PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 26, Fasc. 1 (2006), pp. 97–112

CURSE OF DIMENSIONALITY IN APPROXIMATION OF RANDOM FIELDS

BY

M. A. LIFSHITS AND E. V. TULYAKOVA (ST. PETERSBURG)

Abstract. Consider a random field of tensor product-type $X(t)$, $t \in [0, 1]^d$, given by

$$X(t) = \sum_{k \in N^d} \prod_{l=1}^{d} \lambda(k_l)\, \xi_k \prod_{l=1}^{d} \varphi_{k_l}(t_l),$$

where $(\lambda(i))_{i>0} \in l_2$, $(\varphi_i)_{i>0}$ is an orthonormal system in $L_2[0, 1]$ and $(\xi_k)_{k\in N^d}$ are non-correlated random variables with zero mean and unit variance. We investigate the quality of approximation (both in the average and in the probabilistic sense) to X by the n-term partial sums X_n minimizing the quadratic error $E\|X-X_n\|^2$. In the first part of the paper we consider the case of fixed dimension d. In the second part, following the suggestion of H. Woźniakowski, we consider the same problem for $d \to \infty$. We show that, for any fixed level of relative error, approximation complexity increases exponentially and we find the explosion coefficient. We also show that the behavior of the probabilistic and average complexity is essentially the same in the large domain of parameters.

AMS Mathematics Subject Classification: 41A63, 60G15.

Key words and phrases: Random fields, Gaussian processes, fractional Brownian sheet, linear approximation error, complexity.

1. INTRODUCTION

Let $X(t) = \sum_{k=1}^{\infty} \xi_k \varphi_k(t)$, $t \in T$, be a random function represented via random variables ξ_k and the deterministic real functions φ_k. Let $X_n(t) = \sum_{k=1}^{n} \xi_k \varphi_k(t)$ be the approximation to X of rank n. How large should n be in order to make approximation error small enough? Provided a functional norm $\|\cdot\|$ is given on the sample paths' space, the question can be stated in the average and in the probabilistic setting. Namely, find

$$n^{\text{avg}}(\varepsilon) := \inf\{n:\ E\|X-X_n\|^2 \leqslant \varepsilon^2\}$$

or

$$n^{\mathrm{pr}}(\varepsilon, \delta) := \inf\{n\colon \boldsymbol{P}\{\|X - X_n\| \geqslant \varepsilon\} \leqslant \delta\}.$$

In this work we mostly consider the random fields of tensor product-type with $T \subset \boldsymbol{R}^d$. In the first part we investigate the problem for fixed X, T, and d. Our main goal is to show that two regimes are possible: either (when δ is not very small) $n^{\mathrm{pr}}(\varepsilon, \delta)$ behaves exactly as $n^{\mathrm{avg}}(\varepsilon)$ or (when δ is very small) both parameters are important. We find a precise border between two regimes and give exact asymptotic formulas for $n^{\mathrm{avg}}(\varepsilon)$ and $n^{\mathrm{pr}}(\varepsilon, \delta)$.

In the second part of the paper we consider sequences of related tensor product-type fields $X^{(d)}(t)$, $t \in T^{(d)} \subset \boldsymbol{R}^d$, with $d \to \infty$ and study the influence of dimension parameter d. It turns out that the rank n which is necessary to obtain a relative error ε increases exponentially in d for any fixed ε. The explosion coefficient admits a simple explicit expression and does not depend on ε. Interestingly, the phenomenon of exponential explosion does not depend on the smoothness properties of the underlying fields.

Exponential explosion of the difficulty in approximation problems that include dimension parameter is well known as "dimensionality curse" or "intractability"; see e.g. [13]. Therefore, we essentially add a new probabilistic problem to the list of intractable ones.

2. APPROXIMATION IN FIXED DIMENSION

2.1. Main objects and results. We consider a random field $X(t)$, $t \in [0, 1]^d$, given by

$$X(t) = \sum_{k \in \boldsymbol{N}^d} \prod_{l=1}^{d} \lambda(k_l)\, \xi_k \prod_{l=1}^{d} \varphi_{k_l}(t_l), \tag{2.1}$$

where $(\varphi_i)_{i>0}$ is an orthonormal system in $L_2[0, 1]$ and ξ_k are non-correlated random variables with zero mean and unit variance. Therefore X is a rather typical field of so-called tensor product-type. Under the assumption

$$\sum_{i=1}^{\infty} \lambda(i)^2 < \infty \tag{2.2}$$

its sample paths belong to $L_2[0, 1]^d$ almost surely. Actually we assume more, namely

$$\lambda(i) \sim \mu i^{-r} (\log i)^q \quad \text{as } i \to \infty \tag{2.3}$$

for some $\mu > 0$, $r > 1/2$ and $q \neq -r$. For the sake of simplicity of exposition we exclude the cases $r = 1/2$, $q < -1/2$ and $r > 1/2$, $q = -r$ that satisfy (2.2) and can be investigated in the same way but lead to different and a bit more complicated formulas; cf. Example 3 in [4]. Recall that, for example, the Wiener–Chentsov's Brownian sheet belongs to the class (2.1) with $r = 1$, $q = 0$.

The covariance operator of X has the system of eigenvalues

(2.4) $$\lambda_k^2 := \prod_{l=1}^{d} \lambda(k_l)^2, \quad k \in N^d.$$

We approximate X with the finite sum of series (2.1) corresponding to n maximal eigenvalues. Let us denote this sum by X_n. It is well known (see, for example, [1], [5] or [9]) that X_n provides the minimal average quadratic error among all linear approximations to X having rank n.

