Predictors for High Frequency Signals Based on Rational Polynomial Approximation of Periodic Exponentials

N. G. Dokuchaev

ZJU-UIUC (Zhejiang University/University of Illinois at Urbana-Champaign) Institute, Zhejiang University, Haining, Zhejiang Province, China
e-mail: Dokuchaev@intl.zju.edu.cn

Received August 1, 2022; revised November 12, 2022; accepted November 14, 2022

Abstract—We present linear integral predictors for continuous-time high-frequency signals with a finite spectrum gap. The predictors are based on approximation of a complex-valued periodic exponential (complex sinusoid) by rational polynomials.

Key words: forecasting, linear predictors, transfer functions, periodic exponentials, high-frequency signals.

DOI: 10.1134/S003294602204007X

1. INTRODUCTION

We consider prediction and predictors for continuous time signals. A common approach to forecasting of signals is based on removing high-frequency component regarded as noise and using some filters and forecasting of smooth low-frequency component, which is presumed to be easier. This approach assumes the loss of information contained in the high-frequency component that is regarded as noise. But there are also works focused at the extraction of information contained in the high-frequency component. These works are based on a variety of statistical methods and learning models; see, e.g., [1–5] and references therein.

We study pathwise predictability and predictors of continuous-time signals in deterministic setting and in the framework of frequency analysis. It is well known that certain restrictions on the spectrum can ensure opportunities for prediction and interpolation of the signals; see, e.g., [7–12]. These works considered predictability of band-limited signals; the predictors obtained therein were non-robust with respect to small noise in high frequencies; see, e.g., the discussion in [13, ch. 17]. We study predictors for high-frequency signals, i.e., for signals without any restrictions on the rate of the spectrum decay on higher frequencies. We consider signals such that their spectrum have a finite spectral gap, i.e., an interval where its Fourier transform vanishes. It is known that these signals allow unique extrapolations from their past observations. However, feasibility of predicting algorithm is not implied by this uniqueness. In general, uniqueness of a path does not ensure a possibility to predict this path; some discussion on this can be found in [14,15].

Predictors for anticausal convolutions (i.e., integrals including future values) of high-frequency signals were obtained in [16] for signals with finite spectral gap and in [14] for signals with a single-point spectral degeneracy. The predictors therein were independent of spectral characteristics of input signals from a class with a certain spectral degeneracy for low frequencies. These predictors depended on kernels of the corresponding anticausal convolutions.

The present paper offers some principally new predictors for high-frequency signals. The transfer functions for these predictors are polynomials of the inverse $1/\omega$, approximating a periodic
exponential $e^{i\omega T}$, where $\omega \in \mathbb{R}$ represents the frequency, and where $T > 0$ represents a preselected prediction horizon. These predictors allow a compact explicit representation in the time domain and in the frequency domain. Again, the predictors are independent of spectral characteristics of input signals with fixed and known finite spectral gap. The method is based on the approach from [17] for prediction of signals with fast-decaying spectrum, where polynomial approximations of input signals with fixed and known finite spectral gap. The method is based on the approach from [17] for prediction of signals with fast-decaying spectrum, where polynomial approximations of the periodic exponential have been used.

The paper is organized as follows. In Section 2, we formulate the definitions and background facts related to linear weak predictability. In Section 3, we formulate main theorems on predictability and predictors (Theorems 1 and 2). In Section 4, we discuss some implementation problems. Section 5 contains the proofs.

2. PROBLEM SETTING AND DEFINITIONS

Let $x(t)$ be a currently observable complex-valued continuous-time process, $t \in \mathbb{R}$. The goal is to estimate, at current times $t$, the values $x(t + T)$ using historical values of the observable process $x(s)|_{s \leq t}$. Here $T > 0$ is a given prediction horizon.

We need some notation and definitions.

For $p \in [1, +\infty)$ and for a domain $G \subset \mathbb{R}$, we denote by $L_p(G, \mathbb{R})$ and $L_p(G, \mathbb{C})$ the usual $L_p$-spaces of functions $x : G \to \mathbb{R}$ and $x : G \to \mathbb{C}$, respectively. We denote by $C(G, \mathbb{R})$ and $C(G, \mathbb{C})$ the usual linear normed spaces of bounded continuous functions $x : G \to \mathbb{R}$ and $x : G \to \mathbb{C}$, respectively, with the supremum norm.

