

Weierstraß-Institut für Angewandte Analysis und Stochastik

im Forschungsverbund Berlin e.V.

Preprint

ISSN 0946 – 8633

Dependence on the Dimension for Complexity of Approximation of Random Fields

Nora Serdyukova ¹

submitted: 16th September 2008

- ¹ Weierstrass-Institut für Angewandte Analysis und Stochastik,
Mohrenstr. 39, 10117 Berlin, Germany.
E-Mail: serdyuko@wias-berlin.de
URL: <http://www.wias-berlin.de/~serdyuko>

No. 1360
Berlin 2008



2000 *Mathematics Subject Classification.* 41A63, 60G60, 60G15, 41A58.

Key words and phrases. Gaussian processes, random fields, Karhunen-Loève expansion, linear approximation error, information-based complexity, tractability, curse of dimensionality, multivariate linear problems.

This work is supported by the RFBR under grant number 05-01-00911 and RFBR-DFG under grant number 04-01-04000.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Mohrenstraße 39
10117 Berlin
Germany

Fax: + 49 30 2044975
E-Mail: preprint@wias-berlin.de
World Wide Web: <http://www.wias-berlin.de/>

Abstract

We consider the ε -approximation by n -term partial sums of the Karhunen-Loève expansion to d -parametric random fields of tensor product-type in the average case setting. We investigate the behavior, as $d \rightarrow \infty$, of the information complexity $n(\varepsilon, d)$ of approximation with error not exceeding a given level ε . It was recently shown by M. A. Lifshits and E. V. Tulyakova that for this problem one observes the curse of dimensionality (intractability) phenomenon. The aim of this paper is to give the exact asymptotic expression for $n(\varepsilon, d)$.

1 Introduction

Suppose we have a random function $X(t)$, with t in some compact parametric set T , admitting a series representation via random variables ξ_k and the deterministic real functions φ_k , namely,

$$X(t) = \sum_{k=1}^{\infty} \xi_k \varphi_k(t),$$

where the series converges in the mean and a.s. for each $t \in T$. A more precise description will be given later. For any finite set of positive integers $K \subset \mathbb{N}$ let $X_K(t) = \sum_{k \in K} \xi_k \varphi_k(t)$. In many problems one needs to approximate X , for instance under L_2 -norm, with finite rank process X_K . Natural questions arise then: how large should be K that yields a given small approximation error? Given the size of K , which K provides the smallest error?

In this article we address the first of these questions for a specific class of random functions, namely *tensor product-type random fields* with high-dimensional parameter sets. The tensor product-type field is a separable zero-mean random function $X = \{X(t)\}_{t \in T}$, $T \subset \mathbb{R}^d$ with rectangular parametric set T and covariance function $\mathcal{K}^{(d)}$ which can

be decomposed in a product of equal “marginal” covariances depending on different arguments. Namely, let $T = [0, 1]^d$ and

$$\mathcal{K}^{(d)}(s, t) = \prod_{l=1}^d \mathcal{K}_l(s_l, t_l) \quad (1.1)$$

for all $s_l, t_l \in [0, 1]$, $s = (s_1, \dots, s_d)$, $t = (t_1, \dots, t_d)$. Obviously, the integral operator with the kernel (1.1) is the tensor product of the integral operators with the kernels $\mathcal{K}_l(s_l, t_l)$.

Let $\{\lambda_i\}_{i \geq 1}$ be a non-negative sequence satisfying

$$\sum_{i=1}^{\infty} \lambda_i^2 < \infty \quad (1.2)$$

and let $\{\varphi_i\}_{i > 0}$ be an orthonormal system in $L_2[0, 1]$.

Consider a family of tensor product-type random fields

$$\mathbb{X} = \{X^{(d)}(t), t \in [0, 1]^d\}, \quad d = 1, 2, \dots \quad (1.3)$$

According to the multiparametric Karhunen-Loève expansion (see [1] for details), the following equality in distribution holds

$$\begin{aligned} X^{(d)}(t) &= \sum_{\mathbf{k} \in \mathbb{N}^d} \xi_{\mathbf{k}} \prod_{l=1}^d \lambda_{k_l} \prod_{l=1}^d \varphi_{k_l}(t_l) \\ &= \sum_{k_1=1}^{\infty} \cdots \sum_{k_d=1}^{\infty} \xi_{k_1, \dots, k_d} \lambda_{k_1} \cdots \lambda_{k_d} \varphi_{k_1}(t_1) \cdots \varphi_{k_d}(t_d), \end{aligned} \quad (1.4)$$

where the series converges a.s. for every $t = (t_1, \dots, t_d) \in [0, 1]^d$. The collection $\{\xi_{\mathbf{k}}\}$ is an array of non-correlated random variables with zero mean and unit variance and $\lambda_{k_l}^2$ and φ_{k_l} are, respectively, the eigenvalues and eigenfunctions of the family of integral equations

$$\lambda_{k_l}^2 \varphi_{k_l}(t_l) = \int_0^1 \mathcal{K}_l(s_l, t_l) \varphi_{k_l}(s_l) ds_l, \quad t_l \in [0, 1], \quad l = 1, \dots, d,$$

corresponding to the “marginal” covariance operators. Obviously, under assumption (1.2) the sample paths of $X^{(d)}$ belong to $L_2([0, 1]^d)$ almost surely and the covariance operator of $X^{(d)}$ has the system of eigenvalues

$$\lambda_{\mathbf{k}}^2 := \prod_{l=1}^d \lambda_{k_l}^2, \quad \mathbf{k} \in \mathbb{N}^d. \quad (1.5)$$

As it was mentioned in [27], the Karhunen-Loève expansion or the proper orthogonal decomposition of random functions was introduced independently and almost simultaneously by D. D. Kosambi [16], M. Loève [20], K. Karhunen [13] and [14], A. M. Obukhov [21] and V. S. Pougachev [24].

