Adaptive energy-saving approximation for stationary processes

Z. A. Kabluchko and M. A. Lifshits

Abstract. We consider a stationary process (with either discrete or continuous time) and find an adaptive approximating stationary process combining high quality approximation and other good properties that can be interpreted as additional smoothness or small expense of energy. The problem is solved in terms of spectral characteristics of the original process using the classical analytic methods of prediction theory.

Keywords: least-energy approximation, prediction, stationary process, stationary sequence.

§ 1. Main objects and statement of the problem

Consider a random process $B(t)$ with continuous ($t \in \mathbb{R}$) or discrete ($t \in \mathbb{Z}$) time. We try to approximate $B$ by another process $X$ that, being close to $B$, would have better sample path properties (in an appropriate sense). For example, one might imagine that a sample path of $B$ models the trajectory of a chaotically moving target while the sample path of $X$ is a pursuit trajectory built upon observations of $B$. In the most interesting cases (when the time is continuous), the trajectories of $B$ are non-differentiable while the trajectories of $X$ are required to be smooth.

In this paper we assume that $B$ and $X$ are stationary processes in the wide sense. Additional requirements on $X$ will be stated in terms of small average expense of generalized energy, whose definition will be given below.

Continuous time, stationary process. Let $(B(t))_{t \in \mathbb{R}}$ be a centred complex-valued process stationary in the wide sense. The last condition means that $E|B(t)|^2 < \infty$ and the covariance function of $B$ depends only on the difference of times, namely,

\[ \text{cov}(B(t_1), B(t_2)) = \text{cov}(B(t_1 - t_2), B(0)) =: K_B(t_1 - t_2). \]

As usual, we assume that $K_B(\cdot)$ is continuous.

We seek a process $(X(t))_{t \in \mathbb{R}}$ such that the pair $(B, X)$ is jointly stationary in the wide sense, the processes are close to each other, and $X$ expends a small amount of energy (a notion to be defined soon).
The *instant energy* of $X$ at time $t \in \mathbb{R}$ is the expression

$$
E[X](t) := \left| \sum_{m=0}^{M} \ell_m X^{(m)}(t) \right|^2,
$$

where $X^{(m)}$ stands for the $m$th derivative of $X$, and the $\ell_m$ are certain fixed complex coefficients. The most natural type of energy is the *kinetic energy*, which is just $\alpha^2 |X^{(1)}(t)|^2$ for some $\alpha > 0$.

Our natural goal is the minimization of the functional

$$
\lim_{T \to \infty} \frac{1}{T} \int_0^T \left[ |X(t) - B(t)|^2 + E[X](t) \right] dt,
$$

which combines the approximation and energy properties with averaging in time. If we also assume that the process $X(t) - B(t)$ and all the derivatives $X^{(m)}(t)$ are stationary processes in the strict sense, then the ergodic theorem applies in many cases and the limit above is equal to $E[X(0) - B(0)]^2 + \mathbb{E}E[X](0)$. In the wide-sense theory, we simplify our task to solving the problem

$$
E[X(0) - B(0)]^2 + \mathbb{E}E[X](0) \searrow \min
$$

and setting aside issues of ergodicity. From the point of view of control theory, the term $\mathbb{E}E[X](0)$ may be regarded as a kind of penalty which imposes certain smoothness conditions on $X$.

Notice that once a problem of the form (1.1) has been solved, one can easily separate the two terms in (1.1) by solving (by the Lagrange multiplier method) the following somewhat more natural problems.

(a) Find a process $X$ with minimal expense of energy and with a prescribed closeness to $B$:

$$
\mathbb{E}E[X](0) \searrow \min \quad \text{(over } X \text{ such that } \mathbb{E}|X(0) - B(0)|^2 \leq \delta)\)

for any given $\delta > 0$.

b) Find a process $X$ which is the closest to $B$ using a given amount of energy:

$$
\mathbb{E}|X(0) - B(0)|^2 \searrow \min \quad \text{(over } X \text{ such that } \mathbb{E}E[X](0) \leq \mathcal{E})\)

for any given $\mathcal{E} > 0$.

The following standard notation will be used throughout the paper. Given an arbitrary Hilbert space $H$ and any set $A \subset H$, we write $\text{span}\{A \mid H\}$ for the closed linear span of $A$ in $H$.

We consider the problem (1.1) either in the simpler *non-adaptive setting* (that is, without any further restrictions on $X$), or in the *adaptive setting* by requiring that

$$
X(t) \in \text{span}\{B(\tau), \tau \leq t \mid L_2(\Omega, \mathbb{P})\}, \quad t \in \mathbb{R},
$$

1In what follows we also consider more general forms of energy corresponding to rather arbitrary functions instead of polynomials.
In other words, we allow only approximations based on the current and past values of $B$. The non-adaptive setting was considered in [1], [2]. In the present paper we briefly recall the corresponding results and concentrate on the much more interesting and difficult adaptive setting.

Our approach to solving (1.1) is based on the spectral representations of stationary processes. We recall some basic facts of this theory (see, for example, [3],[4]).

Being non-negative definite, the covariance function $K_B(\cdot)$ admits a spectral representation of the following form by the Bochner–Khintchine theorem:

$$K_B(t) = \int_{\mathbb{R}} e^{itu} \mu(du),$$

where $\mu$ is a finite measure called the spectral measure of $B$. Moreover, the process $B$ itself admits a spectral representation

$$B(t) = \int_{\mathbb{R}} e^{itu} \mathcal{W}(du),$$

where $\mathcal{W}(du)$ is a centred complex-valued random measure with uncorrelated values on $\mathbb{R}$ controlled by the spectral measure $\mu$, that is, $\mu(A) = \mathbb{E} |\mathcal{W}(A)|^2$ for every Borel set $A \subset \mathbb{R}$.

There is no loss of generality in restricting our optimization problem to the class of approximating processes of the form

$$X(t) = \int_{\mathbb{R}} \widehat{g}(u) e^{itu} \mathcal{W}(du), \quad (1.2)$$

where $\widehat{g}(\cdot) \in L_2(\mathbb{R}, \mu)$ is an unknown function. For example, if $X$ is a moving average process, that is,

$$X(t) = \int_{\mathbb{R}} g(\tau) B(t + \tau) d\tau \quad (1.3)$$

for some weight $g \in L_1(\mathbb{R})$, then we have

$$X(t) = \int_{\mathbb{R}} g(\tau) \int_{\mathbb{R}} e^{i(t+\tau)u} \mathcal{W}(du) d\tau = \int_{\mathbb{R}} \widehat{g}(u) e^{itu} \mathcal{W}(du),$$

where

$$\widehat{g}(u) := \int_{\mathbb{R}} g(\tau) e^{i\tau u} d\tau$$

is the inverse Fourier transform of $g$.

Indeed, it is easy to show that every process $X$ which is jointly stationary in the wide sense with $B$ can be written as the sum of two wide-sense stationary processes,

$$X(t) = X_1(t) + X_2(t), \quad t \in \mathbb{R},$$

where $X_1$ is a process of the class (1.2) and $X_2(t)$ is uncorrelated with $B(s)$ for all $s, t \in \mathbb{R}$. It follows that

$$\mathbb{E} |X_1(0) - B(0)|^2 + \mathbb{E} \mathcal{E}[X_1](0) \leq \mathbb{E} |X(0) - B(0)|^2 + \mathbb{E} \mathcal{E}[X](0),$$

and the reduction to the class (1.2) is justified.
Furthermore, if \( X \) is a process of the form (1.2) and
\[
\int_{\mathbb{R}} |\hat{g}(u)|^2 u^{2m} \mu(du) < \infty
\]
for some positive integer \( m \), then the \( m \)th derivative of \( X \) in the mean-square sense exists and admits a representation
\[
X^{(m)}(t) = \int_{\mathbb{R}} \hat{g}(u)(iu)^m e^{itu} \mathcal{W}(du).
\]
Hence,
\[
\mathbb{E}|X(0) - B(0)|^2 + \mathbb{E} \mathcal{E}[X](0)
\]
\[
= \mathbb{E}\left| \int_{\mathbb{R}} (\hat{g}(u) - 1) \mathcal{W}(du) \right|^2 + \mathbb{E}\left| \int_{\mathbb{R}} \hat{g}(u) \sum_{m=0}^{M} \ell_m(iu)^m \mathcal{W}(du) \right|^2
\]
\[
= \int_{\mathbb{R}} \left[ |\hat{g}(u) - 1|^2 + |\hat{g}(u)|^2 |\ell(iu)|^2 \right] \mu(du)
\]
with energy polynomial
\[
\ell(z) := \sum_{m=0}^{M} \ell_m z^m.
\]
The problem (1.1) can now be restated analytically as
\[
\int_{\mathbb{R}} \left[ |\hat{g}(u) - 1|^2 + |\hat{g}(u)|^2 |\ell(iu)|^2 \right] \mu(du) \searrow \min. \quad (1.4)
\]
Notice that one can also consider this problem for a more or less arbitrary function \( \ell(\cdot) \) instead of a polynomial.