For $x \in L_p(\mathbb{R}, \mathbb{C})$, $p = 1, 2$, we denote by $X = \mathcal{F}x$ the function defined on $i\mathbb{R}$ as the Fourier transform of $x$:

$$X(i\omega) = (\mathcal{F}x)(i\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} x(t) \, dt, \quad \omega \in \mathbb{R}.$$ 

It is known that if $x \in L_2(\mathbb{R}, \mathbb{C})$, then $X(i \, \cdot \,) \in L_2(\mathbb{R}, \mathbb{C})$.

Let $\mathcal{X}$ be the set of signals $x : \mathbb{R} \to \mathbb{R}$ such that their Fourier transforms $X(i \cdot) \in L_1(\mathbb{R}, \mathbb{C})$. In particular, the class $\mathcal{X}$ includes signals formed as

$$x(t) = \int_{t-\delta}^{t} y(s) \, ds \quad \text{for} \quad y \in L_2(\mathbb{R}, \mathbb{R}), \quad \delta \in \mathbb{R}.$$ 

Clearly, $\mathcal{X} \subset C(\mathbb{R}, \mathbb{C})$, i.e., these signals are bounded and continuous.

We consider $\mathcal{X}$ as a linear normed space equipped with the norm $\|X(i \cdot)\|_{L_1(\mathbb{R}, \mathbb{C})}$, where $X(i \cdot) = \mathcal{F}x$ for $x \in \mathcal{X}$.

Let $\mathcal{P}$ be the set of all continuous mappings $p : \mathcal{X} \to C(\mathbb{R}, \mathbb{C})$ such that, for any $x_1, x_2 \in \mathcal{X}$ and $\tau \in \mathbb{R}$, we have that $p(x_1(\cdot))(t) = p(x_2(\cdot))(t)$ for all $t \leq \tau$ if $x_1(t) = x_2(t)$ for all $t \leq \tau$. In other words, this is the set of “causal” mappings; we will look for predictors in this class.

Let $T > 0$ be given.

**Definition 1.** We say that a class $\mathcal{X} \subset \mathcal{X}$ is linearly predictable with the prediction horizon $T$ if there exists a sequence $\{\tilde{p}_d(\cdot)\}_{d=1}^{+\infty} \subset \mathcal{P}$ such that

$$\sup_{t \in \mathbb{R}} |x(t + T) - \tilde{y}_d(t)| \to 0 \quad \text{as} \quad d \to +\infty, \quad \forall x \in \mathcal{X},$$

where

$$\tilde{y}_d = \tilde{p}_d(x(\cdot)).$$
Definition 2. We say that a class $\mathcal{X} \subset \bar{\mathcal{X}}$ is uniformly linearly predictable with the prediction horizon $T$ if there exists a sequence $\{\tilde{y}_d(\cdot)\}_{d=1}^{+\infty} \subset \mathcal{P}$ such that
\[
\sup_{t \in \mathbb{R}} |x(t + T) - \tilde{y}_d(t)| \to 0 \quad \text{uniformly in } x \in \mathcal{X},
\]
where $\tilde{y}_d(\cdot)$ is as in Definition 1 above.

Functions $\tilde{y}_d(t)$ in the above definitions can be considered as approximate predictions of the process $x(t + T)$.

3. MAIN RESULT

Let $\Omega > 0$ be given, and let $\mathcal{X}_\Omega$ be the set of all signals $x(\cdot) \in \bar{\mathcal{X}}$ such that $X(i\omega) = 0$ for $\omega \in (-\Omega, \Omega)$, for $X = \mathcal{F}x$.

Let $\mathcal{U}_\Omega$ be some set of signals $x \in \mathcal{X}_\Omega$ such that
\[
\|X(i\cdot)\|_{L_1(\mathbb{R}, \mathbb{C})} \leq 1 \quad \text{and} \quad \int_{\omega: |\omega| \geq M} |X(i\omega)| d\omega \to 0 \quad \text{as} \quad M \to +\infty
\]
uniformly over $x \in \mathcal{U}_\Omega$, for $X = \mathcal{F}x$.

Theorem 1. For any $T > 0$, we have the following:
(i) The class $\mathcal{X}_\Omega$ is linearly predictable with the prediction horizon $T$;
(ii) The class $\mathcal{U}_\Omega$ is uniformly linearly predictable with the prediction horizon $T$.

3.1. A Family of Predictors

In this section, we introduce some predictors.

For $d = 1, 2, \ldots$, let $\Psi_d$ be the set of all functions $\sum_{k=1}^{d} \frac{a_k}{z^k}$ defined for $z \in \mathbb{C} \setminus \{0\}$, for all $a_k \in \mathbb{R}$. Let $\Psi := \bigcup_{d} \Psi_d$.

