Известия НАН Армении, Математика, том 55, н. 2, 2020, стр. 9 – 34

ON HYPERBOLIC DECAY OF PREDICTION ERROR VARIANCE FOR DETERMINISTIC STATIONARY SEQUENCES

N. M. BABAYAN, M. S. GINOVYAN

Russian-Armenian University, Yerevan, Armenia
Boston University, Boston, USA, Institute of Mathematics, Yerevan, Armenia
E-mails: *nmbabayan@gmail.com*; *ginovyan@math.bu.edu*

Abstract. One of the main problems in prediction theory of second-order stationary processes, called direct prediction problem, is to describe the asymptotic behavior of the best linear mean squared one-step ahead prediction error variance in predicting the value $X(0)$ of a stationary process $X(t)$ by the observed past of finite length n as n goes to infinity, depending on the regularity nature (deterministic or non-deterministic) of the underlying observed process $X(t)$. In this paper, we obtain sufficient conditions for hyperbolic decay of prediction error variance for deterministic stationary sequences, generalizing a result obtained by M. Rosenblatt (Some Purely Deterministic Processes, J. of Math. and Mech., 6(6), 801-810, 1957).

MSC2010 numbers: 60G10, 60G25, 62M15, 62M20.
Keywords: prediction problem; deterministic stationary process; singular spectral density; Rosenblatt's theorem.

1. Introduction

1.1. **The prediction problem.** One of the main problems in prediction theory of second-order stationary processes, called direct prediction problem, is to describe the asymptotic behavior of the best linear mean squared one-step ahead prediction error variance in predicting the value $X(0)$ of the stationary process $X(t)$ by the observed past of finite length n as n goes to infinity, depending on the regularity nature (deterministic or nondeterministic) of the underlying observed process $X(t)$.

Let $X(t)$, $t \in \mathbb{Z} = \{0, \pm 1, \dots\}$, be a wide sense stationary stochastic sequence with spectral function $F(\lambda)$ and spectral density function $f(\lambda)$, $\lambda \in \Lambda = [-\pi, \pi]$. Denote by $\sigma_n^2(F)$ the best linear mean squared one-step ahead prediction error variance in predicting the random variable $X(0)$ by the past of $X(t)$ of finite length n: $X(t)$, $-n \le t \le -1$, and let $\sigma^2(F) = \sigma_\infty^2(F)$ be the prediction error variance of $X(0)$ by the entire infinite past: $X(t)$, $t \le -1$. Define the relative prediction error $\delta_n(F) = \sigma_n^2(F) - \sigma^2(F)$, and observe that it is nonnegative and tends to zero as $n \to \infty$. The direct prediction problem is to describe the rate of decrease of $\delta_n(F)$ to zero as $n \to \infty$, depending on the regularity nature (deterministic or nondeterministic) and on the dependence structure of the underlying observed process $X(t)$.

Notice that the aforementioned prediction problem goes back to the classical works by A. Kolmogorov, G. Szegö and N. Wiener, and later for different classes of stationary models has been considered by many authors. The problem has been studied most intensively for nondeterministic processes, that is, in the case where the prediction error is known to be positive ($\sigma^2(F) > 0$) (see Baxter [2], Devinatz [9], Doob [10], Golinski [14], Grenander and Rosenblatt [17], Grenander and Szegö [18], Helson and Szegö [19], Hirshman [21], Ibragimov [23], Ibragimov and Solev [25], Kolmogorov [27], [28], Pourahmadi [29], Rozanov [32], Wiener [34] and others (more references can be found in Bingham [5] and Ginovyan [13]). This is not surprising because from application point of view the nondeterministic models are more realistic and represent great interest.

The case of deterministic processes, that is, when $\sigma^2(F) = 0$, represents mostly theoretical interest. However, it is also important from application point of view. For example, as it was pointed out by M. Rosenblatt [31], situations of this type arise in Neumann's theoretical model of storm-generated ocean waves. Also, such models are of interest for meteorology, because the meteorological spectra often have a gap in the mesoscale region (see Fortus [11]).

There are only few works devoted to the study of asymptotic behavior of prediction error for deterministic processes. It goes back to the classical work by M. Rosenblatt [31], where using the technique of orthogonal polynomials and Szegö's results, M. Rosenblatt has investigated the asymptotic behavior of the prediction error variance $\delta_n(F) = \sigma_n^2(F)$ for discrete-time deterministic processes in the following two cases:

(a) the spectral density $f(\lambda)$ is continuous and vanishes on an interval,

(b) the spectral density $f(\lambda)$ has a high order contact with zero.

Later the problem (a) was studied by Babayan [3], [4], Davisson [8], and Fortus [11], where some generalizations and extensions of Rosenblatt's result have been obtained.

In this paper we consider the case (b), that is, when the spectral density $f(\lambda)$ has a high order contact with zero, and obtain sufficient conditions for hyperbolic decay of prediction error variance, generalizing the corresponding result of Rosenblatt [31], obtained in this case.

Throughout the paper we will use the following notation. The letters C, c, M and m with or without indices are used to denote positive constants, the values of which can vary from line to line. For two functions $f(\lambda)$ and $g(\lambda)$, $\lambda \in \Lambda$, we will write

$f(\lambda) \underset{\lambda\to\lambda_0}{\cong} g(\lambda)$ if $\lim_{\lambda\to\lambda_0} \frac{f(\lambda}{g(\lambda)} = c$, $c \neq 0$, and $f(\lambda) \underset{\lambda\to\lambda_0}{\sim} g(\lambda)$ if $c = 1$. A similar notation we will use for sequences: for two sequences $\{a_n > 0, n \in \mathbb{N} = \{1, 2, \ldots\}\}$ and $\{b_n > 0, n \in \mathbb{N}\}$, we will write $a_n \cong b_n$ if $\lim_{n\to\infty} \frac{a_n}{b_n} = c$, $c \neq 0$, and $a_n \sim b_n$ if $c = 1$.

The paper is organized as follows. In the remainder of this section we introduce the model of interest - a stationary process, recall some key notions and results from the theory of stationary process, and state the infinite prediction problem. In Section 2 we state the finite prediction problem, present a formula for finite prediction error in terms of orthogonal polynomials on the unit circle, and state the Kolmogorov-Szegö theorem. Section 3 is devoted to the asymptotic behavior of the finite prediction error for nondeterministic processes. Here we briefly review some important known results. Section 4 is devoted to the asymptotic behavior of the finite prediction error for deterministic processes. Here we state and prove a number of new theorems.

1.2. **The Model.** In this subsection we introduce the model of interest - a stationary process, recall some key notions and results from the theory of stationary process (Kolmogorov's isometric isomorphism theorem, spectral representations of the covariance function and the process, etc.)

Let $\{X(t),\, t \in \mathbb{Z}\}$ be a centered, real-valued, discrete-time, second-order stationary random process defined on a probability space $(\Omega, \mathcal{F}, P)$ with covariance function $r(t)$, that is, $\mathbb{E}|\mathrm{X(t)}|^2 < \infty$, $\mathbb{E}[\mathrm{X(t)}] = 0$, $r(t) = \mathbb{E}[\mathrm{X(t)X(0)}]$, $t \in \mathbb{Z}$, where $\mathbb{E}[\cdot]$ stands for the expectation operator with respect to measure P.

By the well-known Herglotz' theorem (see [33], p. 421), there is a finite measure μ on $(\Lambda, \mathfrak{B}(\Lambda))$, where $\Lambda = [-\pi, \pi]$ and $\mathfrak{B}(\Lambda)$ is the Borel σ-algebra on Λ, such that for any $t \in \mathbb{Z}$ the covariance function $r(t)$ admits the following *spectral* representation:

$$r(t) = \int_{-\pi}^{\pi} e^{-it\lambda} d\mu(\lambda). \tag{1.1}$$

The measure μ in (1.1) is called the *spectral measure* of the process $X(t)$. The function $F(\lambda) = \mu[-\pi, \lambda]$, $\lambda \in \Lambda$, is called the *spectral function* of the process $X(t)$. If $F(\lambda)$ is absolutely continuous (with respect to Lebesgue measure), then the function $f(\lambda) = dF(\lambda)/d\lambda$ is called the *spectral density* of the process $X(t)$. Notice that $f(\lambda) \geq 0$ and $f(\lambda) \in L^1(\Lambda)$. The set $E_f = \{e^{i\lambda} : f(\lambda) > 0\}$ is called the *spectrum* of the process $X(t)$.

We assume that $X(t)$ is a *non-degenerate process*, that is, $\mathrm{Var}[X(0)] = E|X(0)|^2 = r(0) > 0$. Also, to avoid the trivial cases, we will assume that the spectral measure

μ is non-trivial, that is, μ has infinite support. We write

$$\mu(\lambda) = \mu_{AC}(\lambda) + \mu_S(\lambda) = \int_{-\pi}^{\lambda} f(u)du + \mu_S(\lambda), \tag{1.2}$$

so $f(\lambda)$ is the spectral density and μ_S is the singular part of μ, that is, $\mu_S = \mu_{SC} + \mu_{PP}$, where $\mu = \mu_{AC} + \mu_{SC} + \mu_{PP}$ is the Lebesgue decomposition of μ into an absolutely continuous (with respect to Lebesgue measure) part (μ_{AC}), a singular continuous part (μ_{SC}), and a pure point part (μ_{PP}). The same representations we have also for spectral function $F(\lambda)$.

By the well-known Cramér theorem (see [33], p. 430), for any stationary process $\{X(t),\, t \in \mathbb{Z}\}$ with spectral measure μ there exists an orthogonal stochastic measure $Z = Z(B)$, $B \in \mathfrak{B}(\Lambda)$, such that for every $t \in \mathbb{Z}$ the process $X(t)$ admits the following *spectral* representation:

$$X(t) = \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda). \tag{1.3}$$

Moreover, $\mathbb{E}\left[|\mathrm{Z(B)}|^2\right] = \mu(\mathrm{B})$ for every $B \in \mathfrak{B}(\Lambda)$. For definition and properties of orthogonal stochastic measures and stochastic integral in (1.3) we refer, e.g., [33], Chapter VI.