Consider the average and probabilistic approximation cardinality (complexity) defined as follows:

(2.5) $$n^{\mathrm{avg}}(\varepsilon, d) := \inf\{n\colon\ E\|X - X_n\|^2_{L_2([0,1]^d)} \leqslant \varepsilon^2\},$$

(2.6) $$n^{\mathrm{pr}}(\varepsilon, \delta, d) := \inf\{n\colon\ P\{\|X - X_n\|_{L_2([0,1]^d)} \geqslant \varepsilon\} \leqslant \delta\}.$$

Let $\alpha = q/r$. The form of the results turns out to be significantly different for $\alpha < -1$ and $\alpha > -1$.

We first study the behavior of average approximation cardinality.

THEOREM 2.1. *Under the assumption* (2.3) *we have*

(2.7) $$n^{\mathrm{avg}}(\varepsilon, d) \sim \left(\frac{B_d}{\sqrt{2}\,(r-1/2)^{r\beta+1/2}} \frac{|\log \varepsilon|^{r\beta}}{\varepsilon}\right)^{1/(r-1/2)} \quad \textit{as } \varepsilon \to 0,$$

where

(2.8) $$B_d = \begin{cases} \mu^d \Pi_d^r & \textit{for } \alpha > -1, \\ \mu (dS^{d-1})^r & \textit{for } \alpha < -1, \end{cases}$$

and

(2.9) $$\beta = \begin{cases} (d-1) + d\alpha & \textit{for } \alpha > -1, \\ \alpha & \textit{for } \alpha < -1, \end{cases}$$

while

(2.10) $$S = \sum_{i=1}^{\infty} \lambda(i)^{1/r} \quad \textit{and} \quad \Pi_d = \frac{\Gamma(\alpha+1)^d}{\Gamma(d(\alpha+1))}.$$

Remark. In the simplest case, $\alpha = q = 0$, we have $\beta = d-1$, $B_d = \mu^d/(d-1)!^r$ and we obtain

$$n^{\mathrm{avg}}(\varepsilon, d) \sim \left(\frac{\mu^{2d}}{(2r-1)^{2r(d-1)+1}(d-1)!^{2r}} \frac{(2|\log \varepsilon|)^{2r(d-1)}}{\varepsilon^2}\right)^{1/(2r-1)}.$$

Furthermore, in the case of Brownian sheet we have $r = 1$, $\mu = 1/\pi$, and this leads to

$$n^{\mathrm{avg}}(\varepsilon, d) \sim \frac{1}{\pi^{2d}(d-1)!^2} \frac{(2|\log \varepsilon|)^{2(d-1)}}{\varepsilon^2}.$$

These formulas were obtained in [11].

Now we describe the behavior of probabilistic approximation cardinality in fixed dimension. In this setting we assume that our non-correlated random variables ξ_k are Gaussian (hence independent).

It turns out that two regimes are possible in the behavior of probabilistic approximation cardinality. If δ is decreasing slowly to zero (or not decreasing at all), then $n^{\mathrm{pr}}(\varepsilon, \delta, d)$ behaves like $n^{\mathrm{avg}}(\varepsilon, d)$, i.e. it does not depend on δ. On the other hand, when δ is decreasing to zero quickly, then the behavior of $n^{\mathrm{pr}}(\varepsilon, \delta, d)$ depends on both parameters.

THEOREM 2.2. *For Gaussian variables* (ξ_k), *under the assumption* (2.3) *we have two cases.*

(a) *If* $\varepsilon \to 0$ *and if we let* $\delta = \delta(\varepsilon) \in (0, 1/2)$ *vary in such a way that*

$$|\log \delta|^{r-1/2} \frac{\varepsilon}{|\log \varepsilon|^{r\beta}} \to 0, \tag{2.11}$$

then

$$\lim_{\varepsilon \to 0} \frac{n^{\mathrm{pr}}(\varepsilon, \delta, d)}{n^{\mathrm{avg}}(\varepsilon, d)} = 1. \tag{2.12}$$

(b) *If* $\delta \to 0$ *and if we let* $\varepsilon = \varepsilon(\delta) \in (0, 1/2)$ *vary in such a way that*

$$|\log \delta|^{r-1/2} \frac{\varepsilon}{|\log \varepsilon|^{r\beta}} \to \infty, \tag{2.13}$$

then

$$n^{\mathrm{pr}}(\varepsilon, \delta, d) \sim \left(\frac{B_d \sqrt{2\,|\log \delta|}}{\varepsilon}\right)^{1/r} \left(\frac{\log(|\log \delta|/\varepsilon^2)}{2r}\right)^{\beta} \quad \text{as } \delta \to 0 \tag{2.14}$$

with B_d *and* β *given in* (2.8).

Remarks. 1. A slightly weaker form of Theorem 2.2 was obtained in [11].

2. For fixed δ the result (2.12) was obtained by S. Kwapień long ago (see [10], Theorem 5.4.3, p. 339). Our theorem thus shows the limits of the validity for this effect when $\delta \to 0$ is allowed. In part (a) we do not assume $\delta \to 0$; the case $\delta = \mathrm{const}$ is therefore included.

3. The case when ε is fixed and $\delta \to 0$ was considered in [10], Theorem 5.4.2, p. 337. It is a special case of part (b).

4. It is easy to see from the proof that one can replace 1/2 with any other fixed number in (0, 1).

5. In the critical case, i.e. when

$$|\log \delta|^{r-1/2} \approx \frac{|\log \varepsilon|^{r\beta}}{\varepsilon},$$

our method yields a bilateral asymptotic estimate (but not the equivalence)

$$n^{\mathrm{pr}}(\varepsilon, \delta, d) \approx \left(\frac{|\log \varepsilon|^{r\beta}}{\varepsilon}\right)^{1/(r-1/2)}.$$

6. An interesting set of tractability problems arises when one considers tensor products of "weighted" processes, see e.g. [12].