For $d = 0, 1, 2, \ldots$ and $s \in \mathbb{R}$, let $\mathcal{X}(d)(s)$ be the set of all signals $x \in \bar{\mathcal{X}}$ satisfying the condition
\[
\int_{-\infty}^{s} |t^d x(t)| dt < +\infty.
\]
It can be noted that this class includes, in particular, signals $x \in \bar{\mathcal{X}}$ such that
\[
\int_{\mathbb{R}} \left| \frac{d^k X(i\omega)}{d\omega^k} \right|^2 d\omega < +\infty \quad \text{for} \quad k = 0, 1, \ldots, d + 1, \quad X = \mathcal{F}x.
\]

Let $r: \mathbb{R} \to (0, 1]$ be a continuous function such that $r(0) = 1$, $r(\omega) \equiv r(-\omega)$, the function $r(\omega)$ is monotonically nonincreasing on $(0, +\infty)$, and $r(\omega) \to 0$ as $|\omega| \to +\infty$. Let $r_\nu(\omega) := r(\nu \omega)$, $\nu \in (0, 1]$.

Theorem 2. We have the following:
(i) For any $\varepsilon_1 > 0$ and any $x \in \mathcal{X}_\Omega$ such that $\|X(i\cdot)\|_{L_1(\mathbb{R}, \mathbb{C})} \leq 1$, there exists $\nu_0 = \nu_0(\varepsilon_1, x) > 0$ such that, for $X = \mathcal{F}x$ and any $\nu \in (0, \nu_0]$, we have
\[
\int_{\omega: |\omega| \geq \Omega} (1 - r_\nu(\omega))|X(i\omega)| d\omega \leq \varepsilon_1.
\]
Moreover, one can select the same $\nu_0 = \nu_0(\varepsilon_1)$ for all $x \in \mathcal{U}_\Omega$;
(ii) For any \( \varepsilon_2 > 0 \) and \( \nu > 0 \), there exist an integer \( d = d(\nu, \varepsilon_2, T) > 0 \) and a \( \psi_d \in \Psi_d \) such that
\[
\sup_{\omega: |\omega| \geq \Omega} |e^{i\omega T}r_{\nu}(\omega) - \psi_d(i\omega)| \leq \varepsilon_2; \tag{2}
\]

(iii) The predictability considered in part (i) of Theorem 1 for \( x \in X_\Omega \), as well as the predictability considered in part (ii) of Theorem 1 for \( x \in U_\Omega \), can be ensured with the sequence of predictors
\[
p_d: X_\Omega \to C(\mathbb{R}, \mathbb{C}), \quad d = 1, 2, \ldots,
\]
defined by their transfer functions \( \psi_d(i\omega) \). More precisely, for any \( \varepsilon > 0 \) and \( \hat{y}_d(t) = p_d(x(\cdot))(t) \), the estimate
\[
\sup_{t \in \mathbb{R}} |x(t + T) - \hat{y}_d(t)| \leq \varepsilon
\]
holds if \( \nu, d, \) and \( \psi_d \) are such that (1) and (2) hold for sufficiently small \( \varepsilon_1 \) and \( \varepsilon_2 \) such that
\[
\varepsilon_1 + \varepsilon_2 \leq 2\pi \varepsilon.
\]
It can be noted that, for inputs \( x \in X_\Omega \), the transfer functions \( \psi_d(i\omega) \) can be replaced by the functions \( \psi_d(i\omega)\mathbb{I}_{|\omega| \geq \Omega} \), where \( \mathbb{I} \) denotes the indicator function;

(iv) For \( x \in X^{(d-1)}(t) \cap X_\Omega \), the predictors described above can be represented as
\[
p_d(x(\cdot))(t) = \int_{-\infty}^{t} K(t - \tau)x(\tau)\,d\tau, \tag{3}
\]
where
\[
K(t) = \sum_{k=1}^{d} a_k \frac{t^{k-1}}{(k-1)!}.
\]
Here, \( a_k \in \mathbb{C} \) are the coefficients for a function \( \psi_d(z) = \sum_{k=1}^{d} a_k z^{-k} \) from part (ii).

### 3.2. Integral Representation of Predictors for General-Type \( x \in X_\Omega \)

Representation (3) for the above predictors requires that \( x \in X^{(d-1)}(t) \). Let us discuss possibilities of representations in the time domain for general-type \( x \in X_\Omega \).