In the following we drop the index d and write $X(t)$ instead of $X^{(d)}(t)$. For any $n > 0$, let X_n be the partial sum of (1.4) corresponding to n maximal eigenvalues. We study the *average case error* of approximation to X by X_n

$$e(X, X_n; d) = (\mathbb{E}\|X - X_n\|_{L_2(T)}^2)^{1/2},$$

as $d \rightarrow \infty$. Since in the following we consider only $L_2(T)$ -norms, we will write $\|\cdot\|$ instead of $\|\cdot\|_{L_2(T)}$. It is well known (see, for example, [6], [17] or [26]) that X_n provides the minimal average quadratic error among all linear approximations to X having rank n .

As we are going to explore a *family* of random functions, it is more natural to investigate *relative* errors, that is to compare the error size with the size of the function itself.

Let

$$\Lambda := \sum_{i=1}^{\infty} \lambda_i^2,$$

then

$$\mathbb{E}\|X\|^2 = \sum_{\mathbf{k} \in \mathbb{N}^d} \lambda_{\mathbf{k}}^2 = \Lambda^d.$$

Then the *average case information complexity* for the normalized error criterion reads as the minimal number of terms in X_n (or, equivalently, of maximal eigenvalues, if they would be ordered) needed to approximate X with the error not exceeding a given level ε :

$$n(\varepsilon, d) := \min\left\{n : \frac{e(X, X_n; d)}{(\mathbb{E}\|X\|^2)^{1/2}} \leq \varepsilon\right\} = \min\{n : \mathbb{E}\|X - X_n\|^2 \leq \varepsilon^2 \Lambda^d\}.$$

The study of $n(\varepsilon, d)$ we are interested in here belongs to the class of problems dealing with the dependence of the information complexity for linear multivariate problems on the dimension, see the works of H. Woźniakowski ([30], [31], [32], [33]) and the references therein.

It was suggested in [19] to use an auxiliary probabilistic construction for studying the properties of deterministic array of eigenvalues (1.5). We follow this approach.

Consider a sequence of independent identically distributed random variables $\{U_l\}$, $l = 1, 2, \dots$ with the common distribution given by

$$\mathbb{P}(U_l = -\log \lambda_i) = \frac{\lambda_i^2}{\Lambda}, \quad i = 1, 2, \dots \quad (1.6)$$

Under the assumption

$$\sum_{i=1}^{\infty} |\log \lambda_i|^3 \lambda_i^2 < \infty, \quad (1.7)$$

the condition $\mathbb{E}|U_l|^3 < \infty$ is obviously satisfied.

Let M and σ^2 denote, respectively, the mean and the variance of U_l . Clearly,

$$\begin{aligned} M &= -\sum_{i=1}^{\infty} \log \lambda_i \frac{\lambda_i^2}{\Lambda}, \\ \sigma^2 &= \sum_{i=1}^{\infty} |\log \lambda_i|^2 \frac{\lambda_i^2}{\Lambda} - M^2. \end{aligned}$$

Then the third central moment of U_l is given by

$$\alpha^3 := \mathbb{E}(U_l - M)^3 = -\sum_{i=1}^{\infty} (\log \lambda_i)^3 \frac{\lambda_i^2}{\Lambda} - 3M\sigma^2 - M^3.$$

If (1.7) is verified, we have $|M| < \infty$, $0 \leq \sigma^2 < \infty$ and $|\alpha| < \infty$.

In the sequel the explosion coefficient

$$\mathcal{E} := \Lambda e^{2M} \quad (1.8)$$

will play a significant role, because its contribution into the ‘‘curse of dimensionality’’ is the largest. It was shown in [19] that by concavity of the logarithmic function $\mathcal{E} > 1$, except for the totally degenerate case when the number of strictly positive eigenvalues is zero or one. In other words, $\mathcal{E} = 1$ iff $\sigma = 0$. Henceforth we will exclude this degenerate case.

The following result was obtained in [19], Theorem 3.2.

Theorem 1.1 *Assume that the sequence $\{\lambda_i\}_{i \geq 1}$ satisfies the condition*

$$\sum_{i=1}^{\infty} |\log \lambda_i|^2 \lambda_i^2 < \infty.$$

Then for every $\varepsilon \in (0, 1)$ we have

$$\lim_{d \rightarrow \infty} \frac{\log n(\varepsilon, d) - d \log \mathcal{E}}{\sqrt{d}} = 2q,$$

where the quantile $q = q(\varepsilon)$ is chosen from the equation

$$1 - \Phi\left(\frac{q}{\sigma}\right) = \varepsilon^2. \quad (1.9)$$

The authors of [19] conjectured that under further assumptions on the sequence $\{\lambda_i\}$ one can prove that

$$n(\varepsilon, d) \approx \frac{C(\varepsilon)\mathcal{E}^d e^{2q\sqrt{d}}}{\sqrt{d}}, \quad d \rightarrow \infty.$$

We are going to confirm this conjecture.

2 Main result

It turns out that two different cases depending on the nature of the distribution of U_l should be distinguished. The proof and the final result depend on whether this distribution is a *lattice* one or not.