The spectral condition equivalent to the adaptive setting is
\[
\hat{g}(u) e^{itu} \in \text{span}\{e^{i\tau u}, \tau \leq t \mid \text{L}_2(\mathbb{R}, \mu)\}, \quad t \in \mathbb{R}.
\]
This condition clearly holds for all \( t \in \mathbb{R} \) if and only if it holds for \( t = 0 \), that is,
\[
\hat{g} \in \hat{G}_{\leq 0} := \text{span}\{e^{i\tau u}, \tau \leq 0 \mid \text{L}_2(\mathbb{R}, \mu)\}.
\]

**Discrete time, stationary sequence.** Let \((B(t))_{t \in \mathbb{Z}}\) be a centred stationary in the wide sense sequence. This means that \( \mathbb{E}|B(t)|^2 < \infty \) and the covariance function of \( B \) depends only on the time-difference:
\[
\text{cov}(B(t_1), B(t_2)) = \text{cov}(B(t_1 - t_2), B(0)) =: K_B(t_1 - t_2).
\]
Being non-negative definite, the covariance function \( K_B(\cdot) \) admits a spectral representation of the following form by Herglotz’ theorem:
\[
K_B(t) = \int_{\mathbb{T}} e^{itu} \mu(du), \quad t \in \mathbb{Z},
\]
where $\mu$ is a finite measure on $\mathbb{T} := [−\pi, \pi)$ called the spectral measure of $B$. The sequence $B$ itself admits a spectral representation

$$B(t) = \int_{\mathbb{T}} e^{itu} \mathcal{W}(du), \quad t \in \mathbb{Z},$$

where $\mathcal{W}(du)$ is a centred complex-valued random measure with uncorrelated values on $\mathbb{T}$ controlled by $\mu$.

As in (1.2), we seek an approximating sequence $(X(t))_{t \in \mathbb{Z}}$ in the form

$$X(t) = \int_{\mathbb{T}} \widehat{g}(u) e^{itu} \mathcal{W}(du), \quad t \in \mathbb{Z},$$

where $\widehat{g}(\cdot) \in L_2(\mathbb{T}, \mu)$. For example, if $X$ is a moving average sequence, that is,

$$X(t) = \sum_{\tau \in \mathbb{Z}} g(\tau) B(t + \tau)$$

for a certain summable weight $g$, then

$$X(t) = \int_{\mathbb{T}} \widehat{g}(u) e^{itu} \mathcal{W}(du), \quad t \in \mathbb{Z},$$

where

$$\widehat{g}(u) := \sum_{\tau \in \mathbb{Z}} g(\tau) e^{i\tau u}$$

is the inverse Fourier transform of $g$.

In the discrete case, the notion of energy must be modified by replacing the (right) derivatives by their discrete analogues, that is, we write $X(t + 1) − X(t)$ for $X^{(1)}(t)$, $X(t + 2) − 2X(t + 1) + X(t)$ for $X^{(2)}(t)$ and so on. Hence the instant energy of $X$ takes the form

$$\mathcal{E}[X](t) := \left| \sum_{m=0}^{M} \ell_m X(t + m) \right|^2.$$

Using the integral representation

$$\sum_{m=0}^{M} \ell_m X(t + m) = \int_{\mathbb{T}} \widehat{g}(u) \left( \sum_{m=0}^{M} \ell_m e^{imu} \right) e^{itu} \mathcal{W}(du) := \int_{\mathbb{T}} \widehat{g}(u) \ell(e^{iu}) e^{itu} \mathcal{W}(du)$$

with the polynomial

$$\ell(z) := \sum_{m=0}^{M} \ell_m z^m,$$

we have

$$\mathbb{E} \mathcal{E}[X](t) = \int_{\mathbb{T}} |\widehat{g}(u)|^2 |\ell(e^{iu})|^2 \mu(du), \quad t \in \mathbb{Z}.$$

The discrete-time version of the problem (1.4) becomes

$$\int_{\mathbb{T}} \left[ |\widehat{g}(u)|^2 + |\widehat{g}(u)|^2 |\ell(e^{iu})|^2 \right] \mu(du) \searrow \min.$$  \hfill (1.5)
Again, one can consider this problem for an arbitrary function $\ell(\cdot)$ instead of a polynomial. The discrete-time analogue of kinetic energy corresponds to the increment $\alpha(X(t+1) - X(t))$, that is, to the polynomial $\ell(z) = \alpha(z - 1)$.

One can consider the problem (1.5) either in the non-adaptive setting, or in the adaptive setting by requiring that $b_g \in bG_6 := \text{span}\{e^{iru}, r \leq 0, \tau \in \mathbb{Z} | L_2(\mathbb{T}, \mu)\}$.

§ 2. First step: solution of the non-adaptive problem

2.1. Continuous time. We have already seen that the continuous-time problem (1.4) is of the form

$$
\int_{\mathbb{R}} [ |\hat{g}(u) - 1|^2 + |\hat{g}(u)|^2 |\ell(iu)|^2] \mu(du) \searrow \min.
$$

(2.1)

For any complex numbers $\hat{g}$ and $\ell$ we have the identity

$$
|\hat{g} - 1|^2 + |\hat{g}|^2 |\ell|^2 = (|\ell|^2 + 1) |\hat{g} - \frac{1}{|\ell|^2 + 1}|^2 + \frac{|\ell|^2}{|\ell|^2 + 1}.
$$

(2.2)

Hence the solution of (2.1) in the non-adaptive setting (when no further restrictions are imposed on $\hat{g}$) is given by the function

$$
\hat{g}_*(u) := \frac{1}{|\ell(iu)|^2 + 1}, \quad u \in \mathbb{R},
$$

(2.3)

which depends on the energy form $\ell$ but not on the spectral measure $\mu$. The minimum in (2.1) is equal to

$$
\text{ERR}_{\text{NA}} := \int_{\mathbb{R}} \frac{|\ell(iu)|^2}{|\ell(iu)|^2 + 1} \mu(du).
$$

This quantity may naturally be called the non-adaptive approximation error. In control theory, the term problem cost is also used.

In the simplest case of the kinetic energy $\ell(z) = \alpha z$, where $\alpha > 0$ is a scaling parameter, we get

$$
\text{ERR}_{\text{NA}} := \int_{\mathbb{R}} \frac{\alpha^2 u^2}{\alpha^2 u^2 + 1} \mu(du).
$$

(2.4)

Since the function $\hat{g}_*(u) = 1/(\alpha^2 u^2 + 1)$ is the inverse Fourier transform of

$$
g_*(t) := \frac{1}{2\alpha} \exp\left\{-\frac{|t|}{\alpha}\right\}, \quad t \in \mathbb{R},
$$

we conclude that the solution of the non-adaptive problem with kinetic energy for stationary processes is given, as suggested by (1.3), by the moving average process

$$
X(t) = \frac{1}{2\alpha} \int_{\mathbb{R}} \exp\left\{-\frac{|\tau|}{\alpha}\right\} B(t + \tau) d\tau.
$$

(2.5)
Notice that this solution is indeed non-adaptive because the future values of $B$ occur in the approximation. The formula (2.3) was obtained in [1]; see also [2] for the case of kinetic energy.

However, if the adaptivity restriction is imposed on $b^g$, then (2.3) does not apply and we have to minimize the spectrum-dependent integral. By (2.2), the problem (2.1) reduces to

$$\int_{\mathbb{R}} \left| \hat{g}(u) - \frac{1}{|\ell(iu)|^2 + 1} \right|^2 \left(|\ell(iu)|^2 + 1\right) \mu(du) \min.$$ \hspace{1cm} (2.6)

This minimum (taken over $\hat{g} \in \hat{G}_{\leq 0}$) will be called the additional adaptivity error and denoted by $\text{ERR}^+_A$. This is the price we must pay for not knowing the future. The total approximation error (that is, the minimum in (2.1) over $\hat{g} \in \hat{G}_{\leq 0}$) is then equal to

$$\text{ERR}_A := \text{ERR}_{NA} + \text{ERR}^+_A.$$  

These formulae were obtained in [1].

2.2. Discrete time. The situation in the discrete-time setting is completely analogous because the problem (1.5) differs from (2.1) only by replacing the spectral domain $\mathbb{R}$ by $\mathbb{T}$ and $\ell(iu)$ by $\ell(e^{iu})$. Therefore we obtain the following expression for the non-adaptive error:

$$\text{ERR}_{NA} := \int_{\mathbb{T}} \frac{|\ell(e^{iu})|^2}{|\ell(e^{iu})|^2 + 1} \mu(du),$$

which is attained at

$$\hat{g}_*(u) := \frac{1}{|\ell(e^{iu})|^2 + 1}, \quad u \in \mathbb{T}.$$ 

In the simplest case of the discrete kinetic energy $\ell(z) = \alpha(z - 1)$ we have

$$\hat{g}_*(u) = \frac{1}{\sqrt{1 + 4\alpha^2}} \left(1 + \sum_{k=1}^{\infty} \beta^{-k}(e^{iku} + e^{-iku})\right),$$ \hspace{1cm} (2.7)

where

$$\beta = \frac{2\alpha^2 + 1 + \sqrt{1 + 4\alpha^2}}{2\alpha^2} > 1.$$ \hspace{1cm} (2.8)

The analogue of (2.5) is

$$X(t) = \frac{1}{\sqrt{1 + 4\alpha^2}} \left(B(t) + \sum_{k=1}^{\infty} \beta^{-k}(B(t + k) + B(t - k))\right)$$

while

$$\text{ERR}_{NA} = \int_{\mathbb{T}} \frac{\alpha^2|e^{iu} - 1|^2}{|\alpha^2|e^{iu} - 1|^2 + 1} \mu(du).$$ \hspace{1cm} (2.9)
§ 3. Solutions of the adaptive approximation problem

We recall that the adaptive approximation problem for continuous-time processes was reduced in (2.6) to the problem

\[
\int_{\mathbb{R}} \left| \hat{g}(u) - \frac{1}{|\ell(iu)|^2 + 1} \right|^2 \left( |\ell(iu)|^2 + 1 \right) \mu(du) \Rightarrow \min
\]

(3.1)

over \( \hat{g} \in \hat{G}_{\leq 0} \). This looks very much like a classical prediction problem except for the function to be approximated, which is \( 1/(|\ell(iu)|^2 + 1) \) in our case and \( e^{i\tau u}, \tau > 0 \), in the prediction problem. Therefore we may either directly reduce the approximation problem to the prediction problem, or solve the approximation problem using the methods developed for the prediction problem. The latter seems to be more efficient and general method.