Given a probability space $(\Omega, \mathcal{F}, P)$, define the L^2-space of random variables $\xi = \xi(\omega)$, $\mathbb{E}[\xi] = 0$:

$$L^2(P) = \left\{\xi :\, ||\xi||^2 = \int_{\Omega} |\xi(\omega)|^2 dP(\omega) < \infty\right\}. \tag{1.4}$$

Then $L^2(P)$ becomes a Hilbert space with the following inner product: for $\xi, \eta \in L^2(P)$

$$(\xi, \eta) = \mathbb{E}[\xi\eta] = \int_{\Omega} \xi(\omega)\overline{\eta(\omega)} \mathrm{dP}(\omega). \tag{1.5}$$

For $a, b \in \mathbb{Z}$, $-\infty \le a \le b \le \infty$, we define the space $H_a^b(X)$ to be the closed linear subspace of the space $L^2(P)$ spanned by the random variables $X(t, \omega)$, $t \in [a, b]$:

$$H_a^b(X) = \overline{sp}\{X(t),\ a \le t \le b\}_{L^2(P)}. \tag{1.6}$$

Observe that the subspace $H_a^b(X)$ consists of all finite linear combinations, $\sum_{k=1}^n c_k X(t_k)$ ($a \le t_k \le b$, $k, n \in \mathbb{N}$), as well as, their $L^2(P)$-limits.

Definition 1.1. The space $H(X) = H_{-\infty}^{\infty}(X)$ is called the *Hilbert space generated by the process* $X(t)$, or the *time-domain* of $X(t)$.

Consider the weighted L^2-space $L^2(\mu)$ of complex-valued functions $\varphi(\lambda)$, $\lambda \in \Lambda$, defined by

$$L^2(\mu) = \left\{\varphi(\lambda) :\, ||\varphi||_\mu^2 := \int_{-\pi}^{\pi} |\varphi(\lambda)|^2 d\mu(\lambda) < \infty\right\}. \tag{1.7}$$

Then $L^2(\mu)$ becomes a Hilbert space with the following inner product: for $\varphi, \psi \in L^2(\mu)$

$$(\varphi, \psi)_\mu = \int_{-\pi}^{\pi} \varphi(\lambda)\overline{\psi}(\lambda)d\mu(\lambda). \tag{1.8}$$

For $a, b \in \mathbb{Z}$, $-\infty \leq a \leq b \leq \infty$ define the space $H_a^b(\mu)$ to be the closed linear subspace of the space $L^2(\mu)$ spanned by the exponents $e^{it\lambda}$, $t \in [a, b]$:

$$H_a^b(\mu) = \overline{sp}\{e^{it\lambda},\ a \leq t \leq b\}_{L^2(\mu)}. \tag{1.9}$$

Definition 1.2. The Hilbert space $H(\mu) := H_{-\infty}^{\infty}(\mu)$ is called the *frequency-domain* of the process $X(t)$.

Kolmogorov's Isometric Isomorphism Theorem states that for any stationary process $X(t)$ with spectral measure μ there exists a unique isometric isomorphism V between the time- and frequency-domains $H(X)$ and $L^2(\mu)$, such that $V[X(t)] = e^{it\lambda}$ for any $t \in \mathbb{Z}$. In particular, we have

1. For any random variable $Y \in H(X)$ there exist a unique function $\varphi(\lambda) \in L^2(\mu)$, such that Y admits the spectral representation

$$Y = \int_{-\pi}^{\pi} \varphi(\lambda)dZ(\lambda), \tag{1.10}$$

where Z is the orthogonal stochastic measure in the spectral representation (1.3) of $X(t)$, and for any function $\varphi(\lambda) \in L^2(\mu)$ the stochastic integral (1.10) defines an element $Y \in H(X)$.

2. For any $Y_i \in H(X)$ and $\varphi_i(\lambda) = V[Y_i] \in L^2(\mu)$, $i = 1, 2$,

$$(Y_1, Y_2) = (\varphi_1, \varphi_2)_\mu. \tag{1.11}$$

3. Any linear problem in the time-domain $H(X)$ can be translated into one in the frequency-domain $L^2(\mu)$, and vice versa. This fact allows to study stationary processes using analytic methods.

1.3. **The infinite prediction problem.** Observe first that since by assumption $X(t)$ is a non-degenerate process, the time-domain $H(X)$ of $X(t)$ is non-trivial, that is, $H(X)$ contains elements different from zero.

Definition 1.3. The space $H_{t-n}^t(X)$ is called the *finite history*, or *past of length n and present* of the process $X(u)$ up to time t. The space $H_t(X) = H_{-\infty}^t(X)$ is called the *entire history*, or *infinite past and present* of the process $X(u)$ up to time t. The space

$$H_{-\infty}(X) = \cap_t H_{-\infty}^t(X) \tag{1.12}$$

is called the *remote past* of the process $X(u)$.

It is clear that

$$H_{-\infty}(X) \subset \cdots \subset H_{-\infty}^{t}(X) \subset H_{-\infty}^{t+\tau}(X) \subset \cdots \subset H(X), \quad \tau \in \mathbb{N}.$$

The Hilbert space setting provides a natural framework for stating and solving the problem of predicting future values of the process $X(u)$ from the observed past values. Assume that a realization of the process $X(u)$ for times $u \leq t$ is observed and we want to predict the value $X(t+\tau)$ for some $\tau \geq 1$ from the observed values. Since we will never know what particular realization is being observed, it is reasonable to consider as a predictor $\widehat{X}(t,\tau)$ for $X(t+\tau)$ a function of the observed values, $g(\{X(u), u \leq t\})$, which is good "on the average". So, as an optimality criterion for our predictor we take the L^2-distance, that is, the mean squared error, and consider only the linear predictors. With these restrictions, the infinite linear prediction problem can be stated as follows.

The infinite linear prediction problem. Given a "parameter" of the process $X(u)$ (e.g., the covariance function $r(t)$ or the spectral function $F(\lambda)$), the entire history $H_{-\infty}^{t}(X)$ of $X(u)$, and a natural number $\tau \in \mathbb{N}$, find a random variable $\widehat{X}(t,\tau)$ such that

a) $\widehat{X}(t,\tau)$ is *linear*, that is, $\widehat{X}(t,\tau) \in H_{-\infty}^{t}(X)$,
b) $\widehat{X}(t,\tau)$ is *mean-square optimal (best)* among all elements $Y \in H_{-\infty}^{t}(X)$, that is, $\widehat{X}(t,\tau)$ minimizes the mean-squared error $||X(t+\tau)-Y||^2_{L^2(P)}$:

$$||X(t+\tau)-\widehat{X}(t,\tau)||^2_{L^2(P)} = \min_{Y \in H_{-\infty}^{t}(X)} ||X(t+\tau)-Y||^2_{L^2(P)}. \tag{1.13}$$

The solution - the random variable $\widehat{X}(t,\tau)$ satisfying a) and b), is called the *best linear τ-step ahead predictor* for an element $X(t+\tau) \in H(X)$. The quantity

$$\sigma^2(\tau) = ||X(t+\tau)-\widehat{X}(t,\tau)||^2_{L^2(P)} = ||X(t+\tau)||^2_{L^2(P)} - ||\widehat{X}(t,\tau)||^2_{L^2(P)}, \tag{1.14}$$

which is independent of t, is called the *prediction error (variance).*

The advantage of the Hilbert space setting now becomes apparent. Namely, by the *projection theorem* in Hilbert spaces (see [29], p. 312), such a predictor exists, is unique, and is given by

$$\widehat{X}(t,\tau) = P_t X(t+\tau), \tag{1.15}$$

where $P_t := P_{(-\infty,t]}$ is the orthogonal projection operator in $H(X)$ onto $H_{-\infty}^{t}(X)$.

Remark 1.1. The reason for restricting attention to linear predictors is that the best linear predictor $\widehat{X}(t,\tau)$, in this case, depends only on knowledge of the covariance function $r(t)$ or the spectral function $F(\lambda)$. The prediction problem becomes much more difficult when nonlinear predictors are allowed.

1.4. **Deterministic and nondeterministic processes.** From prediction point of view it is natural to distinguish the class of processes for which we have *error-free prediction*, that is, $\sigma^2(\tau) = 0$ for all $\tau \geq 1$, or equivalently, $\widehat{X}(t, \tau) = X(t + \tau)$ for all $t \in \mathbb{Z}$ and $\tau \geq 1$. In this case, the prediction is called *perfect*. It is clear that a process $X(t)$ possessing perfect prediction represents a singular case of *extremely strong dependence* between the random variables forming the process. Such a process $X(t)$ is called *deterministic* or *singular*. From the physical point of view, singular processes are exceptional. From application point of view it is of interest the class of processes for which we have $\sigma^2(\tau) > 0$ for all $\tau \geq 1$. In this case the prediction is called *imperfect*, and the process $X(t)$ is called *nondeterministic*.

Observe that the time-domain $H(X)$ of any non-degenerate stationary process $\{X(t),\ t \in \mathbb{Z}\}$ can be represented as the orthogonal sum $H(X) = H_1(X) \oplus H_{-\infty}(X)$, where $H_{-\infty}(X)$ is the remote past of $X(t)$ defined by (1.12), and $H_1(X)$ is the orthogonal complement of $H_{-\infty}(X)$. So, we can give the following geometric definition of the deterministic (singular), nondeterministic and purely nondeterministic (regular) processes.

Definition 1.4. A stationary process $\{X(t),\ t \in \mathbb{Z}\}$ is called

- *deterministic or singular* if $H_{-\infty}(X) = H(X)$, that is, $H^t_{-\infty}(X) = H^s_{-\infty}(X)$ for all $t, s \in \mathbb{Z}$,
- *nondeterministic* if $H_{-\infty}(X)$ is a proper subspace of $H(X)$, that is, $H_{-\infty}(X) \subset H(X)$,
- *purely nondeterministic (PND) or regular* if $H_{-\infty}(X) = \{0\}$, that is, the remote past $H_{-\infty}(X)$ of $X(t)$ is the trivial subspace, consisting of the singleton zero.

