COLLOQUIUM MATHEMATICUM

VOL. 139 2015 NO. 2

THE MAJORIZING MEASURE APPROACH TO SAMPLE BOUNDEDNESS

BY

WITOLD BEDNORZ (Warszawa)

Abstract. We describe an alternative approach to sample boundedness and continuity of stochastic processes. We show that the regularity of paths can be understood in terms of the distribution of the argument maximum. For a centered Gaussian process $X(t)$, $t \in T$, we obtain a short proof of the exact lower bound on $\mathbb{E}\sup_{t\in T} X(t)$. Finally we prove the equivalence of the usual majorizing measure functional to its conjugate version.

1. Introduction. Consider a Gaussian process $X(t)$, $t \in T$, on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, that is, a jointly Gaussian family of centered r.v. indexed by T. We can then provide T with the canonical distance

$$d(s,t) = (\mathbb{E}(X(s) - X(t))^2)^{1/2}, \quad s,t \in T.$$

If $X(t)$, $t \in T$, is sample bounded then the space (T, d) is totally bounded since otherwise by Slepian's lemma (see e.g. [11, Theorem 3.18]) one can find a countable subset $S \subset T$ such that $\mathbb{E}\sup_{t\in S} X(t) = \infty$. This implies that $\mathrm{Diam}(T) = \sup_{s,t\in T} d(s,t) < \infty$, and taking the Cauchy closure of (T, d) one can assume that (T, d) is a compact metric space. This implies that there exists a separable modification of $X(t)$, $t \in T$ (which we refer to from now on), and therefore $\sup_{t\in T} X(t)$ is well defined. The *sample boundedness* of $X(t)$, $t \in T$, means that $\sup_{t\in T} X(t) < \infty$ almost surely. Due to the Gaussian concentration inequality this is equivalent to the finiteness of the mean value, namely

$$\mathbb{E}\sup_{t\in T} X(t) < \infty. \tag{1.1}$$

On the other hand note that

$$\mathbb{E}\sup_{t\in T} X(t) = \sup_{F\subset T} \mathbb{E}\sup_{t\in F} X(t), \tag{1.2}$$

where the supremum is taken over all finite subsets F of T. Hence (1.2) provides an alternative definition of $\mathbb{E}\sup_{t\in T} X(t)$, which can be used without introducing any modification of the basic process.

2010 *Mathematics Subject Classification*: Primary 60G15; Secondary 60G17.
Key words and phrases: sample boundedness, majorizing measures, Gaussian processes.

DOI: 10.4064/cm139-2-4

The second basic question on Gaussian processes is the continuity of paths. We say that $X(t)$, $t \in T$, is *continuous* if $(T,d) \ni t \mapsto X(t,\omega) \in \mathbb{R}$ is continuous for almost all $\omega \in \Omega$. There exists a natural family of quantities whose analysis implies a complete characterization of the continuity property. For each $\delta > 0$, define

$$\mathcal{S}(\delta) = \mathbb{E} \sup_{s,t \in T,\, d(s,t) \le \delta} |X(s) - X(t)|.$$

Then continuity is equivalent to $\lim_{\delta \to 0} \mathcal{S}(\delta) = 0$ (see e.g. [11, Chapter 12] or [1, Chapter 3]).

In this paper, K denotes a universal constant that may change from line to line. The standard approach to the regularity of Gaussian processes goes through entropy numbers. Let $B(t,\varepsilon)$ be the ball in T of radius ε, centered at t, i.e. $B(t,\varepsilon) = \{x \in T : d(x,t) \le \varepsilon\}$. Denote by $N(T,d,\varepsilon)$ the smallest number of balls of radius $\varepsilon > 0$ that cover T. The simplest upper bound of $\mathbb{E}\sup_{t \in T} X(t)$ was proved in [6, 13] to be

$$\mathbb{E} \sup_{t \in T} X(t) \le K \int_0^\infty \sqrt{\log_2(N(T,d,\varepsilon))}\, d\varepsilon.$$

Therefore $\int_0^\infty \sqrt{\log_2(N(T,d,\varepsilon))}\, d\varepsilon < \infty$ is a sufficient condition for (1.1) to hold. Since it can also be proved that

$$\mathcal{S}(\delta) \le K \int_0^\delta \sqrt{\log_2(N(T,d,\varepsilon))}\, d\varepsilon,$$

the condition $\int_0^\infty \sqrt{\log_2(N(T,d,\varepsilon))}\, d\varepsilon < \infty$ in fact implies the continuity of $X(t)$, $t \in T$. Unfortunately entropy numbers do not solve the regularity questions completely: there are sample bounded Gaussian processes of infinite entropy (e.g. ellipsoids in Hilbert space [17]) and there are discontinuous Gaussian processes that are sample bounded.