2.2. More abstract version. Actually the results of Theorems 2.1 and 2.2 are valid in a more "abstract" setting. Let H be a Hilbert space, $(\varphi_k)_{k>0}$ be an orthonormal system in H, and (ξ_k) a sequence of non-correlated random variables with zero mean and unit variance. Let (λ_k) be a non-increasing sequence of positive numbers satisfying

$$\lambda_n \sim B n^{-r} (\log n)^{\beta r} \quad \text{as } n \to \infty \tag{2.15}$$

for some $B > 0$, $r > 1/2$, $\beta \in \boldsymbol{R}$. Consider a random vector $X \in H$ defined by

$$X = \sum_{n=1}^{\infty} \lambda_n \xi_n \varphi_n.$$

Define the approximation error

$$\Delta_n := \Big\| \sum_{m=n+1}^{\infty} \lambda_m \xi_m \varphi_m \Big\|$$

and introduce the approximation cardinality

$$n^{\mathrm{avg}}(\varepsilon, X) := \inf\{n:\ E\Delta_n^2 \leqslant \varepsilon^2\}$$

and

$$n^{\mathrm{pr}}(\varepsilon, \delta, X) := \inf\{n:\ \boldsymbol{P}\{\Delta_n \geqslant \varepsilon\} \leqslant \delta\}.$$

Our "abstract" version is as follows.

THEOREM 2.3. *It is true that:*

(a) *The behavior of* $n^{\mathrm{avg}}(\varepsilon, X)$ *is described by formula* (2.7) *with* $B_d = B$.

(b) *If the variables* (ξ_k) *are Gaussian, then the behavior of* $n^{\mathrm{pr}}(\varepsilon, \delta, X)$ *is, in the cases* (2.11) *and* (2.13), *exactly the same as in* (2.12) *and* (2.14), *respectively, with* $B_d = B$.

This theorem does not need a special proof, since the proofs of Theorems 2.1 and 2.2 use, starting from a certain point indicated below, only the representation (2.15).

2.3. Proofs

2.3.1. Proof of Theorem 2.1. This proof is based on the following elementary result.

LEMMA 2.4. *Let the eigenvalues λ_k^2 be defined by* (2.4) *and let $N_d(\varepsilon)$ be the eigenvalue distribution, i.e. the number of solutions to the inequality*

$$\lambda_k^2 \geqslant \varepsilon, \quad k \in N^d.$$

Then

$$N_d(\varepsilon) \sim C_d \varepsilon^{-1/2r} |\log \varepsilon|^{\beta} \quad \text{as } \varepsilon \to 0 \tag{2.16}$$

with β from (2.9) *and*

$$C_d = \begin{cases} (2r)^{-d\alpha-(d-1)} \mu^{d/r} \Pi_d & \text{for } \alpha > -1, \\ (2r)^{-\alpha} \mu^{1/r} d S^{d-1} & \text{for } \alpha < -1. \end{cases}$$

Similar results can be found e.g. in Csáki [3], Li [6], Papageorgiou and Wasilkowski [8] (for $q = 0$) and especially in Karol' et al. [4] for even a more general case than we need here.

An inversion of N_ε is defined as follows. Let $(\bar{\lambda}_n^2, n \in N)$ be the decreasing rearrangement of the array $(\lambda_k^2, k \in N^d)$. By inverting (2.16) we find

$$\bar{\lambda}_n^2 \sim C_d^{2r} (2r)^{2r\beta} n^{-2r} (\log n)^{2r\beta} = B_d^2 n^{-2r} (\log n)^{2r\beta} \quad \text{as } n \to \infty. \tag{2.17}$$

From now on, we can forget about tensor structure of the set of eigenvalues. The only property we use is (2.17). This is why Theorem 2.3 is proved simultaneously with other results.

By summing up the terms of (2.17), we have

$$\sum_{m>n} \bar{\lambda}_m^2 \sim B_d^2 (2r-1)^{-1} n^{1-2r} (\log n)^{2r\beta} \quad \text{as } n \to \infty. \tag{2.18}$$

By the definition of average cardinality we have

$$n^{\text{avg}}(\varepsilon, d) = \inf\{n: \sum_{m>n} \bar{\lambda}_m^2 \leqslant \varepsilon^2\},$$

and the result of Theorem 2.1 now follows from (2.18), since

$$n^{\text{avg}}(\varepsilon, d) \sim \left(\frac{B_d}{\sqrt{2}\,(r-1/2)^{r\beta+1/2}} \frac{|\log \varepsilon|^{r\beta}}{\varepsilon} \right)^{1/(r-1/2)}. \quad \blacksquare \tag{2.19}$$

2.3.2. Proof of Theorem 2.2. Let $A = A_{n,d} = \{k \in N^d: \lambda_k < \bar{\lambda}_n\}$. Consider the approximation error

$$\Delta_n = \Big\| \sum_{k \in A} \lambda_k \xi_k \varphi_k \Big\|_{L_2([0,1]^d)},$$

where λ_k are defined in (2.4) and $\varphi_k(t) = \prod_{l=1}^d \varphi_{k_l}(t_l)$. We represent Δ_n as a supremum of a centered Gaussian random function which permits to establish important concentration properties of its distribution. Indeed,

$$\Delta_n = \sup_{f \in L_2([0,1]^d), \|f\|_2 \leqslant 1} \sum_{k \in A} \lambda_k \xi_k \langle \varphi_k, f \rangle := \sup_f Y(f).$$

Let us find the maximal variance of Y. By using the orthogonality of φ_k and independence of ξ_k we have

$$\sigma_n^2 := \sup_f EY(f)^2 \leqslant \sup_f \sum_{k\in A} \lambda_k^2 \langle \varphi_k, f\rangle^2$$

$$\leqslant \sup_f \sup_{k\in A} \lambda_k^2 \sum_{k\in A} \langle \varphi_k, f\rangle^2 \leqslant \sup_f \sup_{k\in A} \lambda_k^2 \|f\|_2^2 = \sup_{k\in A} \lambda_k^2 = \bar{\lambda}_{n+1}^2.$$

Obviously, this bound is attained and we have the equality

$$\sigma_n^2 = \bar{\lambda}_{n+1}^2 \sim \bar{\lambda}_n^2.$$

By using (2.17) we get

$$\sigma_n^2 \sim B_d^2 n^{-2r} (\log n)^{2r\beta}. \tag{2.20}$$

The following inequalities are well-known consequences of the isoperimetric properties and convexity of Gaussian measures; see e.g. [2] and [7].