3.1. Direct reduction to prediction problems.

3.1.1. Continuous time. Consider the continuous-time setting. Assume that there is an appropriate (in the sense to be made precise a bit later) factorization

\[
|\ell(iu)|^2 + 1 = \lambda_\ell(u) = |\lambda_\ell(u)|^2, \quad u \in \mathbb{R}.
\]

(3.2)

Then the left-hand side of (3.1) takes the form

\[
\int_{\mathbb{R}} \left| \lambda_\ell(u) \hat{g}(u) - \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du).
\]

(3.3)

Recall that the classical prediction problem is

\[
\int_{\mathbb{R}} \left| \hat{g}(u) - e^{i\tau u} \right|^2 \mu(du) \Rightarrow \min, \quad \tau \in \mathbb{R},
\]

(3.4)

over \( \hat{g} \in \hat{G}_{\leq 0} \). Let \( \hat{q}^{(\tau, \mu)}_* \) denote the solution of this problem. The solution of the general prediction problem

\[
\int_{\mathbb{R}} \left| \hat{g}(u) - \nu(u) \right|^2 \mu(du) \Rightarrow \min
\]

is linear in \( \nu \). Therefore, if there is a representation

\[
\frac{1}{\lambda_\ell(u)} = \int_{0}^{\infty} e^{-i\tau u} \nu_\ell(d\tau), \quad u \in \mathbb{R},
\]

(3.5)

with some finite complex-valued measure \( \nu_\ell \) depending on the energy polynomial \( \ell(\cdot) \), then the function

\[
\hat{q}^{(\ell, \mu)}_*(u) := \int_{0}^{\infty} \hat{q}^{(\tau, \mu)}_*(u) \nu_\ell(d\tau), \quad u \in \mathbb{R},
\]

belongs to \( \hat{G}_{\leq 0} \) and satisfies the inequality

\[
\int_{\mathbb{R}} \left| \hat{q}^{(\ell, \mu)}_*(u) - \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du) \leq \int_{\mathbb{R}} \left| \hat{g}(u) - \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du), \quad \hat{g} \in \hat{G}_{\leq 0}.
\]

It follows from the representation (3.5) that \( (1/\lambda_\ell) \hat{g} \in \hat{G}_{\leq 0} \) for any \( \hat{g} \in \hat{G}_{\leq 0} \).
We now impose another assumption on the factorization (3.2):

\[ \text{if } \hat{g} \in \hat{G}_{\leq 0} \text{ and } \lambda_\ell \hat{g} \in L_2(\mathbb{R}, \mu), \text{ then one must have } \lambda_\ell \hat{g} \in \hat{G}_{\leq 0}. \quad (3.6) \]

When the conditions (3.5) and (3.6) hold, the function

\[ \hat{g}_*(u) := \frac{1}{\lambda_\ell(u)} \hat{q}_*^{(\ell,\mu)}(u) \]

minimizes (3.3) over \( \hat{g} \in \hat{G}_{\leq 0} \) and thus solves the problem (3.1). Indeed, suppose that \( \hat{g} \in \hat{G}_{\leq 0} \). If \( \lambda_\ell \hat{g} \not\in L_2(\mathbb{R}, \mu) \), then the expression (3.3) is infinite because the bounded function \( 1/\lambda_\ell \) belongs to \( L_2(\mathbb{R}, \mu) \). We now suppose that \( \lambda_\ell \hat{g} \in L_2(\mathbb{R}, \mu) \). Then, using the optimality of \( \hat{q}_*^{(\ell,\mu)} \) and the inclusion \( \lambda_\ell \hat{g} \in \hat{G}_{\leq 0} \), we have

\[
\int_\mathbb{R} \left| \lambda_\ell(u) \hat{g}_*(u) - \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du) = \int_\mathbb{R} \left| \hat{q}_*^{(\ell,\mu)}(u) - \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du) \\
\leq \int_\mathbb{R} \left| \lambda_\ell(u) \hat{g}(u) - \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du)
\]

and the problem is solved.

We stress that a representation with properties (3.5) and (3.6) always exists for polynomials. Indeed, let \( \ell(\cdot) \) be a polynomial of degree \( M \) with complex coefficients. Then for all real \( u \) we have

\[ 1 + |\ell(iu)|^2 = 1 + \ell(iu)\overline{\ell(iu)} =: \mathcal{P}(u), \]

where \( \mathcal{P} \) is a polynomial of degree \( 2M \) with real coefficients. Therefore, if \( \beta \) is a root of \( \mathcal{P} \), then so is \( \overline{\beta} \). We also notice that \( \mathcal{P} \) has no real roots. Thus we may write

\[ \mathcal{P}(u) = C \prod_{m=1}^{M} (u - \beta_m)(u - \overline{\beta_m}), \]

where \( \text{Im} \beta_m > 0 \) and \( C > 0 \) (this follows by putting \( u = 0 \)). We finally obtain that

\[ 1 + |\ell(iu)|^2 = \mathcal{P}(u) = \lambda_\ell(u)\overline{\lambda_\ell(u)}, \quad u \in \mathbb{R}, \]

where

\[ \lambda_\ell(u) := C^{1/2} \prod_{m=1}^{M} (u - \beta_m). \quad (3.7) \]

In this case we have a representation

\[ \frac{1}{u - \beta_m} = i \int_0^\infty \exp\{-i(u - \beta_m)\tau\} \, d\tau, \]

and the existence of a representation (3.5) for \( 1/\lambda_\ell \) follows.

To verify (3.6), we notice that since the polynomial \( \lambda_\ell \) is bounded away from zero, the condition \( \lambda_\ell \hat{g} \in L_2(\mathbb{R}, \mu) \) is equivalent to the condition

\[ \int_\mathbb{R} |u|^{2m} |\hat{g}(u)|^2 \mu(du) < \infty, \quad 1 \leq m \leq M. \]
Using this fact with $m = 1$, we see that the functions
\[
\tilde{g}_\delta(u) := \frac{1 - e^{-i\delta u}}{i\delta} \tilde{g}(u) \in \tilde{G}_{\leq 0}
\]
converge to the function $u\tilde{g}(u)$ in $L_2(\mathbb{R}, \mu)$ as $\delta \to 0$. Therefore the limit also belongs to $\tilde{G}_{\leq 0}$. Proceeding by induction, we conclude that all the functions $u^m\tilde{g}(u)$ for $1 \leq m \leq M$ belong to $\tilde{G}_{\leq 0}$. The same clearly holds for their linear combination $\lambda_\ell \tilde{g} \in \tilde{G}_{\leq 0}$.

For example, consider the continuous-time kinetic energy polynomial $\ell(z) = \alpha z$. Then we can use the factorization
\[
|\ell(iu)|^2 + 1 = \alpha^2 u^2 + 1 =: \lambda_\ell(u)\overline{\lambda_\ell(u)}
\]
where $\lambda_\ell(u) := 1 + i\alpha u$ and
\[
\frac{1}{\lambda_\ell(u)} = \frac{1}{\alpha} \int_0^\infty e^{-\tau/\alpha} e^{-i\tau u} d\tau., \quad (3.8)
\]

3.1.2. Discrete time. In the discrete-time setting, one need only replace $\mathbb{R}$ by $\mathbb{T}$ and $\ell(\text{i}u)$ by $\ell(e^{i\text{u}})$ in (3.1). We now need a factorization
\[
|\ell(e^{i\text{u}})|^2 + 1 = \lambda_\ell(u)\overline{\lambda_\ell(u)} = |\lambda_\ell(u)|^2, \quad u \in \mathbb{T},
\]
where $1/\lambda_\ell(\cdot)$ admits the following representation instead of (3.5):
\[
\frac{1}{\lambda_\ell(u)} = \sum_{\tau=0}^\infty \nu_\ell(\tau) e^{-i\tau u}, \quad u \in \mathbb{T}, \quad (3.9)
\]
and satisfies the following analogue of (3.6):
\[
\text{if } \tilde{g} \in \tilde{G}_{\leq 0} \text{ and } \lambda_\ell \tilde{g} \in L_2(\mathbb{T}, \mu), \text{ then one must have } \lambda_\ell \tilde{g} \in \tilde{G}_{\leq 0}. \quad (3.10)
\]

Then a solution of our problem is given by the function
\[
\tilde{g}_*(u) := \frac{1}{\lambda_\ell(u)} \sum_{\tau=0}^\infty \nu_\ell(\tau) \tilde{g}_*(\tau, \mu)(u), \quad u \in \mathbb{T},
\]
where $\tilde{g}_*(\tau, \mu)$ is the solution of the classical prediction problem (see (3.4)).

