimization problem in (4.1) is based on optimal approximation in \( L_2(I) \) for Z-transforms. This approximation in can be replaced by approximation in a weighted \( L_2 \)-space on \( I \). This leads to modification of the optimization problem; the weight will represent the relative importance of the approximation on different frequencies.

(iv) It is unclear if an analog of property (3.2) can be obtained with \( d = 2 \) instead of \( d = 1, +\infty \).

References