FACT 2.5. *Let $Y(t)$, $t \in T$, be a bounded centered Gaussian random function. Let us write $\sigma^2 = \sup_{t\in T} EY(t)^2$, $S = \sup_{t\in T} Y(t)$, and let m be a median for S. Then $m \leqslant ES$ and for any $r > 0$ we have*

$$P\{S \geqslant m+\sigma r\} \leqslant \hat{\Phi}(r) \quad \textit{and} \quad P\{S \leqslant m-\sigma r\} \leqslant \hat{\Phi}(r),$$

where $\hat{\Phi}(r)$ is the tail of the standard normal distribution function.

It follows immediately from these inequalities that $ES \leqslant m+\sqrt{1/2\pi}\,\sigma$ and $ES^2-(ES)^2 \leqslant E(S-m)^2 \leqslant \sigma^2$. Hence

$$m^2 \leqslant (ES)^2 \leqslant ES^2 \leqslant (m+\sqrt{1/2\pi}\,\sigma)^2+\sigma^2. \tag{2.21}$$

We will also need a trivial estimate

$$P\{\sup_{t\in T} |Y(t)| \geqslant \sigma r\} \geqslant 2\hat{\Phi}(r). \tag{2.22}$$

Now we turn back to Δ_n. We know from (2.18) that

$$E\Delta_n^2 \sim B_d^2 n^{1-2r} (\log n)^{2r\beta}/(2r-1) \quad \text{as } n\to\infty.$$

By comparing this with (2.20), we have $\sigma_n^2 \approx n^{-1} E\Delta_n^2 = o(E\Delta_n^2)$. Let m_n denote a median of Δ_n. By the last observation and (2.21), we have

$$m_n \sim (E\Delta_n^2)^{1/2} \sim B_d n^{1/2-r} (\log n)^{r\beta} (2r-1)^{-1/2}. \tag{2.23}$$

We are now ready to show that if δ is not very small, then $n^{\mathrm{pr}}(\varepsilon, \delta, d)$ behaves like $n^{\mathrm{avg}}(\varepsilon, d)$, i.e. it does not depend on δ. Namely, let $\varepsilon \to 0$ and let $\delta \in (0, 1/2)$ vary in such a way that

$$|\log\delta|^{r-1/2} \frac{\varepsilon}{|\log\varepsilon|^{r\beta}} \to 0. \tag{2.24}$$

Consider first arbitrary $n = n(\varepsilon)$ such that

$$n \approx n^{\mathrm{avg}}(\varepsilon, d) \approx \left(\frac{|\log \varepsilon|^{r\beta}}{\varepsilon}\right)^{1/(r-1/2)}.$$

Then by (2.23) we have $m_n \sim (E\Delta_n^2)^{1/2} \approx \varepsilon$ and, using (2.20) and (2.24) at the last step, we get

$$\begin{aligned}
(2.25) \qquad \sigma_n \sqrt{|\log \delta|} &\approx n^{-r} (\log n)^{r\beta} \sqrt{|\log \delta|} \\
&\approx \left(\frac{|\log \varepsilon|^{r\beta}}{\varepsilon}\right)^{-r/(r-1/2)} |\log \varepsilon|^{r\beta} \sqrt{|\log \delta|}- \\
&= \varepsilon^{r/(r-1/2)} |\log \varepsilon|^{-r\beta/2(r-1/2)} \sqrt{|\log \delta|} \\
&= (\varepsilon |\log \varepsilon|^{-r\beta} |\log \delta|^{r-1/2})^{1/2(r-1/2)} \varepsilon = o(\varepsilon).
\end{aligned}$$

Now specify n more thoroughly by fixing a small $h > 0$ and taking $n \sim (1+h) n^{\mathrm{avg}}(\varepsilon, d)$. Then $m_n \sim (1+h)^{1/2-r} \varepsilon$, and by (2.25) we eventually have

$$\sigma_n \sqrt{2|\log \delta|} \leqslant \tfrac{1}{2}(1-(1+h)^{1/2-r}) \varepsilon.$$

Therefore,

$$m_n + \sigma_n \sqrt{2|\log \delta|} \leqslant \varepsilon,$$

and the contraction inequality in Fact 2.5 yields

$$P(\Delta_n \geqslant \varepsilon) \leqslant P(\Delta_n \geqslant m_n + \sigma_n \sqrt{2|\log \delta|}) \leqslant \hat{\Phi}(\sqrt{2|\log \delta|}) \leqslant \delta.$$

Thus

$$n^{\mathrm{pr}}(\varepsilon, \delta, d) \leqslant n \sim (1+h) n^{\mathrm{avg}}(\varepsilon, d).$$

On the other hand, taking $n \sim (1-h) n^{\mathrm{avg}}(\varepsilon, d)$ we obtain $m_n \sim (1-h)^{1/2-r} \varepsilon > \varepsilon$, so that eventually for any $\delta < 1/2$

$$P(\Delta_n \geqslant \varepsilon) \geqslant P(\Delta_n \geqslant m_n) = \tfrac{1}{2} > \delta.$$

Consequently,

$$n^{\mathrm{pr}}(\varepsilon, \delta, d) \geqslant n \sim (1-h) n^{\mathrm{avg}}(\varepsilon, d).$$

It follows from our lower and upper bounds that

$$(2.26) \qquad \lim_{\varepsilon \to 0} \frac{n^{\mathrm{pr}}(\varepsilon, \delta, d)}{n^{\mathrm{avg}}(\varepsilon, d)} = 1,$$

as claimed in part (a) of the theorem.