We now explain how to construct the required factorizations for the typical form of energy represented by an arbitrary complex polynomial $\ell(\cdot)$. Suppose that
\[
\ell(z) := \sum_{m=0}^M \ell_m z^m,
\]
where $\ell_M \neq 0$. Then, for $z$ on the unit circle,
\[
1 + |\ell(z)|^2 = 1 + \ell(z)\overline{\ell(z)} = 1 + \left( \sum_{m=0}^M \ell_m z^m \right) \left( \sum_{m=0}^M \overline{\ell_m} z^{-m} \right) =: \frac{\mathcal{P}(z)}{z^M},
\]
where
\[ P(z) := \sum_{m=0}^{2M} p_m z^m \]
is a polynomial of degree at most \(2M\) whose coefficients satisfy the Hermitian symmetry condition \(p_{2M-m} = \overline{p_m}\). By this symmetry, if \(\beta \neq 0\) is a root of \(P\), then so is \(1/\beta\). We also notice that \(P\) has no roots on the unit circle. Assume for the moment that \(\ell_0 \neq 0\). Then \(p_0 = \ell_0 \overline{\ell_M} \neq 0\), whence the origin is not a zero of \(P\) and we may write
\[
1 + |\ell(z)|^2 = \frac{P(z)}{z^M} = C \prod_{m=1}^{M} (z - \beta_m) \left( z - \frac{1}{\beta_m} \right) \frac{1}{z} = \frac{(-1)^M C}{\prod_{m=1}^{M} \beta_m} \prod_{m=1}^{M} (z - \beta_m)(\overline{z} - \overline{\beta_m})
\]
for some complex \(C\) and \(|\beta_m| > 1\). Putting \(z = 1\), we see that the outer constant is positive:
\[
R := \frac{(-1)^M C}{\prod_{m=1}^{M} \beta_m} > 0.
\]
Hence we have a factorization
\[
1 + |\ell(\text{e}^{iu})|^2 = \lambda_\ell(u) \overline{\lambda_\ell(u)}, \quad u \in \mathbb{T},
\]
where
\[
\lambda_\ell(u) := R^{1/2} \prod_{m=1}^{M} (e^{-iu} - \overline{\beta_m}). \tag{3.11}
\]
The proof of the required properties is the same as in the continuous-time case. It is therefore omitted.

Finally, notice that the temporary assumption \(\ell_0 \neq 0\) may easily be dropped. Indeed, in the general case, we may always write \(\ell(z) = z^k \tilde{\ell}(z)\) for some \(k \leq M\) and \(\tilde{\ell}(0) \neq 0\). Then the factorization for \(\tilde{\ell}\) also applies to \(\ell\) since \(1 + |\ell(\cdot)|^2 = 1 + |\tilde{\ell}(\cdot)|^2\) on the unit circle.

For the discrete-time kinetic energy \(\ell(z) = \alpha(z - 1)\), we may use the factorization
\[
|\ell(\text{e}^{iu})|^2 + 1 = \frac{\alpha^2}{\beta} (e^{-iu} - \beta)(e^{iu} - \beta) =: \lambda_\ell(u) \overline{\lambda_\ell(u)}, \tag{3.12}
\]
where
\[
\lambda_\ell(u) := \frac{\alpha}{\sqrt{\beta}} (e^{-iu} - \beta)
\]
and \(\beta\) is the constant in (2.8). In this case we have
\[
\frac{1}{\lambda_\ell(u)} = -\frac{1}{\alpha \sqrt{\beta}} \frac{1}{1 - e^{-iu}/\beta} = -\frac{1}{\alpha \sqrt{\beta}} \sum_{\tau=0}^{\infty} \beta^{-\tau} e^{-i\tau u} \tag{3.13}
\]
as a version of (3.9).
3.2. Application of the prediction technique.

3.2.1. Discrete time. We first recall some notions used in the analytical prediction technique. Define the spaces of spectrally negative and spectrally positive functions

\[ L^2_{\leq 0} := \text{span}\{e^{i\tau u}, \tau \leq 0, \tau \in \mathbb{Z} \mid L^2(T, \Lambda)\}, \]

\[ L^2_{> 0} := \text{span}\{e^{i\tau u}, \tau > 0, \tau \in \mathbb{Z} \mid L^2(T, \Lambda)\}, \]

where \( \Lambda \) denotes the Lebesgue measure. We will need a special class of outer functions. We do not recall the formal definition of an outer function ([5], p.342); instead we use the following characterization ([5], Theorem 17.23): a function \( \gamma \in L^2(T, \Lambda) \) is conjugate to an outer function if and only if

\[ \text{span}\{\gamma e^{i\tau u}, \tau \leq 0, \tau \in \mathbb{Z} \mid L^2(T, \Lambda)\} = L^2_{\leq 0}. \] (3.14)

We stress that these functions are complex-conjugate to outer functions in the form chosen in [5]. In what follows, however, we call them simply outer functions; this should not lead to any misunderstanding.

We now pass to the optimization problem. By (2.6), we have to compute

\[ \text{ERR}^+_A = \min_{\tilde{g} \in \tilde{G}_{\leq 0}} \int_T \left| \tilde{g}(u) - \frac{1}{|\ell(e^{iu})|^2 + 1} \right|^2 \left( |\ell(e^{iu})|^2 + 1 \right) \mu(du), \]

where

\[ \tilde{G}_{\leq 0} := \text{span}\{e^{i\tau u}, \tau \leq 0, \tau \in \mathbb{Z} \mid L^2(T, \mu)\}. \]

Assume that the spectral measure on \( T \) has a density satisfying Kolmogorov’s regularity condition, that is, \( \mu(du) = f(u) \, du \) and

\[ \int_T |\ln f(u)| \, du < \infty. \] (3.15)

The classical prediction technique suggests finding a factorization

\[ f(u) = \gamma_f(u)\gamma_f(u) = |\gamma_f(u)|^2, \quad u \in T, \] (3.16)

\[ |\ell(e^{iu})|^2 + 1 = \lambda_\ell(u)\lambda_\ell(u) = |\lambda_\ell(u)|^2, \quad u \in T, \] (3.17)

where \( \gamma_f \) is an outer function and, as above, \( \lambda_\ell \) satisfies the conditions (3.9) and (3.10). Notice that the assumption (3.15) guarantees the existence of a factorization (3.16); see [5], Theorem 17.16. The factorization (3.17) in the case of a polynomial \( \ell \) was given in (3.11).

**Theorem 1.** Let \( Q_{> 0} \) be the orthogonal projection of \( \gamma_f/\lambda_\ell \) to \( L^2_{> 0} \) in the Hilbert space \( L^2(T, \Lambda) \). Then the optimal adaptive approximation is given by the formula

\[ X(t) = \int_T \hat{g}_*(u) e^{itu} \mathcal{W}(du), \]

where

\[ \hat{g}_*(u) = \frac{1}{|\lambda_\ell|^2} - \frac{Q_{> 0}}{\lambda_\ell \gamma_f}. \]

The error in the adaptive approximation is given by \( \text{ERR}^+_A = \|Q_{> 0}\|_2^2. \)
Proof. We have
\[
\text{ERR}^+_A = \min \int_T \left| \lambda_\ell(u) \hat{g}(u) - \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du) = \min \int_T \left| (\lambda_\ell \gamma_f \hat{g})(u) - \frac{\gamma_f(u)}{\lambda_\ell(u)} \right|^2 du.
\] (3.18)

Consider an arbitrary \( \hat{g} \in \hat{G}_{\leq 0} \). There is no loss of generality in assuming that \( \lambda_\ell \hat{g} \in L^2_0 \) (otherwise the integral in (3.18) is infinite). Then by (3.10) we have
\[
\lambda_\ell \gamma_f \hat{g} \in \mathcal{B}_6 \text{ or, equivalently, }
\lambda_\ell \gamma_f \hat{g} \in \text{span}\{\gamma_f e^{i\tau u}, \tau \leq 0, \tau \in \mathbb{Z} | L^2_0(T, \Lambda)\} = L^2_0;
\]
where the equality holds by (3.14) because \( \gamma_f \) is an outer function.

On the other hand, since \( \gamma_f \in L^2_0(T, \Lambda) \) and \( |\lambda_\ell| > 1 \), we have
\[
Q := \frac{\gamma_f}{\lambda_\ell} \in L^2(T, \Lambda).
\]

Consider the following unique orthogonal decomposition in \( L^2(T, \Lambda) \):
\[
Q := \frac{\gamma_f}{\lambda_\ell} = Q_{\leq 0} + Q_{> 0},
\]
where \( Q_{\leq 0} \in L^2_{\leq 0} \) and \( Q_{> 0} \in L^2_{> 0} \). Since the spaces \( L^2_{\leq 0} \) and \( L^2_{> 0} \) are orthogonal, we have
\[
\text{ERR}^+_A \geq \|Q_{> 0}\|_2^2.
\]
Moreover, the equality
\[
\text{ERR}^+_A = \|Q_{> 0}\|_2^2
\] (3.19)
is attained if and only if
\[
\hat{g} = \hat{g}_* := \frac{Q_{\leq 0}}{\lambda_\ell \gamma_f} = \frac{Q - Q_{> 0}}{\lambda_\ell \gamma_f} = \frac{1}{|\lambda_\ell|^2} - \frac{Q_{> 0}}{\lambda_\ell \gamma_f}.
\] (3.20)

It remains to verify that \( \hat{g}_* \in \hat{G}_{\leq 0} \). Since \( \gamma_f \) is an outer function, it follows from (3.14) that
\[
Q_{\leq 0} \in L^2_{\leq 0} = \text{span}\{\gamma_f e^{i\tau u}, \tau \leq 0, \tau \in \mathbb{Z} | L^2(T, \Lambda)\},
\]
which is equivalent to
\[
\frac{Q_{\leq 0}}{\gamma_f} \in \text{span}\{e^{i\tau u}, \tau \leq 0, \tau \in \mathbb{Z} | L^2(T, \mu)\} = \hat{G}_{\leq 0}.
\]
Finally, we obtain from (3.9) the required inclusion \( \hat{g}_* = Q_{\leq 0}/(\lambda_\ell \gamma_f) \in \hat{G}_{\leq 0} \). □