The next result, known as Wold's decomposition theorem (see [1], p. 65), provides a key step for solution of the infinite prediction problem in the time-domain setting, and essentially says that any stationary process can be represented in the form of a sum of two orthogonal stationary components, one of which is perfectly predictable (singular component), while for the other (regular component) an explicit formula for the predictor can be obtained.

Theorem 1.1 (Wold's decomposition)**.** *Every centered non-degenerate discrete-time stationary process $X(t)$ admits a decomposition*

$$X(t) = X_S(t) + X_R(t), \tag{1.16}$$

where

(a) *the processes* $X_R(t)$ *and* $X_S(t)$ *are stationary, centered, mutually uncorrelated (orthogonal), and completely subordinated to* $X(t)$*, that is,* $H^t_{-\infty}(X_R) \subseteq H^t_{-\infty}(X)$ *and* $H^t_{-\infty}(X_S) \subseteq H^t_{-\infty}(X)$ *for all* $t \in \mathbb{Z}$*.*

(b) *the process* $X_S(t)$ *is deterministic (singular),*

(c) *the process* $X_R(t)$ *is purely nondeterministic (regular) and has the infinite moving-average representation:*

$$X_R(t) = \sum_{k=0}^{\infty} b_k \varepsilon_0(t-k), \quad \sum_{k=0}^{\infty} |b_k|^2 < \infty, \tag{1.17}$$

where $\varepsilon_0(t)$ *is an innovation of* $X_R(t)$*, that is,* $\varepsilon_0(t)$ *is a standard white noise process, such that* $H^t_{-\infty}(X_R) = H^t_{-\infty}(\varepsilon_0)$ *for all* $t \in \mathbb{Z}$*.*

(d) *the representation (1.16) is unique.*

The next theorem contains spectral characterizations of deterministic, nondeterministic and purely nondeterministic processes (see [24], p. 35-36, [32], p. 58, 64)).

Theorem 1.2. *Let* $X(t)$ *be a discrete-time non-degenerate stationary process with spectral function* $F(\lambda) = F_R(\lambda) + F_S(\lambda) = \int_{-\pi}^{\lambda} f(u)du + F_S(\lambda)$*. The following assertions hold.*

(a) (Kolmogorov-Szegö alternative). *Either*

$$H^0_{-\infty}(F_R) = H(F_R) \Leftrightarrow \int_{-\pi}^{\pi} \log f(\lambda)\, d\lambda = -\infty \Leftrightarrow \sigma^2(f) = 0 \Leftrightarrow X(t) \textit{ is deterministic},$$

or else

$$H^0_{-\infty}(F_R) \neq H(F_R) \Leftrightarrow \int_{-\pi}^{\pi} \log f(\lambda)\, d\lambda > -\infty \Leftrightarrow \sigma^2(f) > 0 \Leftrightarrow X(t) \textit{ is nondeterministic}.$$

(b) *The process* $X(t)$ *is regular (PND) if and only if it is nondeterministic and* $F_S(\lambda) \equiv 0$*.*

Remark 1.2. The condition

$$\int_{-\pi}^{\pi} \log f(\lambda)\, d\lambda > -\infty \tag{1.18}$$

is called *Szegö condition*. Observe that (1.18) is satisfied if and only if $\log f \in L^1(\Lambda)$, since $\log f(\lambda) \leq f(\lambda)$ and $f(\lambda) \in L^1(\Lambda)$. Also, the Szegö condition (1.18) is connected with the character of zeros of the spectral density $f(\lambda)$, and does not depend on the differential properties of $f(\lambda)$. For example, for any $\alpha \geq 0$ the function $f(\lambda) = \exp\{-|\lambda|^{-\alpha}\}$ is infinitely differentiable, for $\alpha < 1$ the Szegö condition is satisfied, and hence a stationary process $X(t)$ with this spectral density is nondeterministic, while for $\alpha \geq 1$ the Szegö condition is violated, and $X(t)$ is deterministic (see [30], p. 151, [29], p. 68).

Remark 1.3. A stationary process $X(t)$ is deterministic if either it has pure discrete spectrum, or pure singular spectrum, or the Szegö condition is violated: $\log f \notin L^1(\Lambda)$. Thus, for $X(t)$ to be nondeterministic, its spectral density $f(\lambda)$ cannot be zero too often (see [29], p. 68).

2. The Finite Prediction Problem

In practice we never will have the observed entire infinite past, instead will be available only the finite past.

Suppose we have observed the values $X(-n), \dots, X(-1)$ of a centered, real-valued stationary process $X(t)$ with covariance function $r(t)$ and spectral function $F(\lambda)$, the *one-step ahead finite prediction problem* in predicting a random variable $X(0)$ based on the observed values $X(-n), \dots, X(-1)$ is: find the orthogonal projection $\widehat{X}_n(0) = P_{[-n,-1]}X(0)$ of $X(0)$ onto the space $H_n(X) = H_{-n}^{-1}(X) = sp\{X(t),\ -n \le t \le -1\}$, that is, find constants $\hat{c}_k = \hat{c}_{k,n}$, $k = 1, 2, \dots, n$, that minimize the one-step ahead prediction error variance $\sigma_n^2(F) = \sigma_n^2(1, F)$:

$$\begin{aligned} \sigma_n^2(F) &= \min_{\xi \in H_n(X)} \|X(0) - \xi\|^2_{L^2(P)} = \min_{\{c_k\}} \left\| X(0) - \sum_{k=1}^{n} c_k X(-k) \right\|^2_{L^2(P)} \\ &= \left\| X(0) - \sum_{k=1}^{n} \hat{c}_k X(-k) \right\|^2_{L^2(P)} = ||X(0) - \widehat{X}_n(0)||^2_{L^2(P)}. \end{aligned} \tag{2.1}$$

If such constants $\hat{c}_k$ can be found, then the best linear 1-step ahead predictor $\widehat{X}_n(0)$ of a random variable $X(0)$ based on the observed finite past: $X(-n), \dots, X(-1)$ can be computed by

$$\widehat{X}_n(0) = \sum_{k=1}^{n} \hat{c}_k X(-k), \quad \hat{c}_k = \hat{c}_{k,n}, \tag{2.2}$$

and the mean-squared prediction error $\sigma_n^2(F)$ can be computed by formula (2.1).

Using Kolmogorov's isometric isomorphism $V : X(t) \leftrightarrow e^{it\lambda}$ between the time- and frequency-domains $H(X)$ and $L^2(F)$, in view of (2.1), for $\sigma_n^2(F)$ we can write

$$\begin{aligned} \sigma_n^2(F) &= \min_{\{c_k\}} \left\| X(0) - \sum_{k=1}^{n} c_k X(-k) \right\|^2_{L^2(P)} = \min_{\{c_k\}} \int_{-\pi}^{\pi} \left| 1 - \sum_{k=1}^{n} c_k e^{-ik\lambda} \right|^2 dF(\lambda) \\ &= \min_{\{c_k\}} \int_{-\pi}^{\pi} \left| e^{in\lambda} - \sum_{k=1}^{n} c_k e^{i(n-k)\lambda} \right|^2 dF(\lambda) = \min_{\{q_n \in \mathcal{Q}_n\}} \int_{-\pi}^{\pi} \left| q_n(e^{i\lambda}) \right|^2 dF(\lambda), \end{aligned} \tag{2.3}$$

where $\mathcal{Q}_n = \left\{ q_n : q_n(z) = \sum_{k=0}^{n} c_k z^{n-k},\ c_0 = 1 \right\}$ stands for the set of polynomials of degree n with coefficient of the leading term equal to 1.

Thus, the problem of finding $\sigma_n^2(F)$ becomes to the solution of the following minimum problem:

$$\int_{-\pi}^{\pi} |q_n(e^{i\lambda}|^2 dF(\lambda) = \min, \quad q_n(z) \in \mathcal{Q}_n. \tag{2.4}$$

The polynomial $p_n(z) = p_n(z, F)$ that solves the minimum problem (2.4) is called the *optimal polynomial* for $F(\lambda)$ in the class $\mathcal{Q}_n$. This minimum problem was solved by G. Szegö (see [18], Section 2.2) by showing that the optimal polynomial $p_n(z, F)$ exists, is unique and can be expressed in terms of orthogonal polynomials $\varphi_n(z)$, $n \in \mathbb{Z}_+ = \{0, 1, 2, \ldots\}$, on the unit circle $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ with respect to $F(\lambda)$.

Recall that the system of orthogonal polynomials $\{\varphi_n(z) = \varphi_n(z; F),\ z = e^{i\lambda},\ n \in \mathbb{Z}_+\}$ is uniquely determined by the following conditions:

(i) $\varphi_n(z) = \kappa_n(F) z^n +$ lower order terms
is a polynomial of degree n, in which the coefficient $\kappa_n = \kappa_n(F)$ is real and positive;

(ii) for arbitrary nonnegative integers k and j

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi_k(z) \overline{\varphi_j(z)} dF(\lambda) = \delta_{kj} = \begin{cases} 1, & \text{for } k = j \\ 0, & \text{for } k \neq j, \end{cases} \quad z = e^{i\lambda}.$$

Theorem 2.1 (Szegö theorem)**.** *The optimal polynomial for $F(\lambda)$ in the class $\mathcal{Q}_n$, that is, the polynomial $p_n(z) = p_n(z, F)$ that solves the minimum problem* (2.4) *is given by $p_n(z) = \kappa_n^{-1}(F)\varphi_n(z)$, and the minimum itself is equal to $\kappa_n^{-2}(F)$. Thus, we have*

$$\sigma_n^2(F) = \min_{\{q_n \in \mathcal{Q}_n\}} \int_{-\pi}^{\pi} \left|q_n(e^{i\lambda})\right|^2 dF(\lambda) = \tag{2.5}$$

$$= \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda}, F)\right|^2 dF(\lambda) = \kappa_n^{-2}(F).$$

Remark 2.1. Denote $\mathcal{Q}_n^* = \left\{q_n : q_n(z) = \sum_{k=0}^n c_k z^{n-k},\ c_n = 1\right\}$. Then we have (see [18], Section 3.1):

$$\sigma_n^2(F) = \min_{\{q_n \in \mathcal{Q}_n^*\}} \int_{-\pi}^{\pi} \left|q_n(e^{i\lambda})\right|^2 dF(\lambda) = \int_{-\pi}^{\pi} \left|p_n^*(e^{i\lambda}, F)\right|^2 dF(\lambda), \tag{2.6}$$

where $p_n^*(z) = p_n(z, F)$ is the optimal polynomial for $F(\lambda)$ in the class $\mathcal{Q}_n^*$.