Proving part (b), we will work under the assumption

$$(2.27) \qquad |\log \delta|^{r-1/2} \frac{\varepsilon}{|\log \varepsilon|^{r\beta}} \to \infty.$$

Let us first choose $n = n(\varepsilon, \delta)$ so that

$$\sigma_n \sim \frac{\varepsilon}{\sqrt{2|\log\delta|}}.$$

Using the asymptotic expression for σ_n in (2.20), we get

$$n \sim v^{-1/2r}\left(\frac{|\log v|}{2r}\right)^{\beta}, \qquad \text{where } v = \frac{\varepsilon^2}{2|\log\delta|\, B_d^2}.$$

We slightly change n by fixing a small $h > 0$ and taking $n_\pm \sim (1 \pm h)n$. Then, obviously,

$$\sigma_{n_\pm} \sim (1 \pm h)^{-r}\frac{\varepsilon}{\sqrt{2|\log\delta|}}. \tag{2.28}$$

Next, we will derive from (2.27) that

$$m_{n_\pm} \ll \varepsilon. \tag{2.29}$$

Indeed, by straight comparison of (2.20) and (2.23) one observes that

$$m_n \approx \sigma_n^{(2r-1)/2r}|\log\sigma_n|^{\beta/2}.$$

Under the assumption (2.27) it is true that

$$\sigma_{n_\pm} \approx \frac{\varepsilon}{\sqrt{2|\log\delta|}} \ll \frac{\varepsilon^{2r/(2r-1)}}{|\log\varepsilon|^{r\beta/(2r-1)}} := \bar{\varepsilon},$$

and we have

$$m_{n_\pm} \approx \sigma_{n_\pm}^{(2r-1)/2r}|\log\sigma_{n_\pm}|^{\beta/2} \ll \bar{\varepsilon}^{(2r-1)/2r}|\log\bar{\varepsilon}|^{\beta/2} \approx \varepsilon.$$

It follows from (2.28) and (2.29) that for small ε

$$\varepsilon \geqslant m_{n_+} + \sigma_{n_+}\sqrt{2|\log\delta|}.$$

Hence, by the contraction inequality in Fact 2.5,

$$\boldsymbol{P}(\Delta_{n_+} \geqslant \varepsilon) \leqslant \boldsymbol{P}(\Delta_{n_+} \geqslant m_{n_+} + \sigma_{n_+}\sqrt{2|\log\delta|}) \leqslant \hat{\Phi}(2|\log\delta|) \leqslant \delta.$$

Therefore,

$$n^{\mathrm{pr}}(\varepsilon, \delta, d) \leqslant n_+ \sim (1+h)v^{-1/2r}\left(\frac{|\log v|}{2r}\right)^{\beta}.$$

On the other hand, for small δ it is true that $\varepsilon \leqslant (1-h/2)^r\sigma_{n_-}\sqrt{2|\log\delta|}$ and by (2.22) we have, for small δ,

$$\boldsymbol{P}(\Delta_{n_-} \geqslant \varepsilon) \geqslant \boldsymbol{P}\left(\Delta_{n_-} \geqslant (1-h/2)^r\sigma_{n_-}\sqrt{2|\log\delta|}\right) \geqslant 2\hat{\Phi}\left((1-h/2)^r\sqrt{2|\log\delta|}\right) \geqslant \delta.$$

Therefore,

$$n^{\mathrm{pr}}(\varepsilon, \delta, d) \geqslant n_{-} \sim (1-h)\, v^{-1/2r}\left(\frac{|\log v|}{2r}\right)^{\beta},$$

and we obtain part (b) of Theorem 2.2 by letting $h \to 0$. ■

3. APPROXIMATION IN INCREASING DIMENSION

3.1. Setting and results. In this section we study the approximation error of tensor product-type fields $X(t)$, $t \in [0, 1]^d$, given by (2.1), when $d \to \infty$. We still assume that (2.2) holds. On the contrary, there is no need in regularity assumption (2.3). Let us stress that we cannot just take the asymptotic results in fixed dimension and then let $d \to \infty$ even on the heuristical level. This would lead to false conclusions. Our analysis is therefore independent of the results in the previous section.

When dealing with approximation of a *sequence* of random fields, it is more natural to work with *relative* errors, thus taking into account the size of varying approximation target. Therefore, let us first calculate

$$\boldsymbol{E}\|X\|^2 = \sum_{k \in \boldsymbol{N}^d} \lambda_k^2 \Big(\sum_{i=1}^{\infty} \lambda(i)^2\Big)^d := \Lambda^d,$$

then define and evaluate the relative average approximation cardinality

$$\tilde{n}^{\mathrm{avg}}(\varepsilon, d) := \inf\{n\colon\ \boldsymbol{E}\Delta_n^2 \leqslant \varepsilon^2 \Lambda^d\}$$

and the relative probabilistic approximation cardinality

$$\tilde{n}^{\mathrm{pr}}(\varepsilon, \delta, d) := \inf\{n\colon\ \boldsymbol{P}\{\Delta_n^2 > \varepsilon^2 \Lambda^d\} \leqslant \delta\},$$

where Δ_n is, as before, the norm of the error in approximation of X by the n-term partial sum from (2.1). We will show that for any fixed ε the cardinality $\tilde{n}^{\mathrm{avg}}(\varepsilon, d)$ is increasing exponentially in d, but even before stating our result we must explain our approach to the problem. Of course, we have a representation of cardinality via the ordered eigenvalues:

$$\tilde{n}^{\mathrm{avg}}(\varepsilon, d) := \inf\{n\colon\ \sum_{m>n} \bar{\lambda}_m^2 \leqslant \varepsilon^2 \Lambda^d\},$$

and the problem boils down to a study of deterministic arrays $(\lambda_k,\ k \in \boldsymbol{N}^d)$ and inverting their decreasing rearrangements $(\bar{\lambda}_m)$. To a great surprise, the properties of these objects can be properly understood in the language of a simple auxiliary *probabilistic* construction. Namely, let us introduce a sequence of i.i.d. random variables (U_l), $l = 1, 2, \ldots$, with the common distribution given by

$$\boldsymbol{P}\big(U_l = -\log \lambda(i)\big) = \frac{\lambda(i)^2}{\Lambda}, \qquad i = 1, 2, \ldots$$

The role of this sequence is explained by the following lemma.