Consider operators \( h_k \) defined on \( X_\Omega \) by their transfer functions \( (i\omega)^{-k} \), \( k = 1, 2, \ldots \). In other words, if \( y = h_k(x) \) for \( x \in X_\Omega \), then \( Y(i\omega) = (i\omega)^{-k}X(i\omega) \) for \( Y = \mathcal{F}y \) and \( X = \mathcal{F}x \). Clearly, \( h_k(x(\cdot)) \in X_\Omega \), the Fourier transforms of processes \( h_k(x(\cdot)) \), vanish on \([−\Omega, \Omega]\), and the operators \( h_k: X_\Omega \to C(\mathbb{R}, \mathbb{C}) \) are continuous. By the definitions, it follows that
\[
p_d(x(\cdot))(t) = \sum_{k=1}^{d} a_k h_k(x(\cdot))(t).
\]
It can be noted that \( p_d(\cdot) \) depends on \( T \) via the coefficients \( a_k \) defined for functions \( \psi_d(\omega) \) approximating \( e^{i\omega T} \).

Formally, the operator \( h_k(x(\cdot)) \) can be represented as
\[
h_k(x(\cdot))(s_k) = \int_{-\infty}^{s_k} ds_{k-1} \int_{-\infty}^{s_{k-1}} ds_{k-2} \ldots \int_{-\infty}^{s_2} ds_1 \int_{-\infty}^{s_1} x(s)\,ds, \tag{4}
\]
For general-type \( x \in \mathcal{X}_\Omega \), there is no guarantee that \( x \in L_1(\mathbb{R}, \mathbb{R}) \) or \( h_k(x(:)) \in L_1(\mathbb{R}, \mathbb{R}) \). However, the above integrals are well defined, because they can be replaced with integrals over finite time intervals

\[
h_1(x(:))(t) = \int_{-\infty}^{t} x(s) \, ds, \quad h_k(x(:))(t) = \int_{-\infty}^{t} h_{k-1}(x(s)) \, ds, \quad k = 2, 3, \ldots
\]

where the \( R_k \) are roots of signals \( h_k(x(:))(t) \). This is possible because of the special properties of signals with Fourier transform vanishing on an interval: for any \( \tau < 0 \), these signals have infinitely many roots in the interval \((-\infty, \tau)\); see, e.g., [18]. Hence, the predictors defined in Theorem 2 for \( x \in \mathcal{X}_\Omega \) allow an alternative integral representation via (4) or (5).

4. ON NUMERICAL IMPLEMENTATION OF PREDICTORS

Direct implementation of the predictor introduced in Theorem 2 requires evaluation of integrals over semi-infinite intervals, which could be numerically challenging. However, this theorem could lead to predicting methods avoiding this calculation. Let us discuss these possibilities.

Let \( t_1 \in \mathbb{R} \) be given. Let \( x_k := h_k(x) \) for \( x \in \mathcal{X}_\Omega \), \( k = 1, 2, \ldots \), and let \( \eta_k := x_k(t_1) \).

**Lemma 1.** In the notation of Theorem 2, for any \( t \geq t_1 \) we have that \( \hat{y}_d(t) = p_d(x(\cdot)) \) can be represented as

\[
\hat{y}_d(t) = \sum_{k=1}^{d} a_k \left( \sum_{\ell=1}^{k} c_{\ell}(t) \eta_\ell + f_k(t) \right),
\]

where \( c_{\ell}(t) := (t - t_1)^{\ell-1}/(\ell - 1)! \) and

\[
f_k(t) := \int_{t_1}^{t} d\tau_1 \int_{t_1}^{\tau_1} d\tau_2 \ldots \int_{t_1}^{\tau_k} x(s) \, ds.
\]

This lemma shows that calculating the prediction \( \hat{y}_d(t) \) of \( x(t + T) \) is easy for \( t > t_1 \) if we know all \( \eta_k \) and observe \( x_{\mid_{[t_1,t]}} \).

Let us discuss possible ways to evaluate \( \eta_k \) avoiding direct integration over infinite intervals.

First, let us observe that (6) implies a useful property given below.

**Corollary 1.** For any \( \varepsilon > 0 \), there exist an integer \( d = d(\varepsilon) > 0 \) and \( a_1, \ldots, a_d \in \mathbb{R} \) such that, for any \( x \in \mathcal{X}_\Omega \) and \( t_1 \in \mathbb{R} \), there exist \( \bar{\eta}_1, \ldots, \bar{\eta}_d \in \mathbb{R} \) such that \( |x(t + T) - y_d(t)| \leq \varepsilon \) for all \( t \geq t_1 \), where

\[
y_d(t) = y_d(t, \bar{\eta}_1, \ldots, \bar{\eta}_d) := \sum_{k=1}^{d} a_k \left( \sum_{\ell=1}^{k} c_{\ell}(t) \bar{\eta}_\ell + f_k(t) \right).
\]

In this corollary, \( d = d(\varepsilon) \) can be selected as defined in parts (i) and (ii) of Theorem 2, where \( \varepsilon_1 \) and \( \varepsilon_2 \) are such that \( \varepsilon_1 + \varepsilon_2 \leq 2\pi \varepsilon \).