For the discrete-time kinetic energy \( \ell(z) = \alpha(z - 1) \) we may proceed as follows using the decomposition (3.12). Since \( \gamma_f \) is an outer function, it belongs to \( L^2_{\leq 0} \). Taking the Fourier series expansion
\[
\gamma_f(u) := \sum_{j=0}^{\infty} \hat{\gamma}_j e^{-iju}, \quad u \in \mathbb{T},
\] (3.21)
and multiplying (3.13) and (3.21), we obtain

\[ Q(u) = \frac{-1}{\alpha \sqrt{\beta}} \sum_{\tau=0}^{\infty} \beta^{-\tau} e^{i\tau u} \sum_{j=0}^{\infty} \tilde{\gamma}_j e^{-ij u} = \frac{-1}{\alpha \sqrt{\beta}} \sum_{n=-\infty}^{\infty} \left[ \left( \sum_{j=\max(-n,0)}^{\infty} \tilde{\gamma}_j \beta^{-(n+j)} \right) e^{inu} \right] \]

\[ = \frac{-1}{\alpha \sqrt{\beta}} \sum_{n=-\infty}^{\infty} \left[ \left( \sum_{j=\max(-n,0)}^{\infty} \tilde{\gamma}_j \beta^{-j} \right) \beta^{-n} e^{inu} \right]. \]

Hence,

\[ Q_{>0}(u) = \frac{-1}{\alpha \sqrt{\beta}} \left( \sum_{j=0}^{\infty} \tilde{\gamma}_j \beta^{-j} \right) \left( \sum_{n=1}^{\infty} \beta^{-n} e^{inu} \right) \]

\[ = -K \frac{\alpha \sqrt{\beta}}{\beta - e^{iu}} = \frac{\alpha \sqrt{\beta}}{\beta} \frac{e^{iu}}{\lambda(\ell(u))}, \] (3.23)

with the constant

\[ K := \sum_{j=0}^{\infty} \tilde{\gamma}_j \beta^{-j}. \] (3.24)

It follows from (3.19) and (3.22) that

\[ \text{ERR}_A = ||Q_{>0}||_2^2 = \frac{1}{\alpha^2 \beta} |K|^2 \frac{2\pi}{\beta^2 - 1} \]

\[ = \frac{2\pi}{\alpha^2 \beta^2} |K|^2 \frac{1}{\beta - 1/\beta} = \frac{2\pi}{\beta^2 \sqrt{1 + 4\alpha^2}} |K|^2, \] (3.25)

where we have used the identity

\[ \beta - \frac{1}{\beta} = \frac{\sqrt{1 + 4\alpha^2}}{\alpha^2}. \] (3.26)

We also obtain from (3.20) and (3.23) that

\[ \tilde{g}_*(u) = \frac{1}{|\lambda(\ell(u))|^2} \left( 1 - \frac{K e^{iu}}{\beta \gamma_f(u)} \right) = \frac{1}{2\alpha^2(1 - \cos u) + 1} \left( 1 - \frac{K e^{iu}}{\beta \gamma_f(u)} \right), \quad u \in \mathbb{T}. \] (3.27)

In the case when \( \ell \) is an arbitrary polynomial, we can use (3.11) to construct a partial fraction decomposition of \( 1/\lambda(\ell) \) into a linear combination of fractions of the form \( 1/(1 - e^{iu} \beta^{-1}) \) provided that the numbers \( \beta_m \) are pairwise distinct. Then the considerations above can be applied to every fraction separately.

One can express the outer function \( \gamma_f \) and the constant \( K \) explicitly in terms of the spectral density \( f \). To do this, consider the function

\[ q(z) := \exp \left\{ \frac{1}{2\pi} \int_{\mathbb{T}} \frac{e^{iu} + z}{e^{iu} - z} \ln \sqrt{f(u)} \, du \right\}, \quad |z| < 1. \]
By Theorem 17.16 in [5], the radial limits of its absolute value $|q|$ are given almost everywhere with respect to the Lebesgue measure by the formula

$$\lim_{r \uparrow 1} |q(re^{iu})| = \sqrt{f(u)}, \quad u \in \mathbb{T}.$$ 

Since the function $\sqrt{f}$ is square-integrable, it follows from Theorem 17.16(c) in [5] that $q$ belongs to the Hardy space $H^2$ on the unit disc. Defining

$$\gamma_f(u) = \lim_{r \uparrow 1} q(re^{iu}), \quad u \in \mathbb{T},$$

we clearly have $|\gamma_f(u)|^2 = f(u)$ for $u \in \mathbb{T}$. Also, $\gamma_f$ is (the complex-conjugate of) an outer function by Definition 17.14 in [5]. The Fourier series representation (3.21) of $\gamma_f$ translates into the Taylor series representation of $q:

$$q(z) = \sum_{j=0}^{\infty} \tilde{\gamma}_j z^j, \quad |z| < 1.$$ 

Returning to the case when $\ell(z) = \alpha(z - 1)$, we obtain from (3.24) that

$$K = q\left(\frac{1}{\beta}\right) = \exp\left\{\frac{1}{4\pi} \int_{\mathbb{T}} \frac{e^{-iu} + \beta^{-1}}{e^{-iu} - \beta^{-1}} \ln f(u) \, du \right\}.$$ 

Recalling (3.25) and making straightforward transformations, we arrive at the following result.

**Theorem 2.** In the discrete-time case with $\ell(z) = \alpha(z - 1)$, the additional adaptivity error is given by

$$\text{ERR}^+_A = \frac{2\pi}{\beta^2 \sqrt{1 + 4\alpha^2}} \exp\left\{\frac{1}{2\pi} \int_{\mathbb{T}} \frac{\beta^2 - 1}{\beta^2 + 1 - 2\beta \cos u} \ln f(u) \, du \right\}, \quad (3.28)$$

where $\beta = (2\alpha^2 + 1 + \sqrt{1 + 4\alpha^2})/(2\alpha^2)$.

In the arguments above, we assumed that $\int_{\mathbb{T}} |\ln f(u)| \, du < \infty$. But the formula (3.28) remains valid even when $\int_{\mathbb{T}} |\ln f(u)| \, du = \infty$. Indeed, since $f$ is a density, the latter condition is equivalent to

$$\int_{\mathbb{T}} \ln f(u) \, du = -\infty, \quad (3.29)$$

and (3.28) states that $\text{ERR}^+_A = 0$. This result can easily be explained. It is known (in analytical form; see, for example, [6], pp.48–50) that under (3.29) (or if the spectral measure is singular; see [6], Corollary 1 on p.46) the future of the process $X$ can be predicted **perfectly** on the basis of its past, so that there is no difference between the non-adaptive and adaptive approximations.

We now look at the approximation problem in the case when $\alpha \downarrow 0$. This means that we give less importance to the kinetic energy of the approximating process.
than to the closeness of the processes. As $\alpha \downarrow 0$, we have $\beta = \alpha^{-2} + O(1) \to \infty$ and (3.28) yields that

$$\text{ERR}^+_A \sim 2\pi \alpha^4 \exp\left\{ \frac{1}{2\pi} \int_T \ln f(u) \, du \right\}. \tag{3.30}$$

The right-hand side looks very much like the classical Kolmogorov–Szegő formula. Let us explain this similarity. We recall that the classical prediction problem asks us to predict $B(\tau)$ for $\tau = 1, 2, \ldots$ on the basis of $B(0), B(-1), \ldots$. In particular, in the case of one-step prediction $\tau = 1$, the Kolmogorov–Szegő formula ([3], p. 257, [7], Theorem 5.8.1, or [6], pp. 48–50) states that the mean-square error of the optimal prediction is given by

$$\sigma^2_{\text{pred}} := \min_{\hat{h} \in G_{\leq 0}} \int_T |e^{iu} - \hat{h}(u)|^2 f(u) \, du = 2\pi \exp\left\{ \frac{1}{2\pi} \int_T \ln f(u) \, du \right\}.$$  

By (2.7), the optimal non-adaptive strategy is given by

$$\hat{g^*_s}^{(\text{nad})}(u) = 1 + (e^{iu} + e^{-iu} - 2)\alpha^2 + O(\alpha^4).$$

It is therefore natural to make the following ansatz for the optimal adaptive strategy:

$$\hat{g^*_s}^{(\text{ad})}(u) = 1 + (\hat{w}(u) + e^{-iu} - 2)\alpha^2 + O(\alpha^4),$$

where $\hat{w}$ is a function in $L^2_{\leq 0}$. The additional adaptivity error $\text{ERR}^+_A$ is then

$$\int_T \left| \frac{\hat{g^*_s}^{(\text{ad})}(u) - \frac{1}{|e^{iu}|^2 + 1}}{|\ell(e^{iu})|^2 + 1} \right|^2 (|\ell(e^{iu})|^2 + 1) \mu(du)$$

$$= \int_T \left| 1 + (\hat{w}(u) + e^{-iu} - 2)\alpha^2 + O(\alpha^4) - \frac{1}{\alpha^2 |e^{iu} - 1|^2 + 1} \right|^2 (1 + O(\alpha^2)) \mu(du)$$

$$\sim \alpha^4 \int_T |\hat{w}(u) - e^{iu}|^2 \mu(du).$$

Thus the function $\hat{w}$ should be chosen as the solution of the classical prediction problem and we should have $\text{ERR}^+_A \sim \alpha^4 \sigma^2_{\text{pred}}$. This explains the similarity between (3.30) and Kolmogorov’s formula. We finally observe that, by (2.9),

$$\text{ERR}_{\text{NA}} \sim \alpha^2 \int_T |e^{iu} - 1|^2 f(u) \, du$$

as $\alpha \downarrow 0$. Thus, for small $\alpha$, the price for not knowing the future is small compared to the error in the non-adaptive approximation.