LEMMA 3.1. *For any $d \in N$ and any $0 < a < b$, we have*

$$\sum_{k \in N^d: a \leqslant \lambda_k < b} \lambda_k^2 = \Lambda^d \boldsymbol{P}\Big(\sum_{l=1}^{d} U_l \in (-\log b, -\log a]\Big). \tag{3.1}$$

Proof. Let

$$A := \{k \in N^d:\ a \leqslant \lambda_k < b\} = \Big\{k \in N^d:\ a \leqslant \prod_{l=1}^{d} \lambda(k_l) < b\Big\}.$$

$$= \Big\{k \in N^d:\ -\log b < -\sum_{l=1}^{d} \log \lambda(k_l) \leqslant -\log a\Big\}.$$

By the definition of (λ_k), (U_l), and A, we have

$$\begin{aligned}
\sum_{k \in A} \lambda_k^2 &= \sum_{k \in A} \prod_{l=1}^{d} \lambda(k_l)^2 = \Lambda^d \sum_{k \in A} \prod_{l=1}^{d} \boldsymbol{P}\big(U_l = -\log \lambda(k_l)\big) \\
&= \Lambda^d \sum_{k \in A} \boldsymbol{P}\big((U_1, \ldots, U_d) = (-\log \lambda(k_1), \ldots, -\log \lambda(k_d))\big) \\
&= \Lambda^d \boldsymbol{P}\big((U_1, \ldots, U_d) = (-\log \lambda(k_1), \ldots, -\log \lambda(k_d)) \text{ for some } k \in A\big) \\
&= \Lambda^d \boldsymbol{P}\Big(\sum_{l=1}^{d} U_l \in (-\log b, -\log a]\Big). \ \blacksquare
\end{aligned}$$

Combining Lemma 3.1 with Chebyshev's inequality we see that for any $d \in N$ and any $0 \leqslant a < b \leqslant +\infty$ it is true that

$$\#\{k \in N^d:\ a \leqslant \lambda_k < b\} \leqslant \frac{\Lambda^d}{a^2} \boldsymbol{P}\Big(\sum_{l=1}^{d} U_l \in (-\log b, -\log a]\Big) \tag{3.2}$$

and

$$\#\{k \in N^d:\ a \leqslant \lambda_k < b\} \geqslant \frac{\Lambda^d}{b^2} \boldsymbol{P}\Big(\sum_{l=1}^{d} U_l \in (-\log b, -\log a]\Big). \tag{3.3}$$

In the following we assume that our basic sequence $(\lambda(i))$ satisfies

$$\sum_{i=1}^{\infty} |\log \lambda(i)|^2 \lambda(i)^2 < \infty. \tag{3.4}$$

This condition is of course the same as $\boldsymbol{E} U_l^2 < \infty$ and it is true for regular sequences (2.3). Let

$$M := -\sum_{i=1}^{\infty} \log \lambda(i) \frac{\lambda(i)^2}{\Lambda} = \boldsymbol{E} U_l$$

and

$$\sigma^2 = \sum_{i=1}^{\infty} |\log \lambda(i)|^2 \frac{\lambda(i)^2}{\Lambda} - M^2 = \mathbf{Var}\, U_l.$$

If (3.4) is satisfied, we have $|M| < \infty$ and $0 \leqslant \sigma^2 < \infty$. In the following, the role of the explosion coefficient

$$\mathscr{E} := \Lambda e^{2M}$$

will be crucial. Let us immediately check that $\mathscr{E} > 1$ except for the totally degenerate case when the number of strictly positive $\lambda(i)$'s is zero or one (in other words, $\mathscr{E} = 1$ when $\sigma = 0$). This degenerate case is excluded from the subsequent consideration.

By concavity of the logarithmic function, we have

$$-2M = 2\sum_{i=1}^{\infty} \log\lambda(i)\frac{\lambda(i)^2}{\Lambda} = \sum_{i=1}^{\infty} \log\left(\lambda(i)^2\right)\frac{\lambda(i)^2}{\Lambda}$$
$$\leqslant \log\left(\sum_{i=1}^{\infty} \lambda(i)^2\frac{\lambda(i)^2}{\Lambda}\right) < \log\left(\frac{\left(\sum_{i=1}^{\infty}\lambda(i)^2\right)^2}{\Lambda}\right) = \log\frac{\Lambda^2}{\Lambda} = \log\Lambda.$$

Hence $\log\mathscr{E} = \log\Lambda + 2M > 0$, and we obtain $\mathscr{E} > 1$.

Now we can state the main result of this section.

THEOREM 3.2. *Under the assumption* (3.4) *we have for any* $\varepsilon \in (0, 1)$

$$\lim_{d\to\infty} \frac{\log\tilde{n}^{\mathrm{avg}}(\varepsilon, d) - d\log\mathscr{E}}{\sqrt{d}} = 2q,$$

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

$$\hat{\Phi}(q/\sigma) = \varepsilon^2. \tag{3.5}$$

COROLLARY 3.3. *Under the same assumption, for any* $\varepsilon \in (0, 1)$ *we have*

$$\lim_{d\to\infty} \frac{\log\tilde{n}^{\mathrm{avg}}(\varepsilon, d)}{d} = \log\mathscr{E}.$$

Under further assumptions on $(\lambda(i))$ one can prove that

$$\tilde{n}^{\mathrm{avg}}(\varepsilon, d) \approx \frac{C(\varepsilon)\,\mathscr{E}^d\exp(2q\sqrt{d})}{\sqrt{d}} \quad \text{as } d\to\infty,$$

thus giving more than logarithmic behavior, but we are not going to provide further details here.