A better tool than entropy is majorizing measures. We say that a probability Borel measure m is *majorizing* if

$$\sup_{t \in T} \int_0^\infty \sqrt{\log_2(m(B(t,\varepsilon))^{-1})}\, d\varepsilon < \infty. \tag{1.3}$$

Generalizing the notion of majorizing measure let, for probability measures μ, ν,

$$\mathcal{M}(\mu,\nu,\delta) = \int_T \int_0^\delta \sqrt{\log_2(\mu(B(t,\varepsilon))^{-1})}\, d\varepsilon\, \nu(dt)$$

and $\mathcal{M}(\mu,\nu) = \mathcal{M}(\mu,\nu,\mathrm{Diam}(T)) = \mathcal{M}(\mu,\nu,\infty)$. Obviously

$$\mathcal{M}(\mu,\delta_t,\delta) = \int_0^\delta \sqrt{\log_2(m(B(t,\varepsilon))^{-1})}\,d\varepsilon,$$

where δ_t is the delta measure at $t \in T$. A simple chaining argument shows (see [8]) that the existence of a majorizing measure suffices for sample boundedness of $X(t)$, $t \in T$.

THEOREM 1.1. *For each Gaussian $X(t)$, $t \in T$, we have*

$$\mathbb{E}\sup_{t\in T} X(t) \le K \inf_{\mu}\sup_{t\in T} \mathcal{M}(\mu,\delta_t).$$

The idea of using majorizing measures to study sample boundedness was developed in [15] and later in [2]. In the Gaussian setting the difficult part was to establish that the existence of a majorizing measure is necessary when $X(t)$, $t \in T$, satisfies (1.1). This was first proved in [14].

THEOREM 1.2. *For each Gaussian $X(t)$, $t \in T$,*

$$\mathbb{E}\sup_{t\in T} X(t) \ge K^{-1} \inf_{\mu}\sup_{t\in T} \mathcal{M}(\mu,\delta_t).$$

Moreover (see e.g. [11, Chapter 12]) a Gaussian process $X(t)$, $t \in T$, is continuous if and only if for some probability measure μ,

$$\lim_{\delta\to 0}\sup_{t\in T} \mathcal{M}(\mu,\delta_t,\delta) = 0.$$

A simpler argument for Theorem 1.2 appeared in [16], and finally in [17] the language of majorizing measures was replaced by admissible partitions. Each of the methods contains an important constructive part, where one has to construct a suitable admissible partition or a majorizing measure.

In this paper we propose a different approach. From [8] it is known that whenever $\sup_\mu \mathcal{M}(\mu,\mu) < \infty$ then there exists a majorizing measure on T, namely

$$\inf_{\mu}\sup_{t\in T} \mathcal{M}(\mu,\delta_t) \le \sup_{\mu} \mathcal{M}(\mu,\mu).$$

The quantity $\sup_\mu \mathcal{M}(\mu,\mu)$ is a natural upper bound for processes. Note that for each $X(t)$, $t \in F$, where F is a finite subset of T, there exists a random t_F valued in F such that

$$\mathbb{E}\sup_{t\in F} X(t) = \mathbb{E}X(t_F). \tag{1.4}$$

DEFINITION 1.3. Let $\mu_F(t) := \mathbb{P}(t_F = t)$ for $t \in F$, where t_F is given by (1.4).

This measure can be treated as the distribution of the supremum argument on F. We show in Section 2 in the general setting of processes of bounded increments that $\mathcal{M}(\mu_F, \mu_F)$ is the right upper bound for the mean value of the supremum.

THEOREM 1.4. *For each Gaussian $X(t)$, $t \in T$, and finite $F \subset T$,*

$$\mathbb{E} \max_{t \in F} X(t) \le K \mathcal{M}(\mu_F, \mu_F).$$

Note that in the case of Gaussian processes the above property was proved in [1, Theorem 4.2], yet it is also mentioned in [7] and was known to Talagrand [14]. There are many cases (see [3, 4, 5]) where one can prove the lower bound for the supremum of stochastic processes in the form $\sup_\mu \mathcal{M}(\mu, \mu)$. The benefit of the approach is that the lower bound has to be found for a given measure μ on T, which better fits the chaining argument. Moreover one can reduce the constructive part of the lower bound proof to the definition of a natural partitioning sequence on (T, d), which is described in Section 3. Consequently (see Sections 4, 5), using this idea we give a short proof of the following lower bound.