Remark 2.2. From the obvious embedding $\mathcal{Q}_n^* \subset \mathcal{Q}_{n+1}^*$, it follows that the sequence $\{\sigma_n^2(F),\ n \in \mathbb{N}\}$ is non-increasing in n: $\sigma_{n+1}^2(F) \leq \sigma_n^2(F)$. Also, it follows from (2.5) that $\sigma_n^2(F)$ is a non-decreasing functional of $F(\lambda)$:

$$\sigma_n^2(F_1) \leq \sigma_n^2(F_2) \quad \text{when} \quad F_1(\lambda) \leq F_2(\lambda), \quad \lambda \in \Lambda. \tag{2.7}$$

Indeed, by the definition of optimal polynomials $p_n(z,F_1)$ and $p_n(z,F_2)$, corresponding to spectral functions F_1 and F_2, respectively, we have

$$\begin{aligned}\sigma_n^2(F_1) &= \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda},F_1)\right|^2 dF_1(\lambda) \le \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda},F_2)\right|^2 dF_1(\lambda) \\ &\le \int_{-\pi}^{\pi} \left|p_n(e^{i\lambda},F_2)\right|^2 dF_2(\lambda) = \sigma_n^2(F_2).\end{aligned}$$

The *finite prediction problem* is to describe the asymptotic behavior of $\sigma_n^2(F)$ as the length of the observed past increases ($n\to\infty$). The problem was solved by G. Szegö in 1915 in the special case where $F(\lambda)$ is pure absolute continuous, that is, $F_S(\lambda)=0$, and by A. Kolmogorov in 1941 in the general case (see, e.g., [18], p. 44 or [22], p. 49). The solution is given in the theorem that follows, known as Kolmogorov-Szegö theorem.

Remark 2.3. If $F(\lambda)$ is purely absolutely continuous, that is, $dF(\lambda)=f(\lambda)d\lambda$, then instead of $\sigma_n^2(F)$ and $\sigma^2(F)$ we will write $\sigma_n^2(f)$ and $\sigma^2(f)$, respectively.

Theorem 2.2 (Kolmogorov-Szegö theorem)**.** *For any non-trivial spectral function $F(\lambda)$ the following limiting relation hold:*

$$\lim_{n\to\infty}\sigma_n^2(F)=\sigma^2(F)=\sigma^2(f)=2\pi G(f), \tag{2.8}$$

where $f(\lambda)$ is the spectral density, that is, the derivative of the absolutely continuous part of $F(\lambda)$, and $G(f)$ is the geometric mean of $f(\lambda)$, given by

$$G(f)=\begin{cases}\exp\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\log f(\lambda)\,d\lambda\right\} & \text{if } \log f\in L^1(\Lambda)\\ 0, & \text{otherwise.}\end{cases} \tag{2.9}$$

Define the *relative prediction error* $\delta_n(F)$ to be

$$\delta_n(F):=\sigma_n^2(F)-\sigma^2(F). \tag{2.10}$$

Observe that $\delta_n(F)\ge 0$ and $\delta_n(F)\to 0$ as $n\to\infty$. Note that if the underlying process $X(t)$ is deterministic, then $\delta_n(F)=\sigma_n^2(F)$.

The problem of interest is to describe the rate of decrease of relative prediction error $\delta_n(F)$ to zero as $n\to\infty$, depending on the regularity nature (deterministic or nondeterministic) and the dependence (memory) structure of the model $X(t)$. This problem we discuss in Section 3 for nondeterministic processes and in Section 4 for deterministic processes.

3. Asymptotic behavior of the prediction error variance for nondeterministic processes

In this section we study the asymptotic behavior of the finite prediction error for nondeterministic processes, and review some important known results.

We assume that the model process $X(t)$ is regular, or equivalently, is purely nondeterministic (PND), that is, $X(t)$ has a non-trivial spectral function $F(\lambda) = \int_{-\pi}^{\lambda} f(u)du + F_S(\lambda)$ with $dF_s(\lambda) = 0$ and $\ln f(\lambda) \in L^1(\Lambda)$, and describe the rate of decrease of relative prediction error $\delta_n(F)$ to zero as $n \to \infty$, depending on the dependence (memory) structure of the model $X(t)$ and the smoothness properties of its spectral density $f(\lambda)$.

3.1. **Asymptotic behavior of $\delta_n(f)$ for short-memory processes.** Recall that a *short memory* processes is a second order stationary processes possessing a spectral density $f(\lambda)$ which is bounded away from zero and infinity, that is, there are constants m and M such that $0 < m \le f(\lambda) \le M < \infty$. A typical short memory model example is the stationary autoregressive moving average $(\mathrm{ARMA})(p, q)$ process $X(t)$ defined to be a stationary solution of the difference equation: $\psi_p(B)X(t) = \theta_q(B)\varepsilon(t)$, $t \in \mathbb{Z}$, where ψ_p and θ_q are polynomials of degrees p and q, respectively, B is the backward shift operator defined by $BX(t) = X(t-1)$, and $\{\varepsilon(t), t \in \mathbb{Z}\}$ is a discrete-time white noise, that is, a sequence of zero-mean, uncorrelated random variables.

We first give a result that contains a necessary and sufficient condition for exponential rate of decrease to zero for $\delta_n(f) = \sigma_n^2(f) - \sigma^2(f)$. Notice that the first result of this type goes back to the paper by Grenander and Rosenblatt [17]. The next theorem was proved by Ibragimov [23] (see also Golinskii and Ibragimov [15]).

Theorem 3.1. *A necessary and sufficient condition for*

$$\delta_n(f) = O(q^n), \quad q = e^{-b}, \quad b > 0, \quad n \to \infty \tag{3.1}$$

is that $f(\lambda)$ is a spectral density of a short-memory process, and $1/f(\lambda) \in A_b$, where A_b is the class of 2π–periodic continuous functions $\varphi(\lambda)$, $\lambda \in \mathbb{R}$, admitting an analytic continuation into the strip $z = \lambda + i\mu$, $-\infty < \lambda < \infty$, $|\mu| \le b$.

Observe that (3.1) will be true for all $b > 0$ if and only if the analytic continuation of $f(\lambda)$ is an entire function of $z = \lambda + i\mu$.

Thus, to have exponential rate of decrease to zero for $\delta_n(f)$ the underlying model should be short-memory process with sufficiently smooth spectral density $f(\lambda)$.

Now we give a result that contains a necessary and sufficient condition for hyperbolic rate of decrease to zero for $\delta_n(f)$:

$$\delta_n(f) = O(n^{-\gamma}), \quad \gamma > 0, \quad n \to \infty. \tag{3.2}$$

Bounds of type (3.2) with $\gamma > 1$ for different classes of spectral densities were obtained by Baxter [2], Devinatz [9], Geronimus [12], Grenander and Rosenblatt

[17], Grenander and Szegö [18], and others (see [13] and references therein). The most general result in this direction has been obtained by Ibragimov [23]. To state Ibragimov's theorem, we first introduce the Hölder class of functions.

For a function $\varphi(\lambda) \in L^p(\Lambda)$, we define its L^p-modulus of continuity by

$$\omega_p(\varphi; \delta) = \sup_{0<|t|\leq\delta} ||\varphi(\cdot + t) - \varphi(\cdot)||_p, \quad \delta > 0. \tag{3.3}$$

Given numbers $0 < \alpha < 1$, $r \in \mathbb{Z}_+ := \{0, 1, 2, \ldots\}$, and $p \geq 1$, we put $\gamma := r + \alpha$. The Hölder class of functions, denoted by $H_p(\gamma)$, is defined to be the set of those functions $\varphi(\lambda) \in L^p(\Lambda)$ that have r-th derivative $\varphi^{(r)}(\lambda)$, such that $\varphi^{(r)}(\lambda) \in L^p(\Lambda)$ and $\omega_p(\varphi^{(r)}; \delta) = O(\delta^\alpha)$ as $\delta \to 0$.

Theorem 3.2. *A necessary and sufficient condition for*

$$\delta_n(f) = O(n^{-\gamma}), \ \gamma = 2(r + \alpha) > 1; \ 0 < \alpha < 1, \ r \in \mathbb{Z}_+, \text{ as } n \to \infty \tag{3.4}$$

is that $f(\lambda)$ is a spectral density of a short-memory process belonging to $H_2(\gamma)$.

Remark 3.1. It follows from Theorem 3.2 that if $\delta_n(f) = O(n^{-\gamma})$ with $\gamma > 1$, then the underlying model $X(t)$ is necessarily a short-memory process. Moreover, as it was pointed out by Grenander and Rosenblatt [17] (see, also, Devinatz [9], p. 118), if the model is not a short-memory process, that is, the spectral density $f(\lambda)$ has zeros or is unbounded, then, in general, we cannot expect $\delta_n(f)$ to go to zero faster than $1/n$ as $n \to \infty$. This question we discuss in the next subsection.

3.2. **Asymptotic behavior of $\delta_n(f)$ for long memory and antipersistent processes.** Recall that a second order stationary process $X(t)$ is said to be *anti-persistent* if the spectral density $f(\lambda)$ vanishes at frequency zero: $f(0) = 0$. And, we say that $X(t)$ displays *long memory* or *long-range dependence* if the spectral density $f(\lambda)$ has a pole at frequency zero, that is, it is unbounded at the origin.