Now we show that in a very large zone of parameters the probabilistic cardinality behaves in the same way as the average one.

THEOREM 3.4. *Let* $\varepsilon \in (0, 1)$ *be fixed and define* $q = q(\varepsilon)$ *as in* (3.5). *If we let* $\delta = \delta(d) \in (0, 1/2)$ *vary in such a way that*

$$\lim_{d\to\infty} \frac{|\log\delta|}{\mathscr{E}^d\exp(2q\sqrt{d})} = 0, \tag{3.6}$$

then

$$\lim_{d\to\infty} d^{-1/2} \log \frac{\log \tilde{n}^{\mathrm{pr}}(\varepsilon, \delta, d)}{\log \tilde{n}^{\mathrm{avg}}(\varepsilon, d)} = 0. \tag{3.7}$$

We stress that the assumption (3.6) is extremely mild (the lower bound imposed on δ is a double exponential).

COROLLARY 3.5. *Under the same assumption, we have*

$$\lim_{d\to\infty} \frac{\log \tilde{n}^{\mathrm{pr}}(\varepsilon, \delta, d)}{\log \tilde{n}^{\mathrm{avg}}(\varepsilon, d)} = 1.$$

3.2. Proofs

3.2.1. Proof of Theorem 3.2. We start the proof of the upper bound by fixing a small $h > 0$ and setting

$$\zeta = \exp\left(-\left(Md+(q+h)\sqrt{d}\right)\right).$$

Apply the central limit theorem:

$$\lim_{d\to\infty} P\left(\sum_{l=1}^{d} U_l > Md+(q+h)\sqrt{d}\right) = \hat{\Phi}\left(\frac{q+h}{\sigma}\right) < \varepsilon^2. \tag{3.8}$$

It follows now from (3.1) applied with $a = 0$ and $b = \zeta$ that for all d large enough

$$\sum_{k\in N^d:\lambda_k<\zeta} \lambda_k^2 < \Lambda^d \varepsilon^2. \tag{3.9}$$

On the other hand, let us set $a = \zeta$ and $b = +\infty$ and apply (3.2). We get

$$\#\{k\in N^d\colon \lambda_k \geqslant \zeta\} \leqslant \frac{\Lambda^d}{\zeta^2} \Lambda^d \exp\left(2Md+2(q+h)\sqrt{d}\right). \tag{3.10}$$

It follows from (3.9) and (3.10) that for all d large enough

$$\tilde{n}^{\mathrm{avg}}(\varepsilon, d) \leqslant \mathscr{E}^d \exp\left(2(q+h)\sqrt{d}\right),$$

and we are done with the upper estimate.

The lower bound will be obtained similarly. Take $h > 0$ and set

$$\zeta = \exp\left(-\left(Md+(q-h)\sqrt{d}\right)\right).$$

As in (3.8), the central limit theorem yields

$$\lim_{d\to\infty} P\left(\sum_{l=1}^{d} U_l > Md+(q-h)\sqrt{d}\right) = \hat{\Phi}\left(\frac{q-h}{\sigma}\right) > \varepsilon^2. \tag{3.11}$$

It follows from (3.1), applied with $a = 0$ and $b = \zeta$, that

$$\sum_{k\in N^d:\lambda_k<\zeta} \lambda_k^2 > \Lambda^d \varepsilon^2. \tag{3.12}$$

The lower estimate for the number of "large" eigenvalues is just a bit more delicate. This time we set $a = \zeta$, $b = \zeta\exp(h\sqrt{d})$ and apply (3.3). We get

$$\begin{aligned}
\#\{k\in N^d:\ \lambda_k \geqslant \zeta\} &\geqslant \#\{k\in N^d:\ \zeta \leqslant \lambda_k < b\}\\
&\geqslant \frac{\Lambda^d}{b^2} P\Big(\sum_{l=1}^{d} U_l \in (-\log b,\ -\log\zeta]\Big)\\
&= \mathcal{E}^d \exp\big(2(q-2h)\sqrt{d}\big) P\Big(\sum_{l=1}^{d} U_l \in (Md+(q-2h)\sqrt{d},\ Md+(q-h)\sqrt{d}]\Big)\\
&\sim \mathcal{E}^d \exp\big(2(q-2h)\sqrt{d}\big)\Big(\Phi\Big(\frac{q-h}{\sigma}\Big)-\Phi\Big(\frac{q-2h}{\sigma}\Big)\Big).
\end{aligned}$$

Therefore, for large d we have

$$\#\{k\in N^d:\ \lambda_k \geqslant \zeta\} \geqslant \mathcal{E}^d \exp\big(2(q-3h)\sqrt{d}\big). \tag{3.13}$$

Combining (3.12) and (3.13) we see that

$$\tilde{n}^{\text{avg}}(\varepsilon, d) \geqslant \mathcal{E}^d \exp\big((q-3h)\sqrt{d}\big),$$

and we are also done with the lower estimate. ■

3.2.2. Proof of Theorem 3.4. We fix $\varepsilon\in(0,1)$ and choose the quantile $q = q(\varepsilon)$ as in (3.5). Similarly, take $\varepsilon_1 < \varepsilon$ and choose q_1 from the equation $\hat{\Phi}(q_1/\sigma) = \varepsilon_1^2$. Set $h_1 = q_1 - q > 0$.