THEOREM 1.5. *For each Gaussian $X(t)$, $t \in T$,*

$$\mathbb{E} \sup_{t \in T} X(t) \ge K^{-1} \sup_\mu \mathcal{M}(\mu, \mu).$$

In this way we deduce that $\mathbb{E} \sup_{t \in T} X(t)$ is comparable with $\sup_\mu \mathcal{M}(\mu, \mu)$ up to a universal constant. In particular, this shows the well known property

$$K^{-1} \inf_\mu \sup_{t \in T} \mathcal{M}(\mu, \delta_t) \le \sup_\mu \mathcal{M}(\mu, \mu) \le K \inf_\mu \sup_{t \in T} \mathcal{M}(\mu, \delta_t). \tag{1.5}$$

We prove in Section 8 that (1.5) holds in a much generalized setting (of processes under certain increment conditions). Another question is whether or not there exists a measure μ_T such that $\mathbb{E} \sup_{t \in T} X(t)$ is comparable with $\mathcal{M}(\mu_T, \mu_T)$. Such a measure μ_T should be treated as an asymptotic argument supremum distribution, i.e. a weak limit of μ_{F_n} for an increasing sequence of finite F_n that approximates T. Apparently the result requires the continuity of the process $X(t)$, $t \in T$.

THEOREM 1.6. *If $X(t)$, $t \in T$, is a continuous Gaussian process then there exists a measure μ_T on T such that*

$$K^{-1} \mathcal{M}(\mu_T, \mu_T) \le \mathbb{E} \sup_{t \in T} X(t) \le K \mathcal{M}(\mu_T, \mu_T).$$

Moreover μ_T is any cluster point of any sequence (μ_{F_n}) given by Definition 1.3 *where $F_n \subset F_{n+1}$ and $\bigcup_n F_n$ is dense in T.*

The meaning of Theorem 1.6 is that for continuous processes there exists an asymptotic supremum distribution, which also agrees with the result of [9], where it is proved that the supremum argument exists for continuous Gaussian processes at least up to a modification of the probability space. Obviously if there exists a well defined supremum argument, i.e. a random variable t_T such that $\mathbb{E}\sup_{t\in T} X(t) = \mathbb{E}X(t_T)$, then the proof of Theorem 1.5 shows that $\mathbb{E}\sup_{t\in T} X(t) \le K\mathcal{M}(\mu_T, \mu_T)$, where μ_T is the distribution of t_T. Therefore for continuous Gaussian processes there exists a natural measure μ_T that can be used to characterize the finiteness of $\mathbb{E}\sup_{t\in T} X(t)$.

In the proof of Theorem 1.6 we use the following general estimate on $\mathcal{S}(\delta)$.

THEOREM 1.7. *For each Gaussian $X(t)$, $t \in T$,*

$$K^{-1}\sup_{c>0}\sup_{\mu}\big(\mathcal{M}(\mu,\mu,c) - c\sqrt{\log_2(N(T,d,\delta))}\big) \le \mathcal{S}(\delta) \le K\sup_{\mu}\mathcal{M}(\mu,\mu,2\delta).$$

In particular $X(t)$, $t \in T$, is continuous if and only if

$$\lim_{\delta\to 0}\sup_{\mu}\mathcal{M}(\mu,\mu,\delta) = 0.$$

Proof of Theorems 1.6 and 1.7 are provided in Section 6. Then in Section 7 we study the main toy example for the theory—the Hilbert–Schmidt ellipsoid. In the slightly simplified case of Bernoulli random vectors we give a sufficient description of the supremum distribution to obtain the right upper and lower bounds on the expectation of the supremum. We also point out how the theory may be used in the analysis of small value distribution of centered random vectors valued in the Euclidean space. Finally, in Section 8 we show a duality principle. We consider the quantity $\sup_{\mu}\inf_{t\in T}\mathcal{M}(\mu,\delta_t)$ and prove that in the general setting of processes of bounded increments it is comparable with $\inf_{\mu}\sup_{t\in T}\mathcal{M}(\mu,\delta_t)$ and hence also with $\sup_{\mu}\mathcal{M}(\mu,\mu)$. This is an extension of the result discussed in a recent paper [12] and used to prove generalizations of the Dvoretzky theorem to arbitrary metric spaces. In particular in the Gaussian like setting we have

THEOREM 1.8. *We have*

$$K^{-1}\sup_{\mu}\inf_{t\in T}\mathcal{M}(\mu,\delta_t) \le \sup_{\mu}\mathcal{M}(\mu,\mu) \le K\sup_{\mu}\inf_{t\in T}\mathcal{M}(\mu,\delta_t).$$

2. The upper bound. In this section we collect all the upper bounds required in this paper. The basic theory was given in [15] and then slightly developed in [2] and [3]. First note that our measure approach works in a much more generalized setting. Let (T,ρ) be any compact metric space and φ a *Young function*, i.e. convex, increasing, $\varphi(0) = 0$ and $\varphi(1) = 1$. The centered process $X(t)$, $t \in T$, is of *bounded increments* if

$$\mathbb{E}\varphi\Big(\frac{|X(s)-X(t)|}{\rho(s,t)}\Big) \le 1, \qquad s,t \in T, \tag{2.1}$$

i.e. $\|X(t)-X(s)\|_\varphi \le \rho(s,t)$. Let $\mathrm{Diam}_\rho(T)$ and $B_\rho(t,\varepsilon)$ be the diameter and the ball in the ρ metric. Moreover define

$$\sigma_{\mu,\rho,\varphi}(t,\delta)=\int_0^\delta \varphi^{-1}\biggl(\frac{1}{\mu(B(t,\varepsilon))}\biggr)\,d\varepsilon,$$
$$\mathcal{M}_{\rho,\varphi}(\mu,\nu,\delta)=\int_T \sigma_{\mu,\rho,\varphi}(t,\delta)\,\nu(dt).$$