Recall that one calls a discrete distribution of a random variable U a lattice distribution, if there exist numbers a and $h > 0$ such that every possible value of U can be represented in the form $a + \nu h$, where ν is an integer. The number h is called a span of the distribution. In the following, when studying the lattice case, we assume that h is a maximal span of the distribution, i.e. one cannot represent all possible values of U_l in the form $b + \nu h_1$ for some b and $h_1 > h$.

Definition (1.6) yields that the variables U_l have a common lattice distribution iff $\lambda_i = C e^{-n_i h}$ for some positive C , h and $n_i \in \mathbb{N}$. We call this situation the *lattice case* and will assume that h is chosen to be the largest possible. Otherwise we say that the *non-lattice case* takes place.

By $f(d) = o(g(d))$ we mean that $\lim_{d \rightarrow \infty} \frac{f(d)}{g(d)} = 0$. In particular, $f(d) = g(d)(1 + o(1))$ means that $\lim_{d \rightarrow \infty} \frac{f(d)}{g(d)} = 1$.

Theorem 2.1 *Let the sequence $\{\lambda_i\}_{i \geq 1}$ satisfy (1.7).*

Then for every $\varepsilon \in (0, 1)$ it holds

$$n(\varepsilon, d) = K \phi\left(\frac{q}{\sigma}\right) \mathcal{E}^d e^{2q\sqrt{d}} d^{-1/2} (1 + o(1)), \quad d \rightarrow \infty,$$

where

$$\begin{aligned} \phi(x) &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \\ K &= \begin{cases} \frac{h}{\sigma(1-e^{-2h})} & \text{in the lattice case,} \\ \frac{1}{2\sigma} & \text{otherwise,} \end{cases} \end{aligned}$$

and the quantile $q = q(\varepsilon)$ is defined in (1.9).

Remarks:

- One can see that the complexity of approximation increases exponentially as $d \rightarrow \infty$. This phenomenon is referred to as the *curse of dimensionality* or *intractability*, see e.g. [26] and [31]. The notion of “curse of dimensionality” dates back at least to Bellman [4].
- By l’Hospital’s rule

$$\lim_{h \rightarrow 0} \frac{h}{\sigma(1-e^{-2h})} = \frac{1}{2\sigma}.$$

Proof:

Let $\zeta = \zeta(\varepsilon, d)$ be the maximal positive number such that the sum of eigenvalues satisfies

$$\sum_{\mathbf{k} \in \mathbb{N}^d: \lambda_{\mathbf{k}} < \zeta} \lambda_{\mathbf{k}}^2 \leq \varepsilon^2 \Lambda^d.$$

Define a lattice set in \mathbb{N}^d

$$A = A(\varepsilon, d) := \{\mathbf{k} \in \mathbb{N}^d : \lambda_{\mathbf{k}} \geq \zeta\} = \left\{ \mathbf{k} \in \mathbb{N}^d : \prod_{l=1}^d \lambda_{k_l} \geq \zeta \right\}.$$

Since for any $\mathbf{k} \in A$ it holds that $\lambda_{\mathbf{k}} > 0$, one can write

$$\begin{aligned}
n(\varepsilon, d) &= \text{card}(A) = \sum_{\mathbf{k} \in A} \frac{\lambda_{\mathbf{k}}^2}{\lambda_{\mathbf{k}}^2} \\
&= \sum_{\mathbf{k} \in \mathbb{N}^d: -\sum \log \lambda_{k_l} \leq -\log \zeta} \Lambda^d \exp \left\{ -2 \sum_{l=1}^d \log \lambda_{k_l} \right\} \prod_{l=1}^d \mathbb{P}(U_l = -\log \lambda_{k_l}) \\
&= \Lambda^d \mathbb{E} \exp \left\{ 2 \sum_{l=1}^d U_l \right\} \mathbb{I}_{\{\sum_{l=1}^d U_l \leq -\log \zeta\}}.
\end{aligned}$$

For centered and normalized sums

$$Z_d = \frac{\sum_{l=1}^d U_l - dM}{\sigma\sqrt{d}}$$

we have

$$\left\{ \sum_{l=1}^d U_l \leq -\log \zeta \right\} = \{Z_d \leq \theta\},$$

where

$$\theta = \theta(\varepsilon, d) = -\frac{\log \zeta + dM}{\sigma\sqrt{d}}. \quad (2.10)$$

We show now that θ has a useful probabilistic meaning in terms of $\{U_l\}$ and of their sums. Applying Lemma 3.1 from [19] we have for any $d \in \mathbb{N}$ and $z \in \mathbb{R}^1$

$$\begin{aligned}
\sum_{\mathbf{k} \in \mathbb{N}^d: \lambda_{\mathbf{k}} < z} \lambda_{\mathbf{k}}^2 &= \Lambda^d \mathbb{P} \left(\sum_{l=1}^d U_l > -\log z \right) \\
&= \Lambda^d \mathbb{P} \left(Z_d > -\frac{\log z + dM}{\sigma\sqrt{d}} \right) = \Lambda^d \mathbb{P}(Z_d > \theta_z),
\end{aligned}$$

where

$$\theta_z = -\frac{\log z + dM}{\sigma\sqrt{d}}.$$

Fix $\varepsilon \in (0, 1)$. Observe that

$$\sum_{\mathbf{k} \in \mathbb{N}^d: \lambda_{\mathbf{k}} < z} \lambda_{\mathbf{k}}^2 \leq \varepsilon^2 \Lambda^d$$

iff

$$\mathbb{P}(Z_d > \theta_z) \leq \varepsilon^2.$$

Therefore, $\theta = \theta(\varepsilon, d)$ defined by (2.10) is the $(1 - \varepsilon^2)$ -quantile of the distribution of Z_d , namely,

$$\theta(\varepsilon, d) = \min\{\theta : \mathbb{P}(Z_d > \theta) \leq \varepsilon^2\} = \min\{\theta : \mathbb{P}(Z_d \leq \theta) > 1 - \varepsilon^2\}.$$

Let $q = q(\varepsilon)$ be the quantile of the normal distribution function chosen from the equation (1.9). Then in view of the Central Limit Theorem

$$\theta(\varepsilon, d) \rightarrow \frac{q(\varepsilon)}{\sigma}, \quad d \rightarrow \infty, \quad (2.11)$$

for any fixed $\varepsilon \in (0, 1)$.