3.2.2. Continuous time. Here the approach and the result are very much the same as in the case of stationary sequences except for some issues of integrability. We only replace $\mathbb{T}$ by $\mathbb{R}$ and redefine the spaces $L^2_{\leq 0}$ and $L^2_{> 0}$ in $L_2(\mathbb{R}, \Lambda)$ as the spaces of Fourier transforms of functions supported on $\mathbb{R}_-$ and $\mathbb{R}_+$ respectively.
We again use the class of outer functions, this time with respect to the lower half-plane ([8], p.36), with the following characterization ([8], p.39): a function \( \gamma \in L^2(\mathbb{R}, \Lambda) \) is outer for the lower half-plane if and only if
\[
\text{span}\{\gamma e^{i\tau u}, \tau \leq 0 \mid L^2(\mathbb{R}, \Lambda)\} = L^2_{\leq 0}.
\] (3.31)

The Kolmogorov regularity condition now says that \( \mu(du) = f(u) \, du \) and
\[
\int_{\mathbb{R}} \frac{|\ln f(u)|}{1 + u^2} \, du < \infty.
\] (3.32)

This condition ensures the existence of a factorization
\[
f(u) = \gamma_f(u)\overline{\gamma_f(u)} = |\gamma_f(u)|^2, \quad u \in \mathbb{R},
\]
where \( \gamma_f \) is an outer function ([8], p.38).

For the energy function \( \ell(\cdot) \) we need a factorization
\[
|\ell(iu)|^2 + 1 = \lambda_\ell(u)\overline{\lambda_\ell(u)} = |\lambda_\ell(u)|^2, \quad u \in \mathbb{R},
\]
where \( \lambda_\ell \) possesses the properties (3.5) and (3.6). It was shown in (3.7) how to construct such a factorization for polynomials.

The construction of the optimal adaptive approximation and the calculation of the approximation error are now done exactly as in the discrete-time case, but we repeat the approach for completeness of exposition.

**Theorem 3.** Let \( Q_{>0} \) be the orthogonal projection of \( \gamma_f/\lambda_\ell \) to \( L^2_{>0} \) in the Hilbert space \( L^2(\mathbb{R}, \Lambda) \). Then the optimal adaptive approximation is given by \( X(t) = \int_{\mathbb{R}} \hat{g}_*(u)e^{itu} \mathcal{W}(du) \), where
\[
\hat{g}_*(u) = \frac{1}{|\lambda_\ell|^2} - \frac{Q_{>0}}{\lambda_\ell \gamma_f}.
\]

The error in the adaptive approximation is given by \( \text{ERR}_A^+ = \|Q_{>0}\|_2^2 \).

**Proof.** We have to compute
\[
\text{ERR}_A^+ = \min_{\hat{g} \in \hat{G}_{\leq 0}} \int_{\mathbb{R}} \left| \hat{g}(u) - \frac{1}{|\ell(iu)|^2 + 1} \right|^2 \mu(du),
\]
where \( \hat{G}_{\leq 0} := \text{span}\{e^{i\tau u}, \tau \leq 0 \mid L^2(\mathbb{R}, \mu)\} \). Using factorizations, we have
\[
\text{ERR}_A^+ = \min_{\hat{g} \in \hat{G}_{\leq 0}} \int_{\mathbb{R}} \frac{1}{\lambda_\ell(u)} \left| \frac{1}{\lambda_\ell(u)} \right|^2 \mu(du)
\]
\[
= \min_{\hat{g} \in \hat{G}_{\leq 0}} \int_{\mathbb{R}} \left( \lambda_\ell \gamma_f \hat{g} \right)(u) - \frac{\gamma_f(u)}{\lambda_\ell(u)} \right|^2 du.
\] (3.33)

Consider an arbitrary \( \hat{g} \in \hat{G}_{\leq 0} \). There is no loss of generality in assuming that \( \lambda_\ell \hat{g} \in L^2(\mathbb{R}, \mu) \) (otherwise the integral in (3.33) is infinite). Then, by (3.6), we have \( \lambda_\ell \hat{g} \in \hat{G}_{\leq 0} \). This is equivalent to
\[
\lambda_\ell \gamma_f \hat{g} \in \text{span}\{\gamma_f e^{i\tau u}, \tau \leq 0 \mid L^2(\mathbb{R}, \Lambda)\} = L^2_{\leq 0},
\]
where the last equality holds by (3.31) because $\gamma_f$ is an outer function.

On the other hand, since $\gamma_f \in L_2(\mathbb{R}, \Lambda)$ and $|\lambda_\ell| \geq 1$, we have

$$Q := \frac{\gamma_f}{\lambda_\ell} \in L_2(\mathbb{R}, \Lambda).$$

Consider the following unique orthogonal decomposition in $L_2(\mathbb{R}, \Lambda)$:

$$Q := \frac{\gamma_f}{\lambda_\ell} = Q_{\leq 0} + Q_{> 0},$$

where $Q_{\leq 0} \in L^2_{\leq 0}$ and $Q_{> 0} \in L^2_{> 0}$. Since the spaces $L^2_{\leq 0}$ and $L^2_{> 0}$ are orthogonal, we clearly have

$$\text{ERR}_A^+ \geq \|Q_{> 0}\|_2^2.$$ 

Furthermore, the equality

$$\text{ERR}_A^+ = \|Q_{> 0}\|_2^2$$ 

is attained if and only if

$$\hat{g} = \hat{g}_* := \frac{Q_{\leq 0}}{\lambda_\ell \gamma_f} = Q - \frac{Q_{> 0}}{\lambda_\ell \gamma_f} = \frac{1}{|\lambda_\ell|^2} - \frac{Q_{> 0}}{\lambda_\ell \gamma_f}. \quad (3.35)$$

It remains to prove the inclusion $\hat{g}_* \in \hat{G}_{\leq 0}$. Since $\gamma_f$ is an outer function, it follows from (3.31) that

$$Q_{\leq 0} \in L^2_{\leq 0} = \text{span}\{\gamma_f e^{i\tau u}, \tau \leq 0 \mid L_2(\mathbb{R}, \Lambda)\}$$

or, equivalently,

$$\frac{Q_{\leq 0}}{\gamma_f} \in \text{span}\{e^{i\tau u}, \tau \leq 0 \mid L_2(\mathbb{R}, \mu)\} = \hat{G}_{\leq 0}.$$ 

Finally, we obtain from (3.5) the required inclusion $\hat{g}_* = Q_{\leq 0}/(\lambda_\ell \gamma_f) \in \hat{G}_{\leq 0}$. □

For the continuous-time kinetic energy $\ell(z) = \alpha z$, by taking the Fourier integral representation

$$\gamma_f(u) := \int_0^\infty \hat{\gamma}(\tau) e^{-i\tau u} \, d\tau, \quad u \in \mathbb{R}, \quad (3.36)$$

and multiplying (3.8) and (3.36), we obtain

$$Q(u) = \int_0^\infty \hat{\gamma}(\tau_1) e^{-i\tau_1 u} \, d\tau_1 \cdot \frac{1}{\alpha} \int_0^\infty e^{-\tau_2/\alpha} e^{i\tau_2 u} \, d\tau_2$$

$$= \frac{1}{\alpha} \int_0^\infty \int_0^\infty \hat{\gamma}(\tau_1) e^{-\tau_2/\alpha} e^{i(\tau_2 - \tau_1) u} \, d\tau_1 \, d\tau_2$$

$$= \frac{1}{\alpha} \int_{-\infty}^\infty \left( \int_{\max(0,-\tau)}^\infty \hat{\gamma}(\tau_1) e^{-(\tau + \tau_1)/\alpha} \, d\tau_1 \right) e^{i\tau u} \, d\tau.$$
Hence,

\[
Q_{>0}(u) = \frac{1}{\alpha} \int_0^\infty \left( \int_0^\infty \tilde{\gamma}(\tau_1) e^{-\tau_1/\alpha} d\tau_1 \right) e^{i\tau u} d\tau
\]

\[
= \frac{1}{\alpha} \int_0^\infty \tilde{\gamma}(\tau_1) e^{-\tau_1/\alpha} d\tau_1 \cdot \int_0^\infty e^{-\tau/\alpha} e^{i\tau u} d\tau
\]

\[
= : \mathcal{K} \int_0^\infty e^{-\tau/\alpha} e^{i\tau u} d\tau = \frac{\alpha \mathcal{K}}{1 - i\alpha u}
\]

(3.37)

with the constant

\[
\mathcal{K} := \frac{1}{\alpha} \int_0^\infty \tilde{\gamma}(\tau) e^{-\tau/\alpha} d\tau.
\]

(3.38)

It follows from (3.34) and (3.37) that

\[
\text{ERR}_A^+ = \|Q_{>0}\|_2^2 = \alpha^2 |\mathcal{K}|^2 \int_\mathbb{R} \frac{du}{1 + \alpha^2 u^2} = \pi \alpha |\mathcal{K}|^2.
\]

(3.39)

Furthermore, by using (3.35) and (3.37), we obtain a continuous-time analogue of (3.27):

\[
\tilde{g}_\ast(u) = \frac{1}{|\lambda_\ell(u)|^2} \left( 1 - \frac{\alpha \mathcal{K}}{\gamma_f(u)} \right) = \frac{1}{1 + \alpha^2 u^2} \left( 1 - \frac{\alpha \mathcal{K}}{\gamma_f(u)} \right), \quad u \in \mathbb{R}.
\]

(3.40)

To derive an explicit formula for the outer function \(\gamma_f\) and the constant \(\mathcal{K}\) in terms of the spectral density \(f\), we consider the function

\[
q(z) := \exp \left\{ \frac{1}{\pi i} \int_\mathbb{R} \frac{u z + 1}{u - z} \ln \sqrt{f(u)} \frac{1}{u^2 + 1} du \right\}, \quad \text{Im} z > 0.
\]

It is known ([8], p.37) that \(q(z)\) belongs to the Hardy space \(H^2\) on the upper half-plane and the boundary limits of its absolute value are given almost everywhere with respect to the Lebesgue measure by the formula

\[
\lim_{v \downarrow 0} |q(u + vi)| = \sqrt{f(u)}, \quad u \in \mathbb{R}.
\]