A well-known example of processes that displays long memory or is anti-persistent is an autoregressive fractionally integrated moving average ARFIMA(p, d, q) process $X(t)$ defined to be a stationary solution of the difference equation:

$$\psi_p(B)(1 - B)^d X(t) = \theta_q(B)\varepsilon(t), \quad d < 1/2,$$

where B is the backward shift operator, $\varepsilon(t)$ is a discrete-time white noise, and ψ_p and θ_q are polynomials of degrees p and q, respectively. The spectral density $f(\lambda)$ of $X(t)$ is given by

$$f(\lambda) = |1 - e^{-i\lambda}|^{-2d} h(\lambda), \quad d < 1/2, \tag{3.5}$$

where $h(\lambda)$ is the spectral density of an ARMA(p,q) process. Note that the condition $d < 1/2$ ensures that $\int_{-\pi}^{\pi} f(\lambda)d\lambda < \infty$, implying that the process $X(t)$ is well defined because $E|X(t)|^2 = \int_{-\pi}^{\pi} f(\lambda)d\lambda$.

Observe that for $0 < d < 1/2$ the model $X(t)$ specified by (3.5) displays long-memory, for $d < 0$ it is anti-persistent, and for $d = 0$ it displays short-memory. For $d \geq 1/2$ the function $f(\lambda)$ in (3.5) is not integrable, and thus it cannot represent a spectral density of a stationary process (see Brockwell and Davis [7], Section 13.2). The following theorem was proved by A. Inoue (see [26], Theorem 4.3).

Theorem 3.3. *Let $f(\lambda)$ have the form (3.5) with $0 < d < 1/2$, where $h(\lambda)$ is the spectral density of an ARMA(p,q) process. Then*

$$\delta_n(f) \sim \frac{d^2}{n} \quad \text{as} \quad n \to \infty. \tag{3.6}$$

Another well-known example of processes that displays long memory or is anti-persistent is the Jacobian model. We say that a stationary process $X(t)$ is a Jacobian process, and the corresponding model is a Jacobian model, if its spectral density $f(\lambda)$ has the following form:

$$f(\lambda) = h(\lambda) \prod_{k=1}^{m} |e^{i\lambda} - e^{i\lambda_k}|^{-2d_k}, \tag{3.7}$$

where $h(\lambda)$ is the spectral density of a short-memory process, the points $\lambda_k \in [-\pi, \pi]$ are distinct, and $d_k < 1/2$, $k = 1, \ldots, m$.

The asymptotic behavior of $\delta_n(f)$ as $n \to \infty$ for Jacobian model (3.7) has been considered in a number of papers (see Golinskii [14], Grenander and Rosenblatt [17], Ibragimov [23], Ibragimov and Solev [25].)

The following theorem was proved in Ibragimov and Solev [25].

Theorem 3.4. *Let $f(\lambda)$ have the form (3.7), where $h(\lambda)$ is the spectral density of a short-memory process, the points $\lambda_k \in [-\pi, \pi]$ are distinct, and $d_k < 1/2$ ($d_k \neq 0$), $k = 1, \ldots, m$. If $f(\lambda)$ satisfies the Lipschitz condition with exponent $\alpha \geq 1/2$, then*

$$\delta_n(f) \sim \frac{1}{n} \quad \text{as} \quad n \to \infty. \tag{3.8}$$

More results for this case can be found in Ginovyan [13] and in the references therein.

4. Asymptotic behavior of the predictor error for deterministic processes

4.1. **Background.** In this section we discuss the asymptotic behavior of the predictor error for deterministic processes. We assume that the process $X(t)$ possesses a spectral density $f(\lambda)$ and the Szegö condition (1.18) is violated. As it was mentioned in Introduction, this problem was first studied by M. Rosenblatt [31], where using the technique of orthogonal polynomials and Szegö's results, M. Rosenblatt has investigated the asymptotic behavior of the prediction error $\delta_n(f) = \sigma_n^2(f)$ in the following two cases:

(a) the spectral density $f(\lambda)$ is continuous and vanishes on an interval,
(b) the spectral density $f(\lambda)$ has a high order contact with zero, so that the Szegö condition is violated.

For the case (a), in [31] M. Rosenblatt proved the following result.

Theorem 4.1. *Let the spectral density $f(\lambda)$ of a discrete-time stationary process $X(t)$ be positive and continuous on the interval $(\pi/2-\alpha, \pi/2+\alpha)$, $0<\alpha<\pi$, and zero elsewhere, then the prediction error $\sigma_n^2(f)$ approaches zero exponentially as $n\to\infty$. More precisely, the following asymptotic relation holds:*

$$\delta_n(f) := \sigma_n^2(f) \cong \left(\sin\frac{\alpha}{2}\right)^{2n+1} \quad \text{as} \quad n\to\infty, \tag{4.1}$$

implying that

$$\lim_{n\to\infty} (\sigma_n(f))^{1/n} = \sin\frac{\alpha}{2}. \tag{4.2}$$

Later this result has been generalized by Babayan [3], [4] to the case of several arcs, without having to stipulate continuity of the spectral density $f(\lambda)$ (see also Davisson [8]). To state the corresponding result we first introduce the concept of a transfinite diameter of a set (see, e.g., Goluzin [16], Chapter 7). Let E be a bounded closed set in the complex plane. Denote by $T_n(z,E)$ the Chebyshev polynomial which deviates least from zero on the set E in the uniform metric. We set $C_n(E) = \max_{z\in E}|T_n(z,E)|$. Then $\lim_{n\to\infty}(C_n(E))^{1/n} =: \tau(E)$ exists and is called the *transfinite diameter* (or *Chebyshev constant*, or *capacity*) of the set E.

Remark 4.1. Notice that the transfinite diameter of the unit circle $\mathbb{T}$ is equal to 1 (see Goluzin [16], Section 7.1), and the transfinite diameter of an arc of $\mathbb{T}$ of length 2α $(0<\alpha<\pi)$ is equal to $\sin(\alpha/2)$ (see Rosenblatt [31]). Thus, the right hand side of (4.2) is the transfinite diameter of the closure of the spectrum $E_f = \{e^{i\lambda} : \lambda \in [\pi/2-\alpha, \pi/2+\alpha]\}$ of the process $X(t)$.

Using some results from geometric function theory, in [4] was proved the following theorem, extending Theorem 4.1.

Theorem 4.2. *Let the spectrum $E_f = \{e^{i\lambda} : f(\lambda) > 0\}$ of the process $X(t)$ consist of a finite number of arcs of the unit circle. Then the following asymptotic relation* *holds:*

$$\lim_{n\to\infty} (\sigma_n(f))^{1/n} = \tau(\overline{E}_f), \tag{4.3}$$

where $\overline{E}_f$ is the closure of E_f.

Remark 4.2. It follows from Theorem 4.2 and Remark 4.1 that if the spectral density $f(\lambda)$ vanish on an interval, then the prediction error $\sigma_n(f)$ decreases to zero exponentially, that is, $\sigma_n(f) = O(e^{-bn})$, $b > 0$ as $n \to \infty$. Conversely, a necessary condition for $\sigma_n(f)$ to tend to zero exponentially is that $f(\lambda)$ should vanish on a set of positive Lebesgue measure.

Concerning the case (b), in [31] M. Rosenblatt proved that if the spectral density $f(\lambda)$ of a stationary process $X(t)$ is positive away from zero, and has a very high order contact with zero at $\lambda = 0$, so that the Szegö condition (1.18) is violated, then the prediction error $\delta_n(f) = \sigma_n^2(f)$ decreases to zero hyperbolically as $n \to \infty$. More precisely, in [31] was considered a deterministic process $X(t)$ with spectral density $f_a(\lambda)$ given by formula:

$$f_a(\lambda) = \frac{e^{(2\lambda-\pi)\varphi(\lambda)}}{\cos\lambda(\pi\varphi(\lambda))}, \quad f_a(-\lambda) = f_a(\lambda), \quad 0 \le \lambda \le \pi, \tag{4.4}$$

where $\varphi(\lambda) = \frac{a}{2}\cot\lambda$ and a is a fixed positive parameter.

It is easy to show that

$$f_a(\lambda) \sim \exp\left\{-\frac{a\pi}{2|\lambda|}\right\}|\sin(\lambda)| \quad \text{as} \quad \lambda \to 0, \tag{4.5}$$

so that $f_a(\lambda)$ has a very high order contact with zero only at $\lambda = 0$.

In [31], using the formula (2.5) and the technique of orthogonal polynomials on the unit circle, M. Rosenblatt proved the following result.

Theorem 4.3. *For a process $X(t)$ with spectral density $f_a(\lambda)$ given by* (4.4) *the following asymptotic formula for prediction error $\delta_n(f) = \sigma_n^2(f)$ holds:*

$$\delta_n(f_a) = \sigma_n^2(f_a) \cong \frac{\Gamma^2\left(\frac{a+1}{2}\right)}{\pi 2^{2-a}}\; n^{-a} \sim n^{-a} \quad \text{as} \quad n \to \infty. \tag{4.6}$$

In the next subsection we extend Theorem 4.3 to more broad class of spectral densities.

4.2. **The main results.** In this subsection, we analyze the asymptotic behavior of the prediction error in the case where the spectral density $f(\lambda)$ of the model has a high order contact with zero, so that the Szegö condition (1.18) is violated.

Based on Rosenblatt's result for this case - Theorem 4.3, we can expect that for any deterministic process with spectral density possessing a zero of type (4.5), the rate of prediction error $\sigma_n^2(f)$ should be the same as in (4.6). However, the method applied in [31] does not allow to prove this assertion. Here, using a different approach, we extend Rosenblatt's theorem to more broad class of spectral densities.

To this end, we first examine the asymptotic behavior of the ratio $\sigma_n(fg)/\sigma_n(f)$ as $n \to \infty$, where $g(\lambda)$ is some nonnegative function, such that the product $f(\lambda)g(\lambda)$ is a spectral density, that is, $fg \in L^1(\Lambda)$.

To make the approach clear, we first assume that $f(\lambda)$ is a spectral density of a nondeterministic process, in which case the geometric mean $G(f)$ is positive (see (2.8) and (2.9)). Then, in this case, we can write

$$\lim_{n\to\infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = \frac{\sigma_\infty^2(fg)}{\sigma_\infty^2(f)} = \frac{G(fg)}{G(f)} = G(g). \tag{4.7}$$

It turns out that under some additional assumptions imposed on functions f and g, the asymptotic relation (4.7) remains valid also in the case of deterministic process, that is, when $\sigma_\infty^2(f) = 0$, or equivalently, $G(f) = 0$.