Further, let us discuss the use of (6) for evaluating \( \eta_k \) and for prediction. Let \( \theta > t_1 \). Assume that the goal is to forecast the value \( x(\theta + T) \) given observations at times \( t \leq \theta \). It appears that if \( \theta > t_1 + T \), then Corollary 1 gives an opportunity to construct predictors via the fitting parameters \( \eta_1, \ldots, \eta_d \) using past observations available for \( t \in [t_1, \theta - T] \): we can match the values
If (9) holds, we can conclude that
to a nonlinear fitting problem for the unknowns

directly observable (without calculation of integrals of semi-infinite intervals required for

As an approximation of the true \( \eta_1, \ldots, \eta_d \), we can accept a set \( \tilde{\eta}_1, \ldots, \tilde{\eta}_d \) such that

\[
|x(t + T) - y_d(t, \tilde{\eta}_1, \ldots, \tilde{\eta}_d)| \leq \varepsilon, \quad \forall t \in [t_1, \theta - T].
\] (9)

(Recall that, at time \( \theta \), values \( x(t + T) \) and \( y_d(t, \tilde{\eta}_1, \ldots, \tilde{\eta}_d) \) are observable for these \( t \in [t_1, \theta - T] \).) If (9) holds, we can conclude that \( y_d(t, \tilde{\eta}_1, \ldots, \tilde{\eta}_d) \) delivers an acceptable prediction of \( x(t + T) \) for these \( t \). Clearly, Theorem 2 implies that a set \( \tilde{\eta}_1, \ldots, \tilde{\eta}_d \) ensuring (9) exists, since this inequality holds with \( \tilde{\eta}_k = \eta_k \).

The corresponding value \( y_d(\theta, \tilde{\eta}_1, \ldots, \tilde{\eta}_d) \) would give an estimate for \( \tilde{y}_d(\theta) \) and, respectively, for \( x(\theta + T) \).

Furthermore, finding a set \( \tilde{\eta}_1, \ldots, \tilde{\eta}_d \) ensuring (9) could still be difficult. Instead, one can consider fitting predictions and observations at a finite number of points \( t \in [t_1, T - \theta] \).

Let an integer \( d \geq d \) and a set \( \{t_m\}_{m=1}^{d} \subset \mathbb{R} \) be selected such that

\[
t_1 < t_2 < t_3 < \ldots < t_{d-1} < t_d \leq \theta - T.
\]

We suggest to use observations \( x(t) \) at times \( t = t_m \). Consider a system of equations

\[
\sum_{k=1}^{d} a_k \left( \sum_{\ell=1}^{k} c_{\ell}(t_m)\eta_\ell + f_k(t_m) \right) = \zeta_m, \quad m = 1, \ldots, d.
\] (10)

First, consider the case where \( d = d \). In this case, we can select \( \zeta_m = x(t_m + T) \); these values are directly observable (without calculation of integrals of semi-infinite intervals required for \( \tilde{y}_d(t_m) \)). The corresponding choice of \( \tilde{\eta}_k \) ensures zero prediction error for \( x(t_m + T) \), \( m = 1, \ldots, d \).

Including into consideration more observations, i.e., selecting larger \( d > d \) and wider interval \( [t_1, \theta - T] \), would improve the estimation of \( \eta_k \). If we consider \( d > d \), then, in the general case, it would not be feasible to achieve that \( y_d(t, \tilde{\eta}_1, \ldots, \tilde{\eta}_d) = x(t_m + T) \) for all \( m \), since it cannot be guaranteed that system (10) is solvable for \( \zeta_m \equiv x(t_m + T) \): the system will be overdefined. Nevertheless, the estimate presented in (9) can still be achieved for any arbitrarily large \( d \), since (9) holds. A solution could be found using methods for fitting linear models.

Furthermore, instead of calculating the coefficients \( a_k \) via solving the approximation problem for the complex exponential described in parts (i) and (ii) of Theorem 2, one can find these coefficients considering them as additional unknowns in system (10) with \( d \geq 2d \). Theorem 2 implies again that there exist \( \tilde{\eta}_k = \eta_k \in \mathbb{R} \) and \( a_k \in \mathbb{R} \) such that (10) holds with \( \zeta_m \equiv \tilde{y}_d(t_m) \). This would lead to a nonlinear fitting problem for the unknowns \( a_1, \ldots, a_d, \tilde{\eta}_1, \ldots, \tilde{\eta}_d \).