For simplicity let

$$\sigma_{\mu,\rho,\varphi}(t)=\sigma_{\mu,\rho,\varphi}(t,\mathrm{Diam}_\rho(T)),\qquad \mathcal{M}_{\rho,\varphi}(\mu,\nu)=\mathcal{M}_{\rho,\varphi}(\mu,\nu,\mathrm{Diam}_\rho(T)).$$

We use the concept from the Introduction, i.e. let a random t_F valued in a finite $F\subset T$ be such that $\mathbb{E}\max_{t\in F}X(t)=\mathbb{E}X(t_F)$ and $\mu_F(t)=\mathbb{P}(t_F=t)$.

PROPOSITION 2.1. *There exists a universal constant $K<\infty$ such that for any process $X(t)$, $t\in T$, of bounded increments and for every finite set $F\subset T$,*

$$\mathbb{E}\sup_{t\in F}X(t)-\int_T \mathbb{E}X(u)\,\mu_F(du)\le K\mathcal{M}_{\rho,\varphi}(\mu_F,\mu_F).$$

Proof. First Theorem 1.2 from [2] shows that for each $t\in F$,

$$\Bigl|X(t)-\int_T X(u)\,\mu_F(du)\Bigr| \le K_1\sigma_{\mu_F,\rho,\varphi}(t)+K_2\mathcal{M}_{\rho,\varphi}(\mu_F,\mu_F)\int_{T\times T}\varphi\biggl(\frac{|X(u)-X(v)|}{\rho(u,v)}\biggr)\,\nu(du,dv),$$

where K_1,K_2 are absolute constants and ν is a probability measure on $T\times T$. Denote

$$Z=\int_{T\times T}\varphi\biggl(\frac{|X(u)-X(v)|}{\rho(u,v)}\biggr)\,\nu(du,dv);$$

then (2.1) implies that $\mathbb{E}Z\le 1$. Let $\Omega_t=\{t_F=t\}$. Clearly

$$\begin{aligned}\sum_{t\in F}\mathbb{E}1_{\Omega_t}X(t)-\int_T\mathbb{E}X(u)\,\mu_F(dt)&=\sum_{t\in F}\mathbb{E}1_{\Omega_t}\Bigl(X(t)-\int_T X(u)\,\mu_F(du)\Bigr)\\ &\le\sum_{t\in F}\mathbb{E}1_{\Omega_t}\bigl(K_1\sigma_{\mu_F,\rho,\varphi}(t)+K_2\mathcal{M}_{\rho,\varphi}(\mu_F,\mu_F)Z\bigr)\\ &\le K_1\sum_{t\in F}\sigma_{\mu_F,\rho,\varphi}(t)\mu_F(t)+K_2\mathcal{M}(\mu_F,\mu_F)\le(K_1+K_2)\mathcal{M}(\mu_F,\mu_F).\end{aligned}$$

This completes the proof with $K=K_1+K_2$. ■

We recall that in the Gaussian case, i.e. when $\rho(s,t)=d(s,t)$ and $\varphi(x)=2^{x^2}-1$, we simplify the notation and use σ_μ and $\mathcal{M}$ instead of $\sigma_{\mu,\rho,\varphi}$ and $\mathcal{M}_{\rho,\varphi}$. Obviously since Gaussian variables are symmetric, $\mathbb{E}X(u)=0$

and hence Proposition 2.1 implies Theorem 1.4. In the case of non-symmetric processes we have the following bound.

COROLLARY 2.2. *For $X(t)$, $t \in T$, with Gaussian increments, for a finite set $F \subset T$ and for any $s \in T$,*

$$\mathbb{E}\sup_{t\in T}(X(t)-X(s)) \le (K+2)\mathcal{M}(\mu_F,\mu_F).$$

Proof. Clearly

$$\mathbb{E}\sup_{t\in T}(X(t)-X(s)) \le \mathbb{E}\sup_{t\in T}\Big(X(t)-\int_T X(u)\,\mu_F(du)\Big)+\int_T \mathbb{E}(X(u)-X(s))\,\mu_F(du).$$

However due to the convexity of φ,

$$\mathbb{E}\frac{|X(u)-X(s)|}{c}1_{|X(u)-X(s)|\ge c} \le \mathbb{E}\varphi\bigg(\frac{|X(u)-X(s)|}{c}\bigg),$$

and hence

$$\mathbb{E}|X(u)-X(s)| \le 2\|X(u)-X(s)\|_\varphi \le 2\rho(u,v) \le 2\,\mathrm{Diam}_\rho(T).$$