To state the corresponding results we need some definitions.

Definition 4.1. *A sequence of numbers $\{a_n,\, n \in \mathbb{N}\}$ is said to be slowly decreasing if*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = 1. \tag{4.8}$$

It is easy to check that the following simple assertions hold:

1. If $\{a_n,\, n \in \mathbb{N}\}$ is a slowly decreasing sequence, then for any $\nu \in \mathbb{N}$

$$\lim_{n\to\infty} \frac{a_{n+\nu}}{a_n} = 1. \tag{4.9}$$

2. If $\{a_n,\, n \in \mathbb{N}\}$ is a sequence such that $a_n \to a \neq 0$ as $n \to \infty$, then $\{a_n\}$ is a slowly decreasing sequence.
3. If $\{a_n,\, n \in \mathbb{N}\}$ and $\{b_n,\, n \in \mathbb{N},\}$ are non-zero slowly decreasing sequences, then $ca_n, c \neq 0$, $1/a_n$, $a_n^k, k \in \mathbb{N}$, a_nb_n and a_n/b_n also are slowly decreasing sequences.
4. If $\{a_n,\, n \in \mathbb{N}\}$ is a non-zero slowly decreasing sequence, and $\{b_n,\, n \in \mathbb{N}\}$ is a sequence such that

$$\lim_{n\to\infty} \frac{b_n}{a_n} = c \neq 0, \tag{4.10}$$

then $\{b_n,\, n \in \mathbb{N}\}$ is also a slowly decreasing sequence.
5. If $\{a_n,\, n \in \mathbb{N}\}$ is a slowly decreasing sequence of nonnegative numbers, then

$$\lim_{n\to\infty} (a_n)^{1/n} = 1. \tag{4.11}$$

Remark 4.3. It follows from assertion 2 that the notion of slowly decreasing sequence is more significant in the case where $a_n \to 0$ as $n \to \infty$. Also, it follows from assertion 5 that if $\{a_n,\, n \in \mathbb{N}\}$ is a slowly decreasing sequence of nonnegative numbers such that $a_n \to 0$ as $n \to \infty$, then it converges to zero slowly than the geometric progression $\{q^n, n \in \mathbb{N}\}$ for any q, $0 < q < 1$, that is, $q^n = o(a_n)$ as $n \to \infty$.

In what follows we consider the class of processes for which the sequence of prediction errors $\{\sigma_n(f)\}$ is slowly decreasing. Moreover, in view of Remarks 4.2 and 4.3, it is reasonable to consider deterministic processes except those for which the spectral densities vanish on an interval.

Definition 4.2. *We define the class A to be the set of all nonnegative, Riemann integrable functions $h(\lambda)$, $\lambda \in \Lambda$. Also, define $A_+ = \{h \in A : h(\lambda) \geqslant m > 0\}$, $A^- = \{h \in A : h(\lambda) \leqslant M < \infty\}$, and $A_+^- = A_+ \cap A^-$.*

Now we are in position to state the main results of this paper.

The following theorem describes the asymptotic behavior of the ratio $\sigma_n(fg)/\sigma_n(f)$ as $n \to \infty$ for the class of above described processes.

Theorem 4.4. *Let the spectral density $f(\lambda)$ be such that the sequence $\{\sigma_n(f)\}$ is slowly decreasing, and let $g(\lambda) = h(\lambda) \cdot \frac{t_1(\lambda)}{t_2(\lambda)}$, where $h(\lambda) \in A_+^-$ and $t_1(\lambda)$, $t_2(\lambda)$ are nonnegative trigonometric polynomials. If $f(\lambda)g(\lambda) \in A$, then*

$$\lim_{n\to\infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(g), \tag{4.12}$$

where $G(g)$ is the geometric mean of $g(\lambda)$.

The next theorem extends Rosenblatt's Theorem 4.3.

Theorem 4.5. *Let $f(\lambda) = f_a(\lambda)g(\lambda)$, where $f_a(\lambda)$ is defined by (4.4) and $g(\lambda)$ satisfies the assumptions of Theorem 4.4. Then*

$$\delta_n(f) = \sigma_n^2(f) \cong \frac{\Gamma^2\left(\frac{a+1}{2}\right) G(g)}{\pi 2^{2-a}}\, n^{-a} \sim n^{-a} \quad \text{as} \quad n \to \infty, \tag{4.13}$$

where $G(g)$ is the geometric mean of $g(\lambda)$.

4.3. **Auxiliary lemmas.** To prove the theorems, we first establish a number of lemmas.

Lemma 4.1. *Assume that the sequence $\sigma_n(f)$ is slowly decreasing, that is,*

$$\lim_{n\to\infty} \frac{\sigma_{n+1}(f)}{\sigma_n(f)} = 1. \tag{4.14}$$

Then for any nonnegative trigonometric polynomial $t(\lambda)$ we have

$$\liminf_{n\to\infty} \frac{\sigma_n^2(ft)}{\sigma_n^2(f)} \geqslant G(t), \tag{4.15}$$

where $G(t)$ is the geometric mean of $t(\lambda)$.

Proof. Let the polynomial $t(\lambda)$ be of degree ν. Then by Fejér-Riesz theorem (see [18], Section 1.12), there exists an algebraic polynomial $s_\nu(z)$ of degree ν in $z \in \mathbb{C}$, such that

$$t(\lambda) = |s_\nu(e^{i\lambda})|^2, \quad s_\nu(z) \neq 0, \quad |z| < 1. \tag{4.16}$$

Observing that $\ln|s_\nu(z)|^2$ is a harmonic function, we have

$$\ln|s_\nu(0)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \ln|s_\nu(e^{i\lambda})|^2 d\lambda,$$

implying that

$$|s_\nu(0)|^2 = G(t) > 0. \tag{4.17}$$

Let $p_n^*(z, ft)$ be the optimal polynomial of degree n for $f(\lambda)t(\lambda)$ in the class $\mathcal{Q}_n^*$ (see formula (2.6)). We set

$$r_{n+\nu}(z) = \frac{p_n^*(z, ft)s_\nu(z)}{s_\nu(0)}, \tag{4.18}$$

and observe that $r_{n+\nu}(z) \in \mathcal{Q}_{n+\nu}^*$, and

$$\int_{-\pi}^{\pi} |r_{n+\nu}(e^{i\lambda})|^2 f(\lambda) d\lambda \geqslant \int_{-\pi}^{\pi} |p_{n+\nu}^*(e^{i\lambda}, f)|^2 f(\lambda) d\lambda. \tag{4.19}$$

Therefore, in view of (4.16), (4.18) and (4.19),we can write

$$\sigma_n^2(ft) = \int_{-\pi}^{\pi} |p_n^*(e^{i\lambda}, ft)|^2 f(\lambda)t(\lambda) d\lambda = \int_{-\pi}^{\pi} |p_n^*(e^{i\lambda}, ft)s_\nu(e^{i\lambda})|^2 f(\lambda) d\lambda.$$

$$= |s_\nu(0)|^2 \int_{-\pi}^{\pi} |r_{n+\nu}(e^{i\lambda})|^2 f(\lambda) d\lambda \geqslant |s_\nu(0)|^2 \int_{-\pi}^{\pi} |p_{n+\nu}^*(e^{i\lambda}, f)|^2 f(\lambda) d\lambda = |s_\nu(0)|^2 \sigma_{n+\nu}^2(f),$$

which, in view of (4.17), implies that

$$\liminf_{n\to\infty} \frac{\sigma_n^2(ft)}{\sigma_{n+v}^2(f)} \geqslant |s_\nu(0)|^2 = G(t). \tag{4.20}$$

Now, taking into account (4.14) and (4.9), from (4.20) we obtain (4.15). $\square$

Lemma 4.2. *Let the sequence $\sigma_n(f)$ satisfy (4.14), and let $t(\lambda)$ be a nonnegative trigonometric polynomial such that the function $f(\lambda)/t(\lambda) \in A$. Then the following inequality holds:*

$$\limsup_{n\to\infty} \frac{\sigma_n^2(f/t)}{\sigma_n^2(f)} \leqslant G(1/t). \tag{4.21}$$

Proof. Let the polynomial $s_\nu(z)$ be as in (4.16), and let $p_n^*(z, f/t)$ be the optimal polynomial of degree n for $f(\lambda)/t(\lambda)$ in the class $\mathcal{Q}_n^*$ (see formula (2.6)). For $n > \nu$ we set

$$r_n(z) = \frac{p_{n-\nu}^*(z, f)s_\nu(z)}{s_\nu(0)},$$

and observe that $r_n(z) \in \mathcal{Q}_n^*$. Therefore, we have

$$\begin{aligned}\sigma_n^2(f/t) &= \int_{-\pi}^{\pi} |p_n^*(e^{i\lambda}, f/t)|^2 f(\lambda)/t(\lambda) d\lambda \le \int_{-\pi}^{\pi} |r_n(e^{i\lambda})|^2 f(\lambda)/t(\lambda) d\lambda \\ &= \frac{1}{|s_\nu(0)|^2} \int_{-\pi}^{\pi} |p_{n-\nu}^*(e^{i\lambda}, f)|^2 f(\lambda) d\lambda = \frac{1}{|s_\nu(0)|^2} \sigma_{n-\nu}^2(f),\end{aligned}$$

which, in view of (4.17), implies that

$$\limsup_{n\to\infty} \frac{\sigma_n^2(f/t)}{\sigma_{n-\nu}^2(f)} \leqslant \frac{1}{|s_\nu(0)|^2} = G(1/t). \tag{4.22}$$

Finally, taking into account (4.14) and (4.9), from (4.22) we obtain (4.21). □

Lemma 4.3. *Let $h(\lambda)$ be a function from the class A_+^-. Then for any $\varepsilon > 0$ a trigonometric polynomial $t(\lambda)$ can be found to satisfy the following condition:*

$$|h - t\|_1 = \int_{-\pi}^{\pi} |h(\lambda) - t(\lambda)| d\lambda \leqslant \epsilon. \tag{4.23}$$

Moreover, if m and M are the constants from the Definition 4.2, then the polynomial $t(\lambda)$ can be chosen so that for all $\lambda \in [-\pi, \pi]$ one of the following inequalities is satisfied:

$$m - \varepsilon < t(\lambda) < h(\lambda), \tag{4.24}$$

$$h(\lambda) < t(\lambda) < M + \varepsilon. \tag{4.25}$$

Proof. We first prove the inequalities (4.23) with (4.24). Without loss of generality, we can assume that $h(-\pi) = h(\pi)$. Otherwise by changing one of these values we can make them equal as follows: $h(-\pi) = h(\pi) = \min\{h(-\pi), h(\pi)\}$.