So far, the consistency of these estimates is unclear, since a choice of smaller \( \varepsilon \) leads to larger \( d \). We leave analysis of these methods for future research.

5. PROOFS

Theorem 1 immediately follows from Theorem 2.

**Proof of Theorem 2.** Let us prove part (i). Let us select \( M > 0 \) such that

\[
\int_{\omega: |\omega| > M} |X(i\omega)| \, d\omega < \varepsilon_1/2, \quad \forall x \in \mathcal{U}_\Omega.
\]
We have that \( r_\nu(\omega) \to 1 \) uniformly in \( \omega \in [-M, M] \), since
\[
0 < 1 - r_\nu(\omega) \leq 1 - r_\nu(M) \quad \text{for} \quad \omega \in [-M, M].
\]
Hence, one can select \( \nu > 0 \) such that
\[
\int_{-M}^{M} (1 - r_\nu(\omega)) |X(i\omega)| d\omega \leq \varepsilon_1 / 2.
\]
This implies that
\[
\int_{-\infty}^{\infty} (1 - r_\nu(\omega)) |X(i\omega)| d\omega = \int_{-M}^{M} (1 - r_\nu(\omega)) |X(i\omega)| d\omega + \int_{\omega: |\omega| > M} (1 - r_\nu(\omega)) |X(i\omega)| d\omega
\]
\[
\leq \int_{-M}^{M} (1 - r_\nu(\omega)) |X(i\omega)| d\omega + \int_{\omega: |\omega| > M} |X(i\omega)| d\omega \leq \frac{\varepsilon_1}{2} + \frac{\varepsilon_1}{2} \leq \varepsilon_1.
\]
This completes the proof of part (i).

Let us prove part (ii). By the Stone–Weierstrass theorem for real continuous functions on locally compact spaces, it follows that there exist \( \psi_d^c(\omega) \in \Psi_d \) and \( \psi_d^s(\omega) \in \Psi_d \) such that
\[
\sup_{\omega: |\omega| \geq \Omega} |\cos(T \cdot r_\nu(\cdot) - \psi_d^c(\cdot))| \leq \varepsilon_2 / 2, \quad \sup_{\omega: |\omega| \geq \Omega} |\sin(T \cdot r_\nu(\cdot) - \psi_d^s(\cdot))| \leq \varepsilon_2 / 2
\]
(see, e.g., [19, Theorem 12, pp. 240–241]).

It is easy to see that it suffices to select odd functions \( \psi_d^c(\omega) = \sum_{k=1}^{d} \gamma_k^c \omega^{-k} \) and even functions \( \psi_d^s(\omega) = \sum_{k=1}^{d} \gamma_k^s \omega^{-k} \), i.e., \( \gamma_{2m+1}^c = 0 \) and \( \gamma_{2m}^s = 0 \) for integers \( m \geq 0 \). Here \( \gamma_k^c \) and \( \gamma_k^s \) are real.

We construct the desired functions as
\[
\psi_d(i\omega) = \psi_d^c(\omega) + i\psi_d^s(\omega) = \sum_{k=1}^{d} \gamma_k^c \omega^{-k} + i \sum_{k=1}^{d} \gamma_k^s \omega^{-k} = \sum_{k=1}^{d} a_k(i\omega)^{-k},
\]
where the coefficients \( a_k \in \mathbb{R} \) are defined as follows:
- If \( k = 2m \) for an integer \( m \), then \( a_k = (-1)^m \gamma_k^c \);
- If \( k = 2m + 1 \) for an integer \( m \), then \( a_k = -(-1)^m \gamma_k^s \).

This choice of \( \psi_d \) ensures that estimate (2) holds. This completes the proof of part (ii).