Therefore

$$\int_T \mathbb{E}(X(u)-X(s))\,\mu_F(du) \le 2\int_T \mathrm{Diam}_\rho(T)\,\mu_F(du) \le 2\mathcal{M}(\mu_F,\mu_F).$$

Consequently,

$$\mathbb{E}\sup_{t\in T}(X(t)-X(s)) \le (K+2)\mathcal{M}(\mu_F,\mu_F).$$

The result follows since $\sigma_{\mu_F}(t) \ge \mathrm{Diam}_\rho(T)$ for $t \in F$. ∎

Observe that if μ_{-F} denotes the supremum distribution of $-X(t)$ on F then

$$\mathbb{E}\sup_{t\in T}|X(t)-X(s)| \le (2+K)\big(\mathcal{M}(\mu_F,\mu_F)+\mathcal{M}(\mu_{-F},\mu_{-F})\big). \tag{2.2}$$

3. The partition structure. One of the clear consequences of Gaussian sample boundedness is that $\mathrm{Diam}(T) = \sup\{d(s,t) : s,t \in T\}$ is bounded. For simplicity assume that $\mathrm{Diam}(T) = 1$. Recall that we write σ_μ and $\mathcal{M}$ in this case.

Fix $r > 1$. Let $\mathcal{A} = (\mathcal{A}_k)_{k\ge 0}$ be a partition sequence such that for each $A \in \mathcal{A}_k$ there exists $t_A \in A$ such that $A \subset B(t_A, r^{-k}/2)$. Let $A_k(t)$ be the element of $\mathcal{A}_k$ that contains t. We translate the quantities $\mathcal{M}(\mu,\nu)$ into the language of $\mathcal{A}$.

LEMMA 3.1. *For each μ,*

$$\sigma_\mu(t) \le r \sum_{k=1}^{\infty} r^{-k} \sqrt{\log_2\Big(\frac{\mu(A_{k-1}(t))}{\mu(A_k(t))}\Big)}.$$

Proof. First observe that

$$\int_0^\infty \sqrt{\log_2(\mu(B(t,\varepsilon))^{-1})}\, d\varepsilon \le (r-1) \sum_{k=1}^{\infty} r^{-k} \sqrt{\log_2(\mu(B(t,r^{-k}))^{-1})}.$$

Then note that for all $t \in T$, we have $A_k(t) \subset B(t, r^{-k})$ and therefore

$$\sqrt{\log_2(\mu(B(t,r^{-k}))^{-1})} \le \sqrt{\log_2(\mu(A_k(t))^{-1})}.$$

By the property $\sqrt{\log_2(xy)} \le \sqrt{\log_2(x)} + \sqrt{\log_2(y)}$ we obtain

$$\sqrt{\log_2(\mu(A_k(t))^{-1})} \le \sum_{l=1}^{k} \sqrt{\log_2\Big(\frac{\mu(A_{l-1}(t))}{\mu(A_l(t))}\Big)}.$$

Therefore changing the summation order yields

$$\sum_{k=1}^{\infty} r^{-k} \sqrt{\log_2(\mu(B(t,r^{-k}))^{-1})} \le \sum_{k=1}^{\infty} r^{-k} \sum_{l=1}^{k} \sqrt{\log_2\Big(\frac{\mu(A_{l-1}(t))}{\mu(A_l(t))}\Big)}$$
$$= \sum_{l=1}^{\infty} \Big(\sum_{k=l}^{\infty} r^{-k}\Big) \sqrt{\log_2\Big(\frac{\mu(A_{l-1}(t))}{\mu(A_l(t))}\Big)} = \frac{r}{r-1} \sum_{l=1}^{\infty} r^{-l} \sqrt{\log_2\Big(\frac{\mu(A_{l-1}(t))}{\mu(A_l(t))}\Big)}.$$

This completes the proof. ■

COROLLARY 3.2. *We have*

$$\mathcal{M}(\mu,\nu) \le r \sum_{k=1}^{\infty} r^{-k} \sum_{B \in \mathcal{A}_{k-1}} \sum_{A \in \mathcal{A}_k(B)} \nu(A) \sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A)}\Big)}.$$

4. Gaussian tools. In the general theory of Gaussian processes there are two basic properties one can use (see [11, Theorem 3.18], and [10] for concentration inequalities).

LEMMA 4.1 (Sudakov minorization). *For a Gaussian $X(t)$, $t \in T$, suppose that $d(t_i, t_j) \ge a$ for $i, j \le m$, $i \ne j$. Then*

$$\mathbb{E} \sup_{1 \le i \le m} X(t_i) \ge C_1^{-1} a \sqrt{\log_2(m)},$$

where C_1 is a universal constant.