Let $\{\lambda_i\}$ $(-\pi = \lambda_0 < \lambda_1 < \cdots < \lambda_k = \pi)$ be a partition of the segment $[-\pi, \pi]$, and let s be the Darboux lower sum corresponding to this partition:

$$s = \sum_{i=1}^{k} m_i \Delta\lambda_i,\ m_i = \inf_{\lambda\in\Delta_i} h(\lambda),\ \Delta_i = [\lambda_{i-1}, \lambda_i],\ \Delta\lambda_i = \lambda_i - \lambda_{i-1},\ i = 1, \dots, k.$$

On the segment $[-\pi, \pi]$ we define a step-function $\varphi_k(\lambda)$ corresponding to given partition as follows:

$$\varphi_k(\lambda) = \begin{cases} m_i, & \text{if } \ \lambda \in (\lambda_{i-1}, \lambda_i),\ i = 1, \dots, k-1, \\ \min\{m_{i-1}, m_i\}, & \text{if } \lambda = \lambda_i, \\ m_1 (= m_k), & \text{if } \lambda = \lambda_0 \text{ or } \lambda = \lambda_k. \end{cases}$$

It is clear that such defined function $\varphi_k(\lambda)$ satisfies the following conditions:

$$\varphi_k(\lambda) \le h(\lambda),\ \lambda \in [-\pi, \pi] \quad \text{and} \quad \int_{-\pi}^{\pi} \varphi_k(\lambda) d\lambda = s. \tag{4.26}$$

Since the function $h(\lambda)$ is integrable, for an arbitrary given $\epsilon > 0$ a partition of the segment $[-\pi, \pi]$ can be found so that the corresponding Darboux lower sum satisfies the condition:

$$\int_{-\pi}^{\pi} h(\lambda)d\lambda - s = \int_{-\pi}^{\pi}[h(\lambda)-\varphi_k(\lambda)]d\lambda = \|h-\varphi_k\|_1 < \frac{\epsilon}{3}. \tag{4.27}$$

Now using the function $\varphi_k(\lambda)$ we construct a new continuous function. To this end, we connect all the adjacent segments of the graph of $\varphi_k(\lambda)$ (the steps) by line segments as follows: for each interior point of the partition $\lambda_i, i = 1,\dots,k-1$, the endpoint of the lower step with abscissa λ_i we connect with some interior point of the adjacent (from the left or from the right) upper step, the abscissa λ_i^* of which satisfies the condition:

$$|\lambda_i - \lambda_i^*| < \varepsilon/(3kM). \tag{4.28}$$

Then, we remove the part of the upper step lying under the constructed slanting segment. The obtained polygonal line is a graph of some continuous piecewise linear function, denoted by $h_k(\lambda)$, satisfying the condition:

$$h_k(\lambda) \le \varphi_k(\lambda) \le h(\lambda) \le M, \quad \lambda \in [-\pi,\pi]. \tag{4.29}$$

Taking into account that the functions $h_k(\lambda)$ and $\varphi_k(\lambda)$ coincide outside the segments $[\lambda_i, \lambda_i^*]$ (or $[\lambda_i^*, \lambda_i]$), in view of (4.29) and (4.28), we can write

$$\|\varphi_k - h_k\|_1 = \int_{-\pi}^{\pi}[\varphi_k(\lambda)-h_k(\lambda)]d\lambda = \sum_{i=1}^{k-1}\left|\int_{\lambda_i}^{\lambda_i^*}[\varphi_k(\lambda)-h_k(\lambda)]d\lambda\right| < \frac{\epsilon}{3}. \tag{4.30}$$

Next, according to Weierstrass theorem (see, e.g., [18], Section 1.9), for function $h_k(\lambda)$ a trigonometric polynomial $\widetilde{t}(\lambda)$ can be found so that uniformly for all $\lambda \in [-\pi,\pi]$,

$$-\frac{\epsilon}{12\pi} < h_k(\lambda) - \widetilde{t}(\lambda) < \frac{\epsilon}{12\pi}. \tag{4.31}$$

Setting $t(\lambda) = \widetilde{t}(\lambda) - \frac{\epsilon}{12\pi}$, from (4.31) we get

$$0 < h_k(\lambda) - t(\lambda) < \frac{\epsilon}{6\pi}. \tag{4.32}$$

Therefore

$$\|h_k - t\|_1 = \int_{-\pi}^{\pi}[h_k(\lambda)-t(\lambda)]d\lambda < \frac{\epsilon}{3}. \tag{4.33}$$

Combining the inequalities (4.27), (4.30) and (4.33), we obtain

$$\|h-t\|_1 \le \|h-\varphi_k\|_1 + \|\varphi_k - h_k\|_1 + \|h_k - t\|_1 \le \epsilon,$$

and the inequality (4.23) follows.

Now we proceed to prove the inequality (4.24). Observe first that the second inequality in (4.24) follows from (4.32) and (4.29). To prove the first inequality in (4.24), observe that by construction of function $h_k(\lambda)$, we have

$$h_k(\lambda) \ge \min\{m_1,\dots,m_k\} \ge m. \tag{4.34}$$

Next, in view of (4.32), we get

$$t(\lambda) \geq h_k(\lambda) - \frac{\epsilon}{6\pi} > h_k(\lambda) - \epsilon. \tag{4.35}$$

Combining (4.34) and (4.35), we obtain the first inequality in (4.24).

The proof of inequalities (4.23) with (4.25) is completely similar to that of (4.23) with (4.24). The only difference is that now instead of Darboux lower sum should be used the upper sum and in the construction of function $h_k(\lambda)$, the endpoints of the upper steps of the function $\varphi_k(\lambda)$ should be connected with the interior points of the adjacent lower steps. □

Lemma 4.4. *Let $h(\lambda) \in A_+^-$ and let the sequence $\sigma_n(f)$ satisfy (4.14). Then the following asymptotic relation holds:*

$$\lim_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} = G(h). \tag{4.36}$$

Proof. Observe first that together with $h(\lambda)$ the function $1/h(\lambda)$ also belongs to the class A_+^-:

$$m \leq h(\lambda) \leq M \quad \text{and} \quad 1/M \leq 1/h(\lambda) \leq 1/m. \tag{4.37}$$

By Lemma 4.3, for a given small enough $\epsilon > 0$, we can find two trigonometric polynomials $t_1(\lambda)$ and $t_2(\lambda)$ to satisfy the following conditions:

$$\|h - t_1\|_1 < \epsilon, \qquad \frac{m}{2} < t_1(\lambda) < h(\lambda), \tag{4.38}$$

$$\|1/h - t_2\|_1 < \epsilon, \quad \frac{1}{2M} < t_2(\lambda) < \frac{1}{h(\lambda)}. \tag{4.39}$$

Now in view of (2.7) and Lemmas 4.1, 4.2, to obtain

$$\liminf_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \geq \liminf_{n\to\infty} \frac{\sigma_n^2(ft_1)}{\sigma_n^2(f)} \geq G(t_1), \tag{4.40}$$

and

$$\limsup_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leq \limsup_{n\to\infty} \frac{\sigma_n^2(f/t_2)}{\sigma_n^2(f)} \leq G(1/t_2). \tag{4.41}$$

Next, it follows from (4.37) – (4.39) that

$$\|h - 1/t_2\|_1 = \|h/t_2(t_2 - 1/h)\|_1 \leqslant 2M^2\epsilon,$$

$$\|t_1 - 1/t_2\|_1 \leqslant \|t_1 - h\|_1 + \|h - 1/t_2\|_1 \leqslant \epsilon(1 + 2M^2).$$

Hence, in view of (4.37) and (4.39), we can write

$$\left|\ln\frac{G(t_1)}{G(1/t_2)}\right| = |\ln[G(t_1)G(t_2)]| = \left|\int_{-\pi}^{\pi} \ln[t_1(\lambda)t_2(\lambda)]d\lambda\right| \leqslant \int_{-\pi}^{\pi} |t_1(\lambda)t_2(\lambda) - 1|d\lambda$$

$$= \|t_2(t_1 - 1/t_2)\|_1 \leqslant \frac{1}{m}\|t_1 - 1/t_2\|_1 \leqslant \frac{\epsilon}{m}(1 + 2M^2).$$

Thus, the quantities $G(t_1)$ and $G(1/t_2)$ can be made arbitrarily close. Hence, taking into account that $G(t_1) \leqslant G(h) \leqslant G(1/t_2)$, from (4.40) and (4.41) we obtain (4.36). $\square$

Taking into account that $G(h) > 0$, from (4.10) and (4.36) we obtain the following result.

Corollary 4.1. *If the sequence $\sigma_n(f)$ is slowly decreasing and $h(\lambda) \in A_+^-$, then the sequence $\sigma_n(fh)$ is also slowly decreasing.*

Lemma 4.5. *Let the sequence $\sigma_n(f)$ be slowly decreasing, and let $h(\lambda) \in A^-$. Then*

$$\limsup_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leqslant G(h). \tag{4.42}$$

Proof. Observe that the function $h_\epsilon(\lambda) = h(\lambda) + \epsilon$ belongs to the class A_+^-. Then we have the asymptotic relation (see, [18], Section 3.1 (d)):

$$\lim_{\epsilon\to 0} G(h_\epsilon) = G(h). \tag{4.43}$$

Hence, using (2.7) and Lemma 4.4, we obtain

$$\limsup_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \leq \lim_{n\to\infty} \frac{\sigma_n^2(fh_\epsilon)}{\sigma_n^2(f)} = G(h_\epsilon).$$

Passing to the limit as $\epsilon \to 0$, and taking into account (4.43), we obtain the desired inequality (4.42). $\square$

As an immediate consequence of Lemma 4.5, we have the following result.