For each d we choose ζ_d so that

$$\#\{k\in N^d:\ \lambda_k \geqslant \zeta_d\} = \tilde{n}^{\text{avg}}(\varepsilon_1, d).$$

Recall that for all d large enough we have

$$\tilde{n}^{\text{avg}}(\varepsilon_1, d) > \mathcal{E}^d \exp\big(2(q_1-h_1)\sqrt{d}\big) = \mathcal{E}^d \exp\big(2q\sqrt{d}\big).$$

We know from (3.2) that for any $\zeta > 0$

$$\#\{k\in N^d:\ \lambda_k \geqslant \zeta\} \leqslant \Lambda^d/\zeta^2.$$

By setting $\zeta := \exp(-Md - q\sqrt{d})$ we obtain

$$\#\{k\in N^d:\ \lambda_k \geqslant \zeta\} \leqslant \mathcal{E}^d \exp\big(2q\sqrt{d}\big).$$

Hence

$$\zeta_d \leqslant \zeta = \exp(-Md - q\sqrt{d}).$$

Next, we set

$$\Delta^{(d)} := \Big\|\sum_{k\in N^d:\lambda_k<\zeta_d} \lambda_k \xi_k\Big\|.$$

By the definitions of ζ_d and $\tilde{n}^{\text{avg}}(\varepsilon_1, d)$ we know that $E(\Delta^{(d)})^2 \leqslant \varepsilon_1^2 \Lambda^d$. Therefore, by the Gaussian contraction inequality in Fact 2.5 we have

$$P\big(\Delta^{(d)} \geqslant \varepsilon_1 \Lambda^{d/2} + \zeta_d\sqrt{|\log\delta|}\big) \leqslant \delta.$$

It follows from the assumption (3.6) that

$$|\log\delta| \ll \mathcal{E}^d \exp(2q\sqrt{d}) \approx (\varepsilon-\varepsilon_1)^2 \mathcal{E}^d \exp(2q\sqrt{d});$$

hence for d large enough

$$\zeta_d^2 |\log\delta| \leqslant (\varepsilon-\varepsilon_1)^2 \Lambda^d$$

and, finally,

$$\varepsilon_1 \Lambda^{d/2} + \zeta_d \sqrt{|\log\delta|} \leqslant \varepsilon\Lambda^{d/2}.$$

We have

$$\boldsymbol{P}(\Delta^{(d)} \geqslant \varepsilon\Lambda^{d/2}) \leqslant \delta.$$

By the definition of cardinality, it means that $\tilde{n}^{\mathrm{pr}}(\varepsilon, \delta, d) \leqslant \tilde{n}^{\mathrm{avg}}(\varepsilon_1, d)$. Applying twice Theorem 3.2, we have, for d large,

$$\tilde{n}^{\mathrm{pr}}(\varepsilon, \delta, d) \leqslant \tilde{n}^{\mathrm{avg}}(\varepsilon_1, d) \leqslant \mathcal{E}^d \exp\big(2(q_1+h_1)\sqrt{d}\big)$$
$$= \mathcal{E}^d \exp\big(2(q+2h_1)\sqrt{d}\big) \leqslant \tilde{n}^{\mathrm{avg}}(\varepsilon, d) \exp(5h_1\sqrt{d}),$$

and the required upper bound follows, since h_1 could be chosen arbitrarily small.

The lower bound is even much easier since from (2.21) it follows that for $\delta < 1/2$ we have $\tilde{n}^{\mathrm{pr}}(\varepsilon, \delta, d) \geqslant \tilde{n}^{\mathrm{avg}}(\varepsilon, d)$. ■

Acknowledgements. We are much indebted to H. Woźniakowski for introduction to this problem and for useful comments. We are also grateful to the anonymous referee for careful reading of the manuscript. The work of the first-named author was supported by grants RFBR 05-01-00911, RFBR/DFG 04-01-04000 and INTAS 03-51-5018.

REFERENCES

[1] A. P. Buslaev and O. V. Seleznjev, *On certain extremal problems in the theory of approximation of random processes*, East J. Approx. 5 (4) (1999), pp. 467–481.

[2] Yu. A. Davydov, M. A. Lifshits and N. V. Smorodina, *Local Properties of Distributions of Stochastic Functionals*, Transl. Math. Monogr. 173, Amer. Math. Society, Providence 1998.

[3] E. Csáki, *On small values of the square integral of a multiparameter Wiener process*, in: *Statistics and Probability. Proceedings of the 3rd Pannonian Symposium on Mathematical Statistics*, D. Reidel, Boston 1982, pp. 19–26.

[4] A. I. Karol', A. I. Nazarov and Ya. Yu. Nikitin, *Tensor products of compact operators and logarithmic L_2-small ball asymptotics for Gaussian random fields*, Studi Statistici N. 74, Istituto di Metodi Quantitativi, Università 'L. Bocconi', Milano 2003, 30 pp.

[5] Th. Kühn and W. Linde, *Optimal series representation of fractional Brownian sheets*, Bernoulli 8 (5) (2002), pp. 669–696.

[6] W. V. Li, *Comparison results for the lower tail of Gaussian seminorms*, J. Theoret. Probab. 5 (1992), pp. 1–31.
[7] M. A. Lifshits, *Gaussian Random Functions*, Kluwer, 1996.
[8] A. Papageorgiou and G. W. Wasilkowski, *On the average complexity of multivariate problems*, J. Complexity 6 (1990), pp. 1–23.
[9] K. Ritter, *Average-case Analysis of Numerical Problems*, Lecture Notes in Math. No 1733, Springer, Berlin 2000.
[10] J. F. Traub, G. W. Wasilkowski and H. Woźniakowski, *Information-based Complexity*, Computer Science and Scientific Computing, Academic Press, Inc., Boston, MA, 1988.
[11] E. V. Tulyakova, *Curse of dimensionality in approximation of Gaussian random fields*, Master Thesis, St. Petersburg University, 2005.
[12] G. W. Wasilkowski and H. Woźniakowski, *Weighted tensor-product algorithms for linear multivariate problems*, J. Complexity 15 (1999), pp. 1–56.
[13] H. Woźniakowski, *Tractability and strong tractability of linear multivariate problems*, J. Complexity 10 (1994), pp. 96–128.

Corresponding author:
Mikhail Lifshits
St. Petersburg State University
198504 Stary Peterhof
Department of Mathematics and Mechanics
Bibliotechnaya pl., 2
Russia
E-mail: lifts@mail.rcom.ru

Received on 25.5.2006