Now let us return to the information complexity. We obtain

$$\begin{aligned} n(\varepsilon, d) &= \mathcal{E}^d \mathbb{E} \exp\{2\sigma\sqrt{d}Z_d\} \mathbb{I}_{\{Z_d \leq \theta\}} \\ &= \mathcal{E}^d \exp\{2\sigma\sqrt{d}\theta\} \int_{-\infty}^{\theta} \exp\{2\sigma\sqrt{d}(z - \theta)\} dF_d(z), \end{aligned}$$

where $F_d(z) = \mathbb{P}(Z_d < z)$ and \mathcal{E} is defined in (1.8).

Denote

$$\Psi_d(z) := \exp\{2\sigma\sqrt{d}(z - \theta)\}$$

and integrate by parts the integral

$$\int_{-\infty}^{\theta} \Psi_d(z) d[F_d(z) - F_d(\theta)] = \int_{-\infty}^{\theta} [-F_d(z) + F_d(\theta)] d\Psi_d(z).$$

From now on we have to distinguish the lattice and non-lattice cases.

Non-lattice case

In the following part of the proof we will assume that the distribution of $\{U_l\}$ is not a lattice one. This is true in the most interesting cases, such as the Brownian sheet (the Wiener-Chentsov random field), the completely tugged Brownian sheet (the Brownian pillow), the d-variate Hoeffding, Blum, Kiefer and Rosenblatt process (see Appendix for details).

In view of (1.7) we are able to apply the Cramér-Esseen Theorem (cf. Theorem 2 §42 in [10], Theorem 5.21 §5.7 of Chapter V in [23] or

Theorem 4 §3 of Chapter VI in [22]). It leads to

$$\begin{aligned}
\int_{-\infty}^{\theta} [-F_d(z) + F_d(\theta)] d\Psi_d(z) &= \int_{-\infty}^{\theta} [-\Phi(z) + \Phi(\theta)] d\Psi_d(z) \\
+ \frac{\alpha^3}{6\sigma^3\sqrt{2\pi d}} \int_{-\infty}^{\theta} [(z^2 - 1)e^{-z^2/2} - ((\theta^2 - 1)e^{-\theta^2/2})] d\Psi_d(z) &+ o\left(\frac{1}{\sqrt{d}}\right) \\
&= I_1 + I_2 - I_3 - I_4 + o\left(\frac{1}{\sqrt{d}}\right), \tag{2.12}
\end{aligned}$$

where

$$\begin{aligned}
I_1 &= \int_{-\infty}^{\theta} [-\Phi(z) + \Phi(\theta)] d\Psi_d(z), \\
I_2 &= \frac{\alpha^3}{6\sigma^3\sqrt{2\pi d}} \int_{-\infty}^{\theta} z^2 e^{-z^2/2} d\Psi_d(z), \\
I_3 &= \frac{\alpha^3}{6\sigma^3\sqrt{2\pi d}} \int_{-\infty}^{\theta} e^{-z^2/2} d\Psi_d(z), \\
I_4 &= \frac{\alpha^3}{6\sigma^3\sqrt{2\pi d}} (\theta^2 - 1) e^{-\theta^2/2} = \frac{\alpha^3}{6\sigma^3\sqrt{2\pi d}} \left(\left(\frac{q}{\sigma}\right)^2 - 1 \right) \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)).
\end{aligned}$$

The last equivalence is provided by (2.11).

Since $d\Psi_d(z) = 2\sigma\sqrt{d}\Psi_d(z)dz$, the integral I_2 is given, after a change of variable, by

$$I_2 = I_2(d, \theta) = \frac{\alpha^3}{3\sigma^2\sqrt{2\pi d}} \int_0^{\infty} \left(\theta - \frac{y}{\sqrt{d}}\right)^2 \exp\left\{-\frac{1}{2}\left(\theta - \frac{y}{\sqrt{d}}\right)^2\right\} \exp\{-2\sigma y\} dy$$

where $y = -\sqrt{d}(z - \theta)$.

For any $d = 1, 2, \dots$

$$0 \leq \left(\theta - \frac{y}{\sqrt{d}}\right)^2 \exp\left\{-\frac{1}{2}\left(\theta - \frac{y}{\sqrt{d}}\right)^2\right\} \leq (|\theta| + y)^2.$$

This estimate gives us the majorant required in the Lebesgue's dominated convergence theorem. Using (2.11) and passing to the limit in the integral we obtain, as $d \rightarrow \infty$,

$$I_2(d, \theta) = \frac{\alpha^3}{6\sigma^3\sqrt{2\pi d}} \left(\frac{q}{\sigma}\right)^2 \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)).$$

Similarly,

$$I_3(d, \theta) = \frac{\alpha^3}{6\sigma^3\sqrt{2\pi d}} \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)).$$

Thus we obtain that $\sqrt{d}I_4 = \sqrt{d}(I_2 - I_3)(1 + o(1))$, hence, $I_2 - I_3 - I_4 = o\left(\frac{1}{\sqrt{d}}\right)$.