Let us prove part (iii). Assume that estimates (1) and (2) hold for selected \( d, \nu, \) and \( \psi_d \). We have
\[
2\pi(x(t + T) - \hat{y}_d(t)) = \int_{-\infty}^{\infty} e^{i\omega t} (e^{i\omega T} - \psi_d(i\omega)) X(i\omega) d\omega = A(t) + B(t),
\]
where
\[
A(t) = \int_{-\infty}^{\infty} e^{i\omega t} (e^{i\omega T} - e^{i\omega T} r_\nu(\omega)) X(i\omega) d\omega,
\]
\[
B(t) = \int_{-\infty}^{\infty} e^{i\omega t} (e^{i\omega T} r_\nu(\omega) - \psi_d(i\omega)) X(i\omega) d\omega.
\]
Clearly,
\[ |A(t)| \leq \int_{-\infty}^{\infty} (1 - r_\nu(\omega))|X(i\omega)| \, d\omega \leq \varepsilon_1 \]
and
\[ |B(t)| \leq \int_{-\infty}^{\infty} |e^{i\omega T}T_\nu(\omega) - \psi_d(i\omega)||X(i\omega)| \, d\omega \]
\[ \leq \sup_{\omega: |\omega| \geq \Omega} |e^{i\omega T}T_\nu(\omega) - \psi_d(i\omega)| \int_{-\infty}^{\infty} |X(i\omega)| \, d\omega \leq \varepsilon_2. \]

Hence,
\[ 2\pi|x(t + T) - \tilde{g}_d(t)| \leq \varepsilon_1 + \varepsilon_2. \]

This proves the uniform predictability considered in part (ii) of Theorem 1 for signals \( x \in \mathcal{U}_\Omega \).
The predictability considered in part (i) of Theorem 1 immediately follows from the above proof applied to singletons \( \mathcal{U}_\Omega = \{ x(\cdot) \} \) multiplied by a constant, if needed, to bypass the restriction that \( ||X(\cdot)||_{L_1(\mathbb{R},\mathbb{C})} \leq 1 \). This completes the proof of part (iii).

Let us prove part (iv). First, the known properties of Fourier transforms of derivatives and antiderivatives imply representations (4) and (5) (see Section 3.2). Now (iv) can be obtained by successive application of the Fubini’s theorem to signals \( (\tau - s)^\ell x(s) \) integrable in \( L_1((-\infty, t], \mathbb{R}) \), presented in (4) for \( \ell = 1, 2, \ldots, \tau \in (\infty, s] \). △

**Proof of Lemma 1.** In the notation of Theorem 2, we have that \( y_d(t) = \sum_{k=1}^{d} a_k x_k(t) \) for any \( t \geq t_1 \), i.e.,
\[ y_d(t) = \sum_{k=1}^{d} a_k \left( \eta_k + \int_{t_1}^{t} x_{k-1}(s) \, ds \right) \]  
(11)
(here we assume that \( x_0 := x \)). Further, we have
\[ \int_{t_1}^{t} x_1(t) \, dt = \int_{t_1}^{t} \left( \eta_1 + \int_{t_1}^{\tau} x_0(s) \, ds \right) \, d\tau = \eta_1(t - t_1) + \int_{t_1}^{\tau} x(s) \, ds \]
and
\[ \int_{t_1}^{t} x_2(t) \, dt = \int_{t_1}^{t} \left( \eta_2 + \int_{t_1}^{\tau} x_1(s) \, ds \right) \, d\tau_1 = \eta_2(t - t_1) + \int_{t_1}^{\tau_1} x_1(s) \, ds \]
\[ = \eta_2(t - t_1) + \int_{t_1}^{\tau_1} \left[ \eta_1(\tau_1 - t_1) + \int_{t_1}^{\tau_2} x(s) \, ds \right] \, d\tau_1 \]
\[ = \eta_2(t - t_1) + \frac{\eta_1^2}{2}(t - t_1) + \int_{t_1}^{\tau_2} x(s) \, ds. \]

Similarly, we obtain
\[ \int_{t_1}^{t} x_k(t) \, dt = \eta_k(t - t_1) + \frac{\eta_{k-1}}{2}(t - t_1)^2 + \ldots + \frac{\eta_1}{k!}(t - t_1)^k + \int_{t_1}^{t} \int_{\tau_1}^{\tau_2} \ldots \int_{\tau_k}^{t} x(s) \, ds. \]
It follows that
\[ \eta_k + \int_t^t x_{k-1}(s) \, ds = \sum_{\ell=1}^{k} c_{\ell}(t) \eta_\ell + f_k(t). \]
Together with (11), this proves (8), which completes the proof of Lemma 1. △

REFERENCES

1. Brooks, C. and Hinich, M.J., Detecting Intraday Periodicities with Application to High Frequency Exchange Rates, *J. Roy. Statist. Soc. Ser. C*, 2006, vol. 55, no. 2, pp. 241–259. https://doi.org/10.1111/j.1467-9876.2006.00534.x

2. Christensen, H.L., Murphy, J., and Godsill, S.J., Forecasting High-Frequency Futures Returns Using Online Langevin Dynamics, *IEEE J. Sel. Top. Signal Process.*, 2012, vol. 6, no. 4, pp. 366–380. https://doi.org/10.1109/JSTSP.2012.2191532