Defining

\[
\gamma_f(u) = \lim_{v \downarrow 0} q(u + vi), \quad u \in \mathbb{R},
\]

we evidently have \(|\gamma_f(u)|^2 = f(u)\) for \(u \in \mathbb{R}\). Furthermore, \(\gamma_f\) satisfies (3.31) by [8], pp.37–39. The Fourier representation of \(\gamma_f\) (see (3.36)) still holds in the upper half-plane:

\[
q(z) = \int_0^\infty \frac{\tilde{\gamma}(\tau)}{\gamma_f(\tau)} e^{i\tau z} d\tau, \quad \text{Im} z > 0.
\]

It follows from (3.38) that

\[
\mathcal{K} = \frac{1}{\alpha} q\left( \frac{i}{\alpha} \right) = \frac{1}{\alpha} \exp \left\{ \frac{1}{2\pi} \int_\mathbb{R} \frac{\alpha + ui}{1 + ui\alpha} \ln f(u) \frac{1}{u^2 + 1} du \right\}.
\]

Recalling (3.39), we arrive at the following result.
Theorem 4. In the continuous-time case with $\ell(z) = \alpha z$, the additional adaptivity error is given by

$$\text{ERR}_A^+ = \frac{\pi}{\alpha} \exp \left\{ \frac{\alpha}{\pi} \int_{\mathbb{R}} \frac{\ln f(u)}{1 + \alpha^2 u^2} \, du \right\}.$$ 

As in the discrete-time case, if the Kolmogorov condition (3.32) does not hold (or if the spectral measure $\mu$ does not possess a density at all), then the perfect prediction of the future is possible and we therefore have $\text{ERR}_A^+ = 0$.

§ 4. Examples of adaptive least-energy approximation

In the following examples we consider kinetic energy unless otherwise stated. Hence we put $\ell(z) = \alpha z$ for continuous time and $\ell(z) = \alpha (z - 1)$ for discrete time. Here $\alpha > 0$ is a fixed scaling parameter.

4.1. Discrete time.

Autoregressive sequences. A sequence $(B(t))_{t \in \mathbb{Z}}$ of complex random variables is said to be autoregressive if it satisfies the equation $B(t) = \rho B(t - 1) + \xi(t)$, where $|\rho| < 1$ and $(\xi(t))_{t \in \mathbb{Z}}$ is a sequence of centred uncorrelated random variables such that $\sigma^2 := \mathbb{E}|\xi(t)|^2$ is independent of $t$. In this case we have a representation

$$B(t) = \sum_{j=0}^{\infty} \rho^j \xi(t - j), \quad t \in \mathbb{Z}.$$ 

The uncorrelated sequence admits a spectral representation

$$\xi(t) = \int_{\mathbb{T}} e^{itu} \mathcal{W}(du), \quad (4.1)$$

where $\mathcal{W}$ is a complex-valued centred random measure with uncorrelated values on $\mathbb{T}$ controlled by the normalized Lebesgue measure

$$\mu(du) := \frac{\sigma^2 \, du}{2\pi}.$$ 

Therefore,

$$B(t) = \int_{\mathbb{T}} \sum_{j=0}^{\infty} \rho^j e^{i(t-j)u} \mathcal{W}(du) = \int_{\mathbb{T}} \frac{1}{1 - \rho e^{-iu}} e^{itu} \mathcal{W}(du).$$

We see that the spectral measure for $B$ is

$$\mu(du) := \frac{\sigma^2 \, du}{2\pi |1 - \rho e^{-iu}|^2}, \quad (4.2)$$

which can also be found in [9], Ch. 2, § 17, Example 3, or [7], Example 4.4.2. By (4.2) and (2.9), the error in the non-adaptive approximation is equal to

$$\text{ERR}_{\text{NA}} = \frac{\sigma^2}{1 - |\rho|^2} \left( 1 - \frac{1}{\sqrt{1 + 4\alpha^2}} \frac{\beta^2 - |\rho|^2}{|\beta - \rho|^2} \right), \quad (4.3)$$
where $\beta = \beta(\alpha)$ is defined in (2.8) (see [1] for a detailed calculation).

On the other hand, the spectral measure can be factorized as

$$f(u) = \frac{\sigma^2}{2\pi} |1 - \rho e^{-iu}|^{-2} = \gamma_f(u)\overline{\gamma_f(u)},$$

where

$$\gamma_f(u) := \frac{\sigma}{\sqrt{2\pi}} (1 - \rho e^{-iu})^{-1} = \frac{\sigma}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \rho^j e^{-iju}. \quad (4.4)$$

Hence, by (3.24),

$$K = \frac{\sigma^2}{\beta^2 \sqrt{1 + 4\alpha^2}} \frac{\sigma^2}{2\pi(1 - \rho/\beta)^2} = \frac{\sigma^2}{\sqrt{1 + 4\alpha^2} |\beta - \rho|^2}. \quad (4.5)$$

whereas by (3.25) we have

$$\text{ERR}_A^+ = \frac{2\pi}{\beta^2 \sqrt{1 + 4\alpha^2}} \frac{\sigma^2}{2\pi(1 - \rho/\beta)^2} = \frac{\sigma^2}{\sqrt{1 + 4\alpha^2} |\beta - \rho|^2}.$$  

Using (4.3), we conclude that

$$\text{ERR}_A = \text{ERR}_{NA} + \text{ERR}_A^+$$

$$= \frac{\sigma^2}{1 - |\rho|^2} \left(1 - \frac{1}{\sqrt{1 + 4\alpha^2} |\beta - \rho|^2}\right) + \frac{\sigma^2}{\sqrt{1 + 4\alpha^2} |\beta - \rho|^2}.$$  

Now (3.27), (4.4) and (4.5) yield an expression for the optimal prediction:

$$\hat{g}_*(u) = |\lambda|^{-2} \left(1 - \frac{\sigma}{\sqrt{2\pi}} \frac{e^{iu} \sqrt{2\pi} (1 - \rho e^{-iu})}{\beta \sigma}\right)$$

$$= |\lambda|^{-2} \left(1 - \frac{e^{iu} - \rho}{\beta - \rho}\right) = |\lambda|^{-2} \frac{\beta - e^{iu}}{\beta - \rho}$$

$$= \frac{\beta}{\alpha^2(\beta - \rho)(\beta - e^{-iu})} = \frac{1}{\alpha^2(\beta - \rho)} \sum_{j=0}^{\infty} \beta^{-j} e^{-iju}.$$  

**Uncorrelated sequences.** We regard an uncorrelated sequence as a special case of an autoregressive sequence with $\rho = 0$. The best adaptive approximation is given by

$$\hat{g}_*(u) = \frac{1}{\alpha^2} \sum_{j=0}^{\infty} \beta^{-j-1} e^{-iju}$$

and the approximation errors are

$$\text{ERR}_A^+ = \frac{\sigma^2}{\beta^2 \sqrt{1 + 4\alpha^2}},$$

$$\text{ERR}_A = \sigma^2 \left(1 - \frac{1}{\sqrt{1 + 4\alpha^2}}\right) + \frac{\sigma^2}{\beta^2 \sqrt{1 + 4\alpha^2}} = \sigma^2 \left(1 - \frac{1}{\alpha^2 \beta}\right).$$

Here we again use the identity (3.26) in the last step.
The simplest moving average sequence. A sequence \((B(t))_{t \in \mathbb{Z}}\) of complex random variables is called the simplest moving average sequence if it admits a representation \(B(t) = \xi(t) + \rho \xi(t-1)\), where \((\xi(t))_{t \in \mathbb{Z}}\) is a sequence of centred uncorrelated complex random variables such that \(\sigma^2 := \mathbb{E}|\xi(t)|^2\) is independent of \(t\).

Using (4.1), we obtain that
\[
B(t) = \int_{\mathbb{T}} (1 + \rho e^{-iu}) e^{itu} \mathcal{W}(du), \quad t \in \mathbb{Z}.
\]

We conclude that the spectral measure for \(B\) is
\[
\mu(du) := \frac{\sigma^2 |1 + \rho e^{-iu}|^2 du}{2\pi}; \quad (4.6)
\]
see Example 4.4.1 in [7]. By (4.6) and (2.9), the error in the non-adaptive approximation is equal to
\[
\text{ERR}_{NA} = \sigma^2 \left( 1 + |\rho|^2 - \frac{1}{\sqrt{1 + 4\alpha^2}} \left( 1 + |\rho|^2 + \frac{\rho + \overline{\rho}}{\beta} \right) \right) \quad (4.7)
\]
where \(\beta = \beta(\alpha)\) is defined in (2.8) (see [1] for a detailed calculation).