Corollary 4.2. *Let the sequence $\sigma_n(f)$ be slowly decreasing, and let $g(\lambda) \in A^-$ with $G(g) = 0$. Then $\sigma_n(fg) = o(\sigma_n(f))$ as $n \to \infty$.*

Lemma 4.6. *Let the sequence $\sigma_n(f)$ be slowly decreasing, and let $h(\lambda) \in A_+$. Then*

$$\liminf_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \geqslant G(h). \tag{4.44}$$

Proof. Let $h_l(\lambda)$ denote the truncation of $h(\lambda)$ at the level $l \in \mathbb{N}$:

$$h_l(\lambda) = \begin{cases} h(\lambda), & h(\lambda) \leqslant l \\ l, & h(\lambda) > l. \end{cases}$$

Then by Monotone Convergence Theorem of Beppo Levi (see [6], Theorem 2.8.2), we have

$$\lim_{l\to\infty} G(h_l) = G(h). \tag{4.45}$$

Next, by Lemma 4.4 we get

$$\liminf_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} \geqslant \lim_{n\to\infty} \frac{\sigma_n^2(fh_l)}{\sigma_n^2(f)} = G(h_l).$$

Hence passing to the limit as $l \to \infty$, and taking into account (4.45) we obtain the desired inequality (4.44). $\square$

As an immediate consequence of Lemma 4.6, we have the following result.

Corollary 4.3. *Let the sequence $\sigma_n(f)$ be slowly decreasing, $g(\lambda) \in A_+$ with $G(g) = \infty$, and let $fg \in A$. Then $\sigma_n(f) = o(\sigma_n(fg))$ as $n \to \infty$.*

4.4. **Proof of main results.** In this subsection we prove the main results of this paper - Theorems 4.4 and 4.5.

Proof of Theorem 4.4. We have

$$\frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(f)} = \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(fht_1)} \cdot \frac{\sigma_n^2(fht_1)}{\sigma_n^2(fh)} \cdot \frac{\sigma_n^2(fh)}{\sigma_n^2(f)}. \tag{4.46}$$

Next, by Lemma 4.4 we have

$$\lim_{n\to\infty} \frac{\sigma_n^2(fh)}{\sigma_n^2(f)} = G(h) > 0. \tag{4.47}$$

This, in view of Corollary 4.1, implies that the sequence $\sigma_n^2(fh)$ is also slowly decreasing. Therefore, by Lemma 4.1, we have

$$\liminf_{n\to\infty} \frac{\sigma_n^2(fht_1)}{\sigma_n^2(fh)} \geq G(t_1).$$

On the other hand, since $t_1(\lambda) \in A^-$, then according to Lemma 4.5, we get

$$\limsup_{n\to\infty} \frac{\sigma_n^2(fht_1)}{\sigma_n^2(fh)} \leq G(t_1).$$

Therefore

$$\lim_{n\to\infty} \frac{\sigma_n^2(fht_1)}{\sigma_n^2(fh)} = G(t_1) > 0 \tag{4.48}$$

This implies that the sequence $\sigma_n^2(fht_1)$ is also slowly decreasing. Hence we can apply Lemma 4.2, to obtain

$$\limsup_{n\to\infty} \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(fht_1)} \leqslant G(1/t_2).$$

Next, it is easy to see that $1/t_2 \in A_+$. Hence, according to Lemma 4.6, we obtain

$$\liminf_{n\to\infty} \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(fht_1)} \geqslant G(1/t_2).$$

Therefore

$$\lim_{n\to\infty} \frac{\sigma_n^2(fht_1/t_2)}{\sigma_n^2(fht_1)} = G(1/t_2). \tag{4.49}$$

Finally, combining the relations (4.46) - (4.49), we obtain

$$\lim_{n\to\infty} \frac{\sigma_n^2(fg)}{\sigma_n^2(f)} = G(1/t_2)G(t_1)G(h) = G(ht_1/t_2) = G(g).$$

Theorem 4.4 is proved. $\square$

As an immediate consequence of Theorem 4.4, we have the following result.

Corollary 4.4. *If the sequence $\sigma_n(f)$ is slowly decreasing and $g(\lambda)$ satisfies the conditions of Theorem 4.4, then the sequence $\sigma_n(fg)$ is also slowly decreasing.*

The proof of Theorem 4.5 immediately follows from Theorems 4.3 and 4.4.

Список литературы

[1] R. B. Ash, M. F. Gardner, Topics in Stochastic Processes. New York, Academic Press (1975).

[2] G. Baxter, "An asymptotic result for the finite predictor", Math. Scand., **10**, 137 – 144 (1962).

[3] N. Babayan, "On the asymptotic behavior of prediction error", J. of Soviet Mathematics, **27**(6), 3170 – 3181 (1984).

[4] N. Babayan, "On asymptotic behavior of the prediction error in the singular case", Theory Probab. Appl., **29** (1), 147 – 150 (1985).

[5] N. H. Bingham,: "Szegö's theorem and its probabilistic descendants", Probability Surveys, **9**, 287 – 324 (2012), DOI: 10.1214/11-PS178

[6] V. I. Bogachev, Measure Theory, **1**, Springer-Verlag (2007).

[7] P. J. Brockwell,, R. A. Davis, Time Series: Theory and Methods, Second Edition, New York, Springer-Verlag (1991).

[8] L. D. Davisson, "Prediction of time series from finite past", J. Soc. Indust. Appl. Math., **13**, no. 3, 819 – 826 (1965).

[9] A. Devinatz, "Asymptotic estimates for the finite predictor", Math. Scand., **15**, 111 – 120 (1964).

[10] J. L. Doob, Stochastic Processes, John Wiley & Sons, New York (1953).

[11] M. I. Fortus, "Prediction of a stationary time series with the spectrum vanishing on an interval", Akademiia Nauk SSSR, Izvestiia, Fizika Atmosfery i Okeana **26**, 1267 – 1274 (1990).

[12] Ya. L. Geronimus, "On a problem of G. Szegö, M. Kac, G. Baxter and I. Hirshman", Izv. AN SSSR, ser. Matematika, **31**, 289 – 304 (1967).

[13] M. S. Ginovian, "Asymptotic behavior of the prediction error for stationary Random sequences", Journal of Contemporary Math. Anal., **34**, no. 1, 14 – 33 (1999).

[14] B. L. Golinskii, "On asymptotic behavior of the prediction error", Theory Probab. and appl., **19**, no. 4, 724 – 739 (1974).

[15] B. L. Golinskii and I. A. Ibragimov, "On G. Szegö limit theorem", Izv. AN SSSR, ser. Matematika, **35**, 408 – 427 (1971).

[16] G. M. Goluzin, Geometric Theory of Functions of a Complex Variable, Providence, Amer. Math. Soc. (1969).

[17] U. Grenander, M. Rosenblatt, "An extension of a theorem of G. Szegö and its application to the study of stochastic processes", Trans. Amer. Math. Soc., **76**, 112 – 126 (1954).

[18] U. Grenander, G. Szegö, Toeplitz Forms and Their Applications, University of California Press, Berkeley and Los Angeles (1958).

[19] H. Helson, G. Szegö, "A problem in prediction theory", Acta Mat. Pura Appl. **51**, 107 – 138 (1960).

[20] I. I. Hirschman, "Finite sections of Wiener-Hopf equations and Szegö polynomials", Journal of Mathematical Analysis and Applications **11**, 290 – 320 (1965).

[21] I. I. Hirschman, "On a theorem of Szegö, Kac, and Baxter", Journal d'Analyse Mathematique **14**, no. 1, 225 – 234 (1965).

[22] K. Hoffman, Banach Spaces of Analytic Functions, Prentice-Hall, Englewood Cliffs, N. J. (1958).

[23] I. A. Ibragimov, "On asymptotic behavior of the prediction error", Probab. Theory and appl., **9**, no. 4, 695 – 703 (1964).

[24] I. A. Ibragimov, Yu. A. Rozanov, Gaussian Random Processes, New York, Springer-Verlag (1978).

[25] I. A. Ibragimov, V. N. Solev, "The asymptotic behavior of the prediction error of a stationary sequence with the spectral density function of a special form", Probab. Theory and Appl. **13**, no. 4, 746 – 750 (1968).
[26] A. Inoue, "Asymptotic behavior for partial autocorrelation functions of fractional ARIMA processes", The Annals of Applied Probability, **12**(4), 1471 – 1491 (2002).
[27] A. N. Kolmogorov, "Stationary sequences in a Hilbert space", Bull. Moscow State University, **2**, no.6, 1 – 40 (1941).
[28] A. N. Kolmogorov, "Inerpolation and extrapolation of stationary random sequences", Izv. Akad. Nauk SSSR. Ser. Tat., **5**, 3 – 14 (1941).
[29] M. Pourahmadi, Fundamentals of Time Series Analysis and Prediction Theory, New York, Wiley (2001).
[30] E. A. Rakhmanov, "On asymptotic properties of polynomials orthogonal on the circle with weights not satisfying Szegö's condition", Mat. Sb. **130**(172), 151 – 169 (1986); Engl. transl. Math. USSR Sb. **58**, 149 – 167 (1987).
[31] M. Rosenblatt, "Some purely deterministic processes", J. of Math. and Mech., **6**, no. 6, 801 – 810 (1957).
[32] Yu. A. Rozanov, Stationary Random Processes, Holden-Day, San Francisco (1967).
[33] A. N. Shiryaev, Probability, New York, Springer-Verlag (1984).
[34] N. Wiener, Extrapolation, Interpolation and Smoothing of Stationary Time Series, With engineering applications, MIT Press/Wiley (1949).

Поступила 21 марта 2019

После доработки 26 августа 2019

Принята к публикации 19 декабря 2019