LEMMA 4.2 (Gaussian concentration). *For a Gaussian* $X(t)$, $t \in T$, *let* $\sigma = \sup_{s,t\in D}(\mathbb{E}(X(t)-X(s))^2)^{1/2}$, $D \subset T$. *Then*

$$\mathbb{P}\Big(\Big|\sup_{t\in D}\Big(X(t) - \mathbb{E}\sup_{t\in D} X(t)\Big)\Big| \geq u\Big) \leq 2\exp\Big(-\frac{u^2}{2\sigma^2}\Big).$$

The main consequence of these facts is the basic tool we use (see [17, Proposition 2.1.4]).

PROPOSITION 4.3. *Let* $(t_i)_{i=1}^m \subset T$ *satisfy* $d(t_i,t_j) \geq a$ *if* $i \neq j$. *Consider* $\sigma > 0$ *such that* $D_i \subset B(t_i,\sigma)$. *If* $\bigcup_{i=1}^m D_i \subset D$ *then*

$$\mathbb{E}\sup_{t\in D} X(t) \geq C_1^{-1} a\sqrt{\log_2(m)} - C_2\sigma\sqrt{\log_2(m)} + \min_{1\leq i\leq m} \mathbb{E}\sup_{t\in D_i} X(t).$$

Thus for $a \geq (2C_1C_2)\sigma$,

$$\mathbb{E}\sup_{t\in D} X(t) \geq C_3^{-1} a\sqrt{\log_2(m)} + \min_{1\leq i\leq m} \mathbb{E}\sup_{t\in D_i} X(t),$$

where C_1, C_2, C_3 *are universal constants.*

5. The lower bound. In this section for a Gaussian $X(t)$, $t \in T$, we prove Theorem 1.5. Recall that $\mathrm{Diam}(T) = 1$. First define a set functional by

$$F(A) = \mathbb{E}\sup_{t\in A} X(t), \qquad A \text{ is a Borel subset of } T.$$

Using this functional we define a natural partitioning structure for (T,d). Recall that $F(T)$ can be finite only if (T,d) is totally bounded, and hence using compactification we may refer to (T,d) as a compact space.

Fix $r > 1$ and $\varepsilon > 0$. We construct $\mathcal{A} = (\mathcal{A}_k)_{k\geq 0}$ in the following way. Let $\mathcal{A}_0 = \{T\}$. To define $\mathcal{A}_k$, $k \geq 1$, we partition each $B \in \mathcal{A}_{k-1}$ into sets $A_1, \ldots, A_M$ in the following way. Let $B_0 = B$ and $t_1 \in B$ be such that

$$\sup_{s\in B_0} F(C(s)) \leq F(C(t_1)) + \varepsilon r^{-k},$$

where $C(s) = B(s, r^{-k-1}/2) \cap B_0$. Let $A_1 = B(t_1, r^{-k}/2) \cap B_0$ and $B_1 = B_0 \setminus A_1$. We continue the construction, and if for $i \geq 1$, $B_{i-1} \neq \emptyset$ then choose $t_i \in B_{i-1}$ so that

$$\sup_{s\in B_{i-1}} F(C(s)) \leq F(C(t_i)) + \varepsilon r^{-k}, \tag{5.1}$$

where $C(s) = B(s, r^{-k-1}/2) \cap B_{i-1}$. Using t_i we construct $A_i = B(t_i, r^{-k}/2) \cap B_{i-1}$ and $B_i = B_{i-1} \setminus A_i$. Recall that our basic assumption was that there exists $M < \infty$ such that $B_M = \emptyset$, namely by construction $M \leq N(T,d,r^{-k}/2) < \infty$ due to the compactness of T.

For each $B \in \mathcal{A}_k$ and $l \geq k$ denote $\mathcal{A}_l(B) = \{A \in \mathcal{A}_l : A \subset B\}$. Note that by construction for each $A \in \mathcal{A}_k$ there exists $t_A \in A$ such that $A \subset B(t_A, r^{-k}/2)$, so the partition satisfies the requirement from Section 3.

The main result of this section is the following induction scheme. The result is based on Proposition 4.3 and hence we have to choose r suitably large. On the other hand, a parameter ε is introduced to avoid some technical problems with attaining a supremum. This parameter is not particularly important but again we have to make it sufficiently small in order to apply Proposition 4.3 properly.