Consider the main integral I_1 .

$$\begin{aligned} I_1 &= I_1(d, \theta) = \int_{-\infty}^{\theta} [-\Phi(z) + \Phi(\theta)] d\Psi_d(z) \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\theta} \exp\{2\sigma\sqrt{d}(z - \theta)\} \exp\{-z^2/2\} dz \\ &= \frac{1}{\sqrt{2\pi d}} \int_0^{\infty} \exp\left\{-\frac{1}{2}\left(\theta - \frac{y}{\sqrt{d}}\right)^2\right\} \exp\{-2\sigma y\} dy \\ &= \frac{1}{2\sigma\sqrt{2\pi d}} \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)), \quad d \rightarrow \infty. \end{aligned} \quad (2.13)$$

Then

$$n(\varepsilon, d) = \frac{\mathcal{E}^d \exp\{2q\sqrt{d}\}}{2\sigma\sqrt{d}} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)),$$

as asserted.

Lattice case

Now we will proceed under the assumption that the random variables U_l have a lattice distribution. Let possible values of the random variable U_l be

$$\tilde{a} + \nu h, \quad \nu = 0, \pm 1, \pm 2, \dots$$

where $\tilde{a} = M + a$ is a shift, and h is the maximal span of the distribution. Therefore, all possible values of Z_d have the form

$$\frac{da + \nu h}{\sigma\sqrt{d}}, \quad \nu = 0, \pm 1, \pm 2, \dots$$

Introduce the function

$$S(x) = [x] - x + \frac{1}{2},$$

where $[x]$ denotes, as usual, the integer part of x , and consider

$$S_d(x) = \frac{h}{\sigma} S\left(\frac{x\sigma\sqrt{d} - da}{h}\right).$$

Let $F_d(z)$ be as above. Then under the assumption (1.7) Esseen's result (see Theorem 1 § 43 in [10]) yields

$$F_d(z) - \Phi(z) = \frac{e^{-z^2/2}}{\sqrt{2\pi}} \left(\frac{S_d(z)}{\sqrt{d}} - \frac{\alpha^3(z^2 - 1)}{6\sigma^3\sqrt{d}} \right) + o\left(\frac{1}{\sqrt{d}}\right)$$

uniformly in z .

Comparing with (2.12), we observe that one only needs to evaluate the additional term

$$\begin{aligned} J &= \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{\theta} [-S_d(z)e^{-z^2/2} + S_d(\theta)e^{-\theta^2/2}] d\Psi_d(z) \\ &= \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{\theta} \Psi_d(z) d\left(S_d(z)e^{-z^2/2}\right) = J_1 - J_2 + J_3, \end{aligned}$$

where

$$\begin{aligned} J_1 &= \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{\theta} \Psi_d(z) S'_d(z) e^{-z^2/2} dz, \\ J_2 &= \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{\theta} \Psi_d(z) S_d(z) z e^{-z^2/2} dz, \end{aligned}$$

and J_3 is a “discrete part”, which is defined in the following way. Notice that $S(x)$ is a periodic function with period one, therefore $S_d(x)$ possesses the period $h/\sigma\sqrt{d}$ and has jumps at points $\{\frac{kh+da}{\sigma\sqrt{d}}, k \in \mathbb{Z}\}$. If the point θ belongs to this lattice then there exists an integer k' such that $\theta = \frac{k'h+da}{\sigma\sqrt{d}}$. Hence, one can integrate the discontinuous part of the integral J with respect to the measure $\frac{h}{\sigma} \delta_{\frac{kh+da}{\sigma\sqrt{d}}}$ and obtain

$$J_3 = \frac{1}{\sqrt{2\pi d}} \frac{h}{\sigma} \sum_{k=-\infty}^{k'} \Psi_d\left(\frac{kh+da}{\sigma\sqrt{d}}\right) \exp\left\{-\frac{1}{2} \left(\frac{kh+da}{\sigma\sqrt{d}}\right)^2\right\}.$$

We start with the estimation of J_1 . At the points where the derivative $S'_d(z)$ makes sense, one can easily calculate that $S'_d(z) = \frac{h}{\sigma} S\left(\frac{z\sigma\sqrt{d}-da}{h}\right) =$

$-\sqrt{d}$, therefore, similarly to the non-lattice case, by the Lebesgue's dominated convergence theorem we have

$$\begin{aligned}
J_1 &= \frac{-\sqrt{d}}{\sqrt{2\pi d}} \int_{-\infty}^{\theta} \exp\{2\sigma\sqrt{d}(z-\theta)\} \exp\{-z^2/2\} dz \\
&= \frac{-1}{\sqrt{2\pi d}} \int_0^{\infty} \exp\{-\frac{1}{2}(\theta - \frac{y}{\sqrt{d}})^2\} \exp\{-2\sigma y\} dy \\
&= \frac{-1}{2\sigma\sqrt{2\pi d}} \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)) , \quad d \rightarrow \infty, \quad (2.14)
\end{aligned}$$

and it yields $\sqrt{d}J_1 = -\sqrt{d}I_1 (1 + o(1))$.