3. Granger, C.W.J., Extracting Information from Mega-panels and High-Frequency Data, *Statist. Neerlandica*, 1998, vol. 52, no. 3, pp. 258–272. https://doi.org/10.1111/1467-9574.00084

4. Li, Z., Han, J., and Song, Y., On the Forecasting of High-Frequency Financial Time Series Based on ARIMA Model Improved by Deep Learning, *J. Forecast.*, 2020, vol. 39, no. 7, pp. 1081–1097. https://doi.org/10.1002/for.2677

5. Luo, S. and Tian, C., Financial High-Frequency Time Series Forecasting Based on Sub-step Grid Search Long Short-Term Memory Network, *IEEE Access*, 2020, vol. 8, pp. 203183–203189. https://doi.org/10.1109/ACCESS.2020.3037102

6. Engle, R.F., The Econometrics of Ultra-high-Frequency Data, *Econometrica*, 2000, vol. 68, no. 1, pp. 1–22. https://doi.org/10.1111/1468-0262.00091

7. Knab, J.J., Interpolation of Band-Limited Functions Using the Approximate Prolate Series, *IEEE Trans. Inform. Theory*, 1979, vol. 25, no. 6, pp. 717–720. https://doi.org/10.1109/TIT.1979.1056115

8. Lyman, R.J., Edmonson, W.W., McCullough, S., and Rao, M., The Predictability of Continuous-Time, Bandlimited Processes, *IEEE Trans. Signal Process.*, 2000, vol. 48, no. 2, pp. 311–316. https://doi.org/10.1109/78.823959

9. Lyman, R.J. and Edmonson, W.W., Linear Prediction of Bandlimited Processes with Flat Spectral Densities, *IEEE Trans. Signal Process.*, 2001, vol. 49, no. 7, pp. 1564–1569. https://doi.org/10.1109/78.928709

10. Papoulis, A., A Note on the Predictability of Band-limited Processes, *Proc. IEEE*, 1985, vol. 73, no. 8, pp. 1332–1333. https://doi.org/10.1109/PROC.1985.13284

11. Marvasti, F., Comments on “A Note on the Predictability of Band-limited Processes,” *Proc. IEEE*, 1986, vol. 74, no. 11, pp. 1596. https://doi.org/10.1109/PROC.1986.13874

12. Vaidyanathan, P.P., On Predicting a Band-limited Signal Based on Past Sample Values, *Proc. IEEE*, 1987, vol. 75, no. 8, pp. 1125–1127. https://doi.org/10.1109/PROC.1987.13856

13. Higgins, J.R., *Sampling Theory in Fourier and Signal Analysis: Foundations*, Oxford: Clarendon; New York: Oxford Univ. Press, 1996.

14. Dokuchaev, N., On Linear Weak Predictability with Single Point Spectrum Degeneracy, *Appl. Comput. Harmon. Anal.*, 2021, vol. 53, pp. 116–131. https://doi.org/10.1016/j.acha.2021.01.005

15. Dokuchaev, N.G., On Data Compression and Recovery for Sequences Using Constraints on the Spectrum Range, *Probl. Peredachi Inf.*, 2021, vol. 57, no. 4, pp. 74–78 [Probl. Inf. Transm. (Engl. Transl.), 2021, vol. 57, no. 4, pp. 365–372]. https://doi.org/10.1134/S0032946021040062

16. Dokuchaev, N.G., The Predictability of Band-limited, High-Frequency, and Mixed Processes in the Presence of Ideal Low-Pass Filters, *J. Phys. A: Math. Theor.*, 2008, vol. 41, no. 38, pp. 382002 (7 pp.). https://doi.org/10.1088/1751-8113/41/38/382002
17. Dokuchaev, N., Limited Memory Predictors Based on Polynomial Approximation of Periodic Exponentials, *J. Forecast.*, 2022, vol. 41, no. 5, pp. 1037–1045. https://doi.org/10.1002/for.2843

18. Blank, N. and Ulanovskii, A., Paley–Wiener Functions with a Generalized Spectral Gap, *J. Fourier Anal. Appl.*, 2011, vol. 17, no. 5, pp. 899–915. https://doi.org/10.1007/s00041-010-9160-3

19. Stone, M.H., The Generalized Weierstrass Approximation Theorem, *Math. Mag.*, 1948, vol. 21, no. 4, pp. 167–184; no. 5, pp. 237–254 (continued). https://doi.org/10.2307/3029750; https://doi.org/10.2307/3029337