The form of factorization of the spectral density depends on \(|\rho|\). If \(|\rho| < 1\), then we have a factorization
\[
f(u) = \frac{\sigma^2}{2\pi} |1 + \rho e^{-iu}|^2 = \gamma_f(u) \overline{\gamma_f(u)}, \quad \text{where } \gamma_f(u) := \frac{\sigma}{\sqrt{2\pi}} (1 + \rho e^{-iu}).
\]
Hence, by (3.24),
\[
\mathcal{K} = \frac{\sigma}{\sqrt{2\pi}} \left( 1 + \frac{\rho}{\beta} \right),
\]
whereas by (3.25) we have
\[
\text{ERR}_A^+ = \frac{2\pi}{\beta^2 \sqrt{1 + 4\alpha^2}} \frac{\sigma^2 |1 + \rho/\beta|^2}{2\pi} = \frac{\sigma^2 |1 + \rho/\beta|^2}{\beta^2 \sqrt{1 + 4\alpha^2}}.
\]

Using (4.7), we find that
\[
\text{ERR}_A = \text{ERR}_{NA} + \text{ERR}_A^+ = \sigma^2 \left( 1 + |\rho|^2 - \frac{1}{\sqrt{1 + 4\alpha^2}} \left( 1 + |\rho|^2 + \frac{\rho + \overline{\rho}}{\beta} \right) \right) + \frac{\sigma^2 |1 + \rho/\beta|^2}{\beta^2 \sqrt{1 + 4\alpha^2}}
\]
\[
= \sigma^2 \left( 1 + |\rho|^2 - \frac{1}{\beta \alpha^2} - \frac{\rho + \overline{\rho}}{\beta^2 \alpha^2} - \frac{|\rho|^2 (2\alpha^2 + 1)}{\beta^2 \alpha^4} \right).
\]

In our setting,
\[
\gamma_f(\lambda_t)(u) = \frac{\sigma}{\sqrt{2\pi}} (1 + \rho e^{-iu}) \frac{\sqrt{\beta}}{\alpha} \frac{1}{e^{iu} - \beta} = \frac{-\sigma}{\sqrt{2\pi} \alpha \sqrt{\beta}} (1 + \rho e^{-iu}) \frac{1}{1 - e^{iu}/\beta}
\]
\[
= \frac{-\sigma}{\sqrt{2\pi} \alpha \sqrt{\beta}} (1 + \rho e^{-iu}) \sum_{j=0}^{\infty} \beta^{-j} e^{iju}.
\]
It follows that
\[ Q_{\leq 0}(u) = - \frac{\sigma}{\sqrt{2\pi} \alpha \sqrt{\beta}} \left( \rho e^{-iu} + 1 + \frac{\rho}{\beta} \right). \]

Now (3.27) yields the optimal prediction
\[ \hat{g}_*(u) = Q_{\leq 0} \gamma_f = \frac{-(\rho e^{-iu} + 1 + \rho/\beta)}{\alpha^2(1 + \rho e^{-iu})(e^{-iu} - \beta)}. \]

When \(|\rho| < 1\) and \(\rho \neq -1/\beta\), this expression may be rewritten as
\[ \hat{g}_*(u) = \frac{1}{\alpha^2(\rho \beta + 1)} \left[ \frac{\rho^2}{\beta(1 + \rho e^{-iu})} - \frac{1 + \rho/\beta + \rho \beta}{e^{-iu} - \beta} \right] \]
\[ = \frac{1}{\alpha^2 \beta(\rho \beta + 1)} \left[ \frac{\rho^2}{1 + \rho e^{-iu}} + \frac{1 + \rho/\beta + \rho \beta}{1 - e^{-iu}/\beta} \right] \]
\[ = \frac{1}{\alpha^2 \beta(\rho \beta + 1)} \sum_{j=0}^{\infty} \left[ (-1)^j \rho^{j+2} + \beta^{-j} + \rho \beta^{-j-1} + \rho \beta^{-j+1} \right] e^{-iju}. \]

Notice that by letting \(\rho = 0\), we come back to the results for uncorrelated variables.

In the case when \(|\rho| > 1\), we have
\[ \gamma_f(u) := \frac{\sigma}{\sqrt{2\pi}} \left( \frac{\rho \beta + 1}{\rho + e^{-iu}} \right), \quad K = \frac{\sigma}{\sqrt{2\pi}} \left( \frac{\beta + 1}{\beta} \right), \quad \text{ERR}^+_A = \frac{\sigma^2 |\rho + 1/\beta|^2}{\beta^2 \sqrt{1 + 4\alpha^2}}, \]
\[ \text{ERR}_A = \sigma^2 \left( 1 + |\rho|^2 - \frac{2\alpha^2 + 1}{\beta^2 \alpha^4} - \frac{\rho + \overline{\rho}}{\beta^2 \alpha^2} - \frac{|\rho|^2}{\beta \alpha^2} \right). \]

Furthermore,
\[ \hat{g}_*(u) = \frac{\beta + 1}{\alpha^2 (\beta - e^{-iu})(\rho + e^{-iu})} \]
\[ = \left( 1 + \frac{1}{\beta(\rho + \beta)} \right) \frac{1}{\alpha^2 (\beta - e^{-iu})} + \frac{1}{\beta(\rho + \beta)} \frac{1}{\alpha^2 (\rho + e^{-iu})} \]
\[ = \left( 1 + \frac{1}{\beta(\rho + \beta)} \right) \frac{1}{\alpha^2 \beta} \sum_{j=0}^{\infty} \frac{\rho^{-j} e^{-iju}}{\beta^j} + \frac{1}{(\rho + \beta)\alpha^2 \beta \rho} \sum_{j=0}^{\infty} \frac{e^{-iju}}{(-\rho)^j}. \]

4.2. Continuous time.

The Ornstein–Uhlenbeck process. The Ornstein–Uhlenbeck process is a centred Gaussian stationary process with covariance \(K_B(t) = e^{-|t|/2}\) and spectral measure
\[ \mu(du) := \frac{2 \, du}{\pi (4u^2 + 1)}. \tag{4.8} \]

The error in the non-adaptive approximation can easily be calculated by (4.8) and (2.4): \(\text{ERR}_{\text{NA}} = \alpha/(2 + \alpha)\).
The spectral density factorizes as
\[ f(u) = \frac{2}{\pi (4u^2 + 1)} = \gamma_f(u) \overline{\gamma_f(u)}, \]
where
\[ \gamma_f(u) := \sqrt{\frac{2}{\pi}} \frac{1}{1 + 2iu} = \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\tau/2} e^{-i\tau u \tau} d\tau. \]
Hence, by (3.38),
\[ K = \frac{1}{\alpha} \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\tau/2} e^{-\tau/\alpha} d\tau = \frac{\sqrt{2}}{\sqrt{\pi} (2 + \alpha)}, \]
whereas by (3.39) we have
\[ \text{ERR}^+_A = \pi \alpha \frac{2}{\pi (2 + \alpha)^2} = \frac{2\alpha}{(2 + \alpha)^2} \]
and we obtain that
\[ \text{ERR}_A = \text{ERR}_N + \text{ERR}_A^+ = \frac{\alpha}{2 + \alpha} + \frac{2\alpha}{(2 + \alpha)^2} = \frac{4\alpha + \alpha^2}{(2 + \alpha)^2}. \]
The optimal adaptive approximation can easily be found from (3.40):
\[ \hat{g}(u) = \frac{2}{2 + \alpha} \frac{1}{1 + i\alpha u}. \]
Hence, the optimal weight is
\[ g(\tau) = \frac{2}{(2 + \alpha)\alpha} e^{\tau/\alpha} 1_{\{\tau \leq 0\}}. \]
Summarizing, we arrive at the following result.

**Theorem 5.** Suppose that \( \ell(z) = \alpha z \). Then the optimal adaptive approximation of the Ornstein–Uhlenbeck process with covariance function \( K_B(t) = e^{-|t|/2} \) is given by
\[ X(t) = \frac{2}{(2 + \alpha)\alpha} \int_0^\infty B(t-s) e^{-s/\alpha} ds, \]
and the corresponding error is \( \text{ERR}_A = (4\alpha + \alpha^2)/(2 + \alpha)^2 \).

The same results may be obtained formally by \( \delta \)-discretizing the Ornstein–Uhlenbeck process (which gives an autoregressive sequence with \( \rho_\delta = e^{-\delta/2}, \alpha_\delta = \alpha/\delta, \sigma_\delta^2 = 1 - \rho_\delta^2 \)) and letting \( \delta \to 0 \).

**Remark 1.** To conclude, we notice that a similar theory can be developed for processes and sequences with stationary increments and mention in this connection the papers [10] and [11], where similar problems are solved for fractional Brownian motion.
Bibliography

[1] I. Ibragimov, Z. Kabluchko, and M. Lifshits, “Some extensions of linear approximation and prediction problems for stationary processes”, *Stochastic Process. Appl.* **129**:8 (2019), 2758–2782.

[2] Z. Kabluchko and M. Lifshits, “Least energy approximation for processes with stationary increments”, *J. Theoret. Probab.* **30**:1 (2017), 268–296.

[3] A. V. Bulinskii and A. N. Shiryaev, *Theory of stochastic processes*, Fizmatlit, Moscow 2003. (Russian)

[4] R. B. Ash and M. F. Gardner, *Topics in stochastic processes*, Probab. Math. Statist., vol. 27, Academic Press, New York–London 1975.

[5] W. Rudin, *Real and complex analysis*, 3rd ed., McGraw-Hill, New York 1987.

[6] K. Hoffman, *Banach spaces of analytic functions*, Prentice Hall Series in Modern Analysis, Prentice Hall, Englewood Cliffs, NJ 1962; Russian transl., Inostr. Lit., Moscow 1963.

[7] P. J. Brockwell and R. A. Davis, *Time series: theory and methods*, 2nd ed., Springer Ser. Statist., Springer-Verlag, New York 1991.

[8] H. Dym and H. P. McKean, *Gaussian processes, function theory, and the inverse spectral problem*, Reprint of the 1976 ed., Dover Books Math., Dover Publications, Mineola, NY 2008.

[9] A. M. Yaglom, *Correlation theory of stationary random functions with examples from meteorology*, Gidrometeoiizdat, Leningrad 1981. (Russian)

[10] M. L. Kleptsyna, A. Le Breton, and M. Viot, “About the linear-quadratic regulator problem under a fractional Brownian perturbation”, *ESAIM Probab. Stat.* **7** (2003), 161–170.

[11] M. L. Kleptsyna, A. Le Breton, and M. Viot, “On the infinite time horizon linear-quadratic regulator problem under a fractional Brownian perturbation”, *ESAIM Probab. Stat.* **9** (2005), 185–205.

Zakhar A. Kabluchko
Münster University, Münster, Germany
E-mail: zakhar.kabluchko@uni-muenster.de

Mikhail A. Lifshits
St.-Petersburg State University, St.-Petersburg, Russia
E-mail: mikhail@lifshits.org

Received 9/JUL/2018
Translated by THE AUTHORS

25/SEP/2018