As for the integral J_2 , this one, as d is large enough, becomes negligible. Indeed,

$$\begin{aligned}
J_2 &= \frac{1}{\sqrt{2\pi d}} \int_{-\infty}^{\theta} \exp\{2\sigma\sqrt{d}(z-\theta)\} S_d(z) z \exp\{-z^2/2\} dz \\
&= \frac{1}{\sqrt{2\pi d}} \frac{1}{\sqrt{d}} \int_0^{\infty} \exp\{-\frac{1}{2}(\theta - \frac{y}{\sqrt{d}})^2\} (\theta - \frac{y}{\sqrt{d}}) S_d(\theta - \frac{y}{\sqrt{d}}) \exp\{-2\sigma y\} dy \\
&\leq \frac{3h}{2\sigma d\sqrt{2\pi}} \int_0^{\infty} \exp\{-\frac{1}{2}(\theta - \frac{y}{\sqrt{d}})^2\} (\theta - \frac{y}{\sqrt{d}}) \exp\{-2\sigma y\} dy \\
&= \frac{3h}{4\sigma^2 d\sqrt{2\pi}} \left(\frac{q}{\sigma}\right)^2 \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)) , \quad d \rightarrow \infty.
\end{aligned}$$

And, of course, $J_2 = o\left(\frac{1}{\sqrt{d}}\right)$.

Now we consider the most essential summand

$$\begin{aligned}
J_3 &= \frac{1}{\sqrt{2\pi d}} \frac{h}{\sigma} \sum_{k=-\infty}^{k'} \exp\{2\sigma\sqrt{d}\left(\frac{kh+da}{\sigma\sqrt{d}} - \theta\right)\} \exp\left\{-\frac{1}{2}\left(\frac{kh+da}{\sigma\sqrt{d}}\right)^2\right\} \\
&= \frac{1}{\sqrt{2\pi d}} \frac{h}{\sigma} \sum_{k=-\infty}^{k'} \exp\{2h(k-k')\} \exp\left\{-\frac{1}{2}\left(\frac{kh+da}{\sigma\sqrt{d}}\right)^2\right\} \\
&= \frac{1}{\sqrt{2\pi d}} \frac{h}{\sigma} \sum_{l=0}^{\infty} \exp\{-2hl\} \exp\left\{-\frac{1}{2}\left(\frac{(k'-l)h+da}{\sigma\sqrt{d}}\right)^2\right\} \\
&= \frac{1}{\sqrt{2\pi d}} \frac{h}{\sigma} \sum_{l=0}^{\infty} \exp\{-2hl\} \exp\left\{-\frac{1}{2}\left(\theta - \frac{lh}{\sigma\sqrt{d}}\right)^2\right\} \\
&= \frac{1}{\sigma\sqrt{d}} \frac{h}{(1-e^{-2h})} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{q^2}{2\sigma^2}\right\} (1 + o(1)) , \quad d \rightarrow \infty. \quad (2.15)
\end{aligned}$$

We obtain

$$\sqrt{d}J_3 = \sqrt{d} \frac{2h}{(1 - e^{-2h})} I_1 (1 + o(1)).$$

Putting together (2.13), (2.14) and (2.15), we get

$$n(\varepsilon, d) = \frac{\mathcal{E}^d e^{2q\sqrt{d}}}{\sigma\sqrt{d}} \frac{h}{(1 - e^{-2h})} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{q^2}{2\sigma^2} \right\} (1 + o(1)), \quad d \rightarrow \infty.$$

□

3 Appendix. Examples of tensor product-type random fields

This section contains some examples of random fields to which the above general result can be applied.

3.1 Wiener-Chentsov random field

The Wiener-Chentsov field (the Brownian sheet) (see [18]) is a zero-mean Gaussian random function $W^{(d)}$ with the covariance function equal to a product of the covariance functions corresponding to the Wiener process W :

$$\mathcal{K}_{W^{(d)}}(s, t) = \prod_{l=1}^d \min\{s_l, t_l\}, \quad s = (s_1, \dots, s_d), \quad t = (t_1, \dots, t_d) \in T.$$

Therefore the marginal eigenvalues have the following form:

$$\lambda_{W;i}^2 = (\pi(i - 1/2))^{-2}, \quad i = 1, 2, \dots .$$

3.2 Completely tucked Brownian sheet

The completely tucked Brownian sheet (the Brownian pillow) is a zero-mean Gaussian random function $B^{(2)}$ with the covariance function,

equal to a product of the covariance functions corresponding to the standard Brownian bridge $B(t) = W(t) - tW(1)$, namely

$$\mathcal{K}_{B^{(2)}}(s, t) = \prod_{l=1}^2 (\min\{s_l, t_l\} - s_l t_l), \quad s, t \in [0, 1]^2.$$

Correspondingly, the marginal eigenvalues (see. [2]) are equal to

$$\lambda_{B,i}^2 = (\pi i)^{-2}, \quad i = 1, 2, \dots$$

In the literature different terms are in use for this random field. In [28] the term “completely tucked Brownian sheet” is used; in [7] – “tied-down Kiefer process”; in [15] this field is called “the Brownian pillow”.

The notion of “completely tucked Brownian sheet” and its generalization for the case $d > 2$, was introduced by J. R. Blum, J. Kiefer and M. Rosenblatt [5] as the limiting distribution for a functional of empirical process occurring in nonparametric testing of independency, so-called “the independence empirical process” (see [28]). Therefore the d -parametric generalization of the completely tucked Brownian sheet is often referred to as “the d -variate Hoeffding, Blum, Kiefer and Rosenblatt process” (see, for example, [15]). The mention of Hoeffding’s name in the term is motivated by the fact that the test studied in [5] was equivalent to the one suggested earlier by W. Hoeffding in [12]. But the limiting distribution, the covariance function, the eigenvalues and the eigenfunctions of the corresponding integral equation were obtained in [5]. Higher-dimensional generalizations were later treated in [9] and in [8].