PROPOSITION 5.1. *For r sufficiently large, namely $r > \max\{1, 2C_1C_2\}$, and $\varepsilon > 0$ sufficiently small, namely $\varepsilon < (4C_3)^{-1}$, there exists a universal constant $L < \infty$, namely $L = 4C_3$, such that for each measure μ on T and $B \in \mathcal{A}_{k-1}$, $k \geq 1$,*

$$\mu(B)(F(B) + 4r^{-k}) \geq \frac{1}{2L} r^{-k} \sum_{A \in \mathcal{A}_k(B)} \mu(A) \sqrt{\log_2\left(\frac{\mu(B)}{\mu(A)}\right)} + \sum_{C \in \mathcal{A}_{k+1}(B)} \mu(C)F(C).$$

Proof. Fix $B \in \mathcal{A}_{k-1}$, $k \geq 1$. By the above construction $\mathcal{A}_k(B) = \{A_1, \ldots, A_M\}$. There exists the smallest $l_0 \geq 0$ such that $1 \leq M \leq 2^{2^{l_0}}$. For simplicity let $m_{-1} = 0$ and $m_l = 2^{2^l}$ for $l = 0, 1, \ldots, l_0$. We group sets in $\mathcal{A}_k(B)$ using the following scheme. Let $\mathcal{A}_{k,l}(B) = \{A_{m_{l-1}+1}, \ldots, A_{m_l}\}$ for $0 \leq l < l_0$, and $\mathcal{A}_{k,l_0}(B) = \{A_{m_{l_0-1}+1}, \ldots, A_M\}$. Clearly $|\mathcal{A}_{k,l}(B)| = m_l - m_{l-1}$ for $0 \leq l < l_0$, and $|\mathcal{A}_{k,l_0}(B)| = M - m_{l_0-1}$. For simplicity denote $B_l = \bigcup_{A_j \in \mathcal{A}_{k,l}(B)} A_j$, $0 \leq l \leq l_0$.

By the partition construction there exist points t_i, $1 \leq i \leq M$, such that $A_i \subset B(t_i, r^{-k}/2)$ and $d(t_i, t_j) \geq r^{-k}/2$ if $1 \leq i < j \leq M$. Moreover for each $C \in \mathcal{A}_{k+1}(A_i)$ there exists $t_C \in C$ such that $C \subset B(t_C, r^{-k-1}/2) \cap A_i$, and hence by (5.1),

$$F(C) \leq F(D_i) + \varepsilon r^{-k}$$

where

$$D_i = B(t_i, r^{-k-1}/2) \cap A_i \quad \text{for } 1 \leq i \leq M.$$

Again by the partition construction, if $i \leq j$ then

$$F(D_j) \leq F(D_i) + \varepsilon r^{-k}.$$

Fix $l \geq 1$. We apply Proposition 4.3 with $a = r^{-k}/2$, $\sigma = r^{-k-1}/2$ and $m = m_{l-1} + 1$ for the sets D_i, $1 \leq i \leq m$, and deduce that for $r > \max\{1, 2C_1C_2\}$ and $\varepsilon < (4C_3)^{-1}$ there exists a universal constant $L = 4C_3$ such that

$$F(B) \geq \frac{1}{L} r^{-k} 2^{l/2} + F(C) \quad \text{for all } C \in \mathcal{A}_{k+1}(A_j),\ A_j \in \mathcal{A}_{k,l}(B).$$

Consequently,

$$\mu(B_l)F(B) \geq \frac{1}{L}\mu(B_l)r^{-k}2^{l/2} + \sum_{A_j\in\mathcal{A}_{k,l}(B)}\sum_{C\in\mathcal{A}_{k+1}(A_j)} \mu(C)F(C). \tag{5.2}$$

The remaining bound concerns $\mathcal{A}_{k,0}(B)$. Here we cannot do better than the simplest estimate

$$\mu(B_0)F(B) \geq \sum_{A_j\in\mathcal{A}_{k,0}(B)}\sum_{C\in\mathcal{A}_{k+1}(A_j)} \mu(C)F(C). \tag{5.3}$$

By the concavity of $\sqrt{\log_2 x}$ on $[1,\infty)$ we find that for $0 \leq l \leq l_0$,

$$\mu(B_l)2^{l/2} \geq \sum_{A_j\in\mathcal{A}_{k,l}(B)} \mu(A_j)\sqrt{\log_2\Big(\frac{\mu(B_l)}{\mu(A_j)}\Big)}.$$

Moreover for each $0 \leq l \leq l_0$ and $A_j \in \mathcal{A}_{k,l}(B)$,

$$\sqrt{\log_2\Big(\frac{\mu(B_l)}{\mu(A_j)}\Big)} + \sqrt{\log_2\Big(\frac{\mu(B)}{\mu(B_l)}\Big)} \geq \sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A_j)}\Big)},$$

and hence

$$\begin{aligned}\mu(B_l)\bigg(2^{l/2} + \sqrt{\log_2\Big(\frac{\mu(B)}{\mu(B_l)}\Big)}\bigg) \\ \geq \sum_{A_j\in\mathcal{A}_{k,l}(B)} \mu(A_j)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A_j)}\Big)}.\end{aligned} \tag{5.4}$$