3.3 Centered Gaussian processes

In some statistical problems it is convenient to use centered empirical processes and corresponding limiting Gaussian processes.

For any Gaussian process $X = \{X(t)\}$, $t \in [0, 1]$ we define the centered process

$$\mathring{X}(t) := X(t) - \int_0^1 X(u) du.$$

The centered Brownian bridge \mathring{B} , also referred to in the literature as the Watson process, was introduced in [29] for nonparametric goodness-of-fit testing on a circle. G. S. Watson showed that the covariance function

is given by

$$\mathcal{K}_{\mathring{B}}(s, t) = \min\{s, t\} - st + \frac{1}{2}(s^2 + t^2 - s - t) + \frac{1}{12}, \quad s, t \in [0, 1],$$

and the covariance operator with this kernel has a double spectrum, i.e.

$$\lambda_{\mathring{B};2i}^2 = \lambda_{\mathring{B};(2i-1)}^2 = (2\pi i)^{-2}, \quad i = 1, 2, \dots$$

The covariance function of the centered Wiener process \mathring{W} has the form

$$\mathcal{K}_{\mathring{W}}(s, t) = \min\{s, t\} + \frac{1}{2}(s^2 + t^2) - s - t + \frac{1}{3}, \quad s, t \in [0, 1],$$

and the corresponding eigenvalues coincide with those of the standard Brownian bridge, i.e.

$$\lambda_{\mathring{W};i}^2 = \lambda_{B;i}^2 = (\pi i)^{-2}, \quad i = 1, 2, \dots,$$

that is in accordance with the well-known equality in distribution for L_2 -norms of the Brownian bridge and centered Wiener process, see [3].

Centered integrated Brownian bridge

$$\check{B}(t) = \bar{B}(t) - \int_0^1 \bar{B}(u) du,$$

where

$$\bar{B}(t) = \int_0^t B(u) du, \quad t \in [0, 1]$$

was considered in a framework of goodness-of-fit testing and small deviation probabilities under L_2 -norm in [11] and in [3], where its covariance function

$$\mathcal{K}_{\check{B}}(s, t) = \frac{st \min\{s, t\}}{2} - \frac{\min\{s, t\}^3}{6} - \frac{(st)^2}{4} - \frac{s^2 + t^2}{6} - \frac{s^4 + t^4}{24} + \frac{s^3 + t^3}{6} + \frac{1}{45},$$

$s, t \in [0, 1]$ and eigenvalues

$$\lambda_{\check{B};i}^2 = (\pi i)^{-4}, \quad i = 1, 2, \dots$$

were obtained.

3.4 Multivariate extensions of the Anderson-Darling process

The tensor product of Anderson-Darling processes $A^{(d)}(t)$, $t \in [0, 1]^d$ is a zero-mean Gaussian random function $A^{(d)}(t)$, $t \in [0, 1]^d$ with the covariance function

$$\mathcal{K}_{A^{(d)}}(s, t) = \prod_{l=1}^d \frac{\min\{s_l, t_l\} - s_l t_l}{\sqrt{s_l(1-s_l)}\sqrt{t_l(1-t_l)}}, \quad s_l, t_l \in [0, 1].$$

The eigenvalues of the corresponding covariance operator are given by

$$\lambda_{\mathbf{k}}^2 = \prod_{l=1}^d \frac{1}{k_l(k_l + 1)}, \quad \mathbf{k} = (k_1, \dots, k_d) \in \mathbb{N}^d.$$

In the one-dimensional case the Anderson-Darling process coincides in distribution with $\frac{B(t)}{\sqrt{t(1-t)}}$, $t \in [0, 1]$ and was introduced in [2] in the context of goodness-of-fit testing. T. Anderson and D. Darling obtained its covariance function and the exact spectrum.

In [25] another multivariate extension of Anderson-Darling process, defined as a zero-mean Gaussian process with the covariance function

$$\mathcal{K}_A^\mu(s, t) = \left(\frac{\min\{s, t\} - st}{\sqrt{s(1-s)}\sqrt{t(1-t)}} \right)^\mu, \quad s, t \in [0, 1], \mu > 0,$$

is given.

The eigenvalues of its covariance operator are of the form

$$\lambda_{\mu, j}^2 = \frac{\mu}{(\mu + j - 1)(\mu + j)}, \quad j = 1, 2, \dots$$

When parameter μ is positive integer, the random field, defined in such a way, (more precisely, the square of its L_2 -norm) is the limiting distribution for Cramér-von Mises type statistics.

Acknowledgements

This paper was partially written while the author was visiting the Institut für Mathematische Stochastik, Georg-August-Universität, Göttingen. Special thanks are due to Professor M.A. Lifshits for formulation of the problem and constant encouragement and to Professor M. Denker for his support and for providing excellent working conditions.

References

- [1] *Adler R. J.* An Introduction to Continuity, Extrema and Related Topics for General Gaussian Processes. IMS Lect. notes – Monograph Series, 12. Institute of Mathematical Statistics, Hayward, California, 1990, 160 p.
- [2] *Anderson T. W., Darling D. A.* Asymptotic theory of certain “goodness of fit” criteria based on stochastic processes. *Ann. Math. Statist.*, 1952, v. 23 p. 193–212.
- [3] *Beghin L., Nikitin Ya., Orsingher E.* Exact small ball constants for some Gaussian processes under the L^2 -norm., *J. Math. Sciences*, 2005, v. 128 p. 2493–2502.
- [4] *Bellman R.* Adaptive Control Processes: a Guided Tour. Princeton University, Princeton, 1961, 255 p.
- [5] *Blum J. R., Kiefer J., Rosenblatt M.* Distribution free tests of independence based on the sample distribution function. *Ann. Math. Statist.*, 1961, v. 32, no. 2, p. 485–498.
- [6] *Buslaev A. P., Seleznev O. V.* On certain extremal problems in the theory of approximation of random processes. *East J. Approx.*, 1999, v. 5 , no. 4, p. 467–481.
- [7] *Csörgő M., Horváth L.* Limit Theorems in Change-point Analysis., Wiley Series in Probability and Statistics, NY,Wiley, 1997, xvi+414 p.
- [8] *Deheuvels P.* An asymptotic decomposition for multivariate distribution-free tests of independence. *J. of Multivar. Analysis*, 1981, v. 11 , p. 102–113.
- [9] *Dugue D.* Sur des tests d’indépendance indépendants de la loi. *Comptes Rendus des Séances de l’Académie des Sciences Series A*, 1975, v. 281, p. 1103-1104.
- [10] *Gnedenko, B. V., Kolmogorov, A. N.* Limit Distributions for Sums of Independent Random Variables. Addison-Wesley, Cambridge, 1954, ix+264 p (*in Russian*: GTTI, Moscow-Leningrad, 1949).
- [11] *Henze N., Nikitin Ya. Yu.* A new approach to goodness-of-fit testing based on the integrated empirical process. *J. Nonparam. Statist.*, 2000, v. 12, p.391–416.
- [12] *Hoeffding W.* A non-parametric test of independence. *Ann. Math. Statist.*, 1948, v. 19 , p. 546–557.

- [13] *Karhunen K.* Zur Spektraltheorie stochastischer Prozesse. Ann. Acad. Sci. Fennicae, Ser. A I, 1946, no. 34 p. 1–7.
- [14] *Karhunen K.* Über lineare Methoden in der Wahrscheinlichkeitsrechnung. Ann. Acad. Sci. Fennicae, Ser. A I, 1947, no. 37 p. 3–79.
- [15] *Koning A. J., Protasov V.* Tail behaviour of Gaussian processes with applications to the Brownian pillow. J. of Multivar. Analysis, 2003, v. 87, no. 2, p. 370–397.
- [16] *Kosambi D. D.* Statistics in functional space. J. Indian Math. Soc., 1943, v. 7, p. 76–88.
- [17] *Kühn Th., Linde W.* Optimal series representation of fractional Brownian sheets. Bernoulli, 2002, v. 8, no. 5, p. 669–696.
- [18] *Lifshits M. A.* Gaussian Random Functions. Kluwer, Dordrecht, 1995, xi+333 p.
- [19] *Lifshits M. A., Tulyakova E. V.* Curse of dimensionality in approximation of random fields. Probab. Math. Stat., 2006, v. 26, no. 1, p. 83–98.
- [20] *Loève M.* Fonctions aléatoires de second ordre. Revue Scientifique, 1946, t. 84, v. 4, p. 195–206.
- [21] *Obukhov A. M.* Statistical description of continuous fields. Tr. geophys. Inst. Akad. Nauk SSSR, 1954, no. 24 (151), p. 3–42, (*in Russian*).
- [22] *Petrov, V. V.* Sums of Independent Random Variables. Nauka, Moscow, 1972, 416 p., (*in Russian*).
- [23] *Petrov, V. V.* Limit Theorems of Probability Theory: Sequences of Independent Random Variables. Oxford Studies in Probability, 4. Clarendon press, Oxford, 1995, ix+292 p. (*in Russian*: Nauka, Moscow, 1987, 320 p.).
- [24] *Pougachev V. S.* General theory of the correlations of random functions. Izv. Akad. Nauk SSSR, Ser. Math., 1953, v. 17, no. 5., p. 401–420.
- [25] *Pycke J.-R.* Multivariate extensions of the Anderson-Darling process. Stat. Probab. Lett., 2003, v. 63, no. 4, p. 387–399.
- [26] *Ritter K.* Average-case Analysis of Numerical Problems. Lecture Notes in Mathematics, 2000, v. 1733, x+254 p.
- [27] *Sabelfeld K.* Expansion of random boundary excitations for elliptic PDEs. Monte Carlo Methods Appl., 2007, v. 13, no. 5–6, p. 405–453.

- [28] *van der Vaart A. W., Wellner J. A.* Weak Convergence and Empirical Processes with Applications to Statistics. Springer Series in Statistics. NY, Springer, 1996, xvi+508 p.
- [29] *Watson G. S.* Goodness-of-fit tests on a circle. *Biometrika*, 1961, v. 48 , p. 109–114.
- [30] *Woźniakowski H.* Average case complexity of linear multivariate problems. Part 1: Theory. Part 2: Applications. *J. Complexity*, 1992, v. 8, p. 337–372, p. 373–392.
- [31] *Woźniakowski H.* Tractability and strong tractability of linear multivariate problems. *J. Complexity*, 1994, v. 10, p. 96–128.
- [32] *Woźniakowski H.* Tractability and strong tractability of multivariate tensor product problems. *J. of Computing and Information*, 1994, v. 4, p. 1–19.
- [33] *Woźniakowski H.* Tractability of multivariate problems for weighted spaces of functions. *Approximation and Probability, Banach Center Publ.*, v. 72, 2006, p. 407–427.