Thus if $2^{l/2} \geq \sqrt{\log_2(\mu(B)/\mu(B_l))}$ then by (5.4),

$$2\mu(B_l)2^{l/2} \geq \sum_{A_j\in\mathcal{A}_{k,l}(B)} \mu(A_j)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A_j)}\Big)};$$

otherwise $2^{l/2} \leq \sqrt{\log_2(\mu(B)/\mu(B_l))}$, which together with the fact that $x\sqrt{\log_2(1+x^{-1})}$ increases on $[0,1]$ implies

$$\mu(B_l)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(B_l)}\Big)} \leq \frac{2^{l/2}+1}{2^{2^l}}\mu(B).$$

Therefore by (5.4) we obtain

$$\sum_{l=0}^{l_0} 2^{l/2}\mu(B_l) + \sum_{l=0}^{l_0}\frac{2^{l/2}+1}{2^{2^l}}\mu(B) \geq 2^{-1}\sum_{i=1}^{M}\mu(A_i)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A_i)}\Big)}. \tag{5.5}$$

Summing (5.2), (5.3) and (5.5) yields

$$\mu(B)\Big(F(B)+r^{-k}\Big(1+\sum_{l=0}^{l_0}\frac{2^{l/2}+1}{2^{2^l}}\Big)\Big)$$
$$\geq \frac{1}{2L}r^{-k}\sum_{i=1}^{M}\mu(A_i)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A_i)}\Big)}+\sum_{C\in\mathcal{A}_{k+1}(B)}\mu(C)F_{k+1}(C).$$

Clearly $1+\sum_{l=0}^{l_0}\frac{2^{l/2}+1}{2^{2^l}}\leq 4$, which completes the proof. ∎

Proposition 5.1 and a simple induction yield

$$F(T)+4\sum_{k=1}^{\infty}r^{-2k+2}$$
$$\geq \frac{1}{2L}\sum_{k=1}^{\infty}r^{-2k+1}\sum_{B\in\mathcal{A}_{2(k-1)}}\sum_{A\in\mathcal{A}_{2k-1}(B)}\mu(A)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A)}\Big)}.$$

Note that for each $C\in\mathcal{A}_1$, the partition sequence $\mathcal{A}$ defines $\mathcal{C}=(\mathcal{C}_k)_{k\geq 0}$ by $\mathcal{C}_k=\mathcal{A}_{k+1}(C)$. Applying the above inequality to C and $\mathcal{C}$ in place of T and $\mathcal{A}$, and then using the inequality $F(T)\geq\sum_{C\in\mathcal{A}_1}\mu(C)F(C)$ we deduce that

$$F(T)+4\sum_{k=1}^{\infty}r^{-2k+1}\geq\frac{1}{2L}\sum_{k=1}^{\infty}r^{-2k}\sum_{B\in\mathcal{A}_{2k-1}}\sum_{A\in\mathcal{A}_{2k}(B)}\mu(A)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A)}\Big)}.$$

Since $F(T)=\mathbb{E}\sup_{t\in T}X(t)$ and $\sum_{k=1}^{\infty}r^{-k}\leq 1$ for $r\geq 2$, we finally get

$$2\Big(\mathbb{E}\sup_{t\in T}X(t)+4\Big)\geq\frac{1}{2L}\sum_{k=1}^{\infty}\sum_{B\in\mathcal{A}_{k-1}}\sum_{A\in\mathcal{A}_k(B)}\mu(B)\sqrt{\log_2\Big(\frac{\mu(B)}{\mu(A)}\Big)}.$$

Together with Corollary 3.2 and the inequality

$$\mathbb{E}\sup_{t\in T}X(t)=\mathbb{E}\sup_{t\in T}X(t)-X(s)\geq\sup_{t\in T}\mathbb{E}\max(X(t)-X(s),0)$$
$$\geq C\,\mathrm{Diam}(T)=C,$$

where C is an absolute constant, this completes the proof of Theorem 1.5.

6. Continuity of the process. In this section we prove Theorem 1.7, i.e. we estimate

$$\mathcal{S}(\delta)=\mathbb{E}\sup_{s,t\in T,\,d(s,t)\leq\delta}|X(t)-X(s)|$$

in terms of $\sup_\mu\mathcal{M}(\mu,\mu,\delta)$. Fix $0<\delta\leq\mathrm{Diam}(T)=1$ and let $\mathcal{A}$ be a partition of (T,d) that satisfies $A\subset B(t_A,\delta)$ for each $A\in\mathcal{A}$ and some $t_A\in A$. We require that $|\mathcal{A}|=N(T,d,\delta)$, which is clearly possible by the entropy definition. Obviously

$$\{(s,t) : d(s,t) \le \delta\} \supset \bigcup_{A\in\mathcal{A}} A \times \{t_A\}.$$

Therefore

$$\mathcal{S}(\delta) = \mathbb{E} \sup_{s,t\in T, d(s,t)\le\delta} |X(t) - X(s)| \ge \mathbb{E} \max_{A\in\mathcal{A}_k} \sup_{t\in A} |X(t) - X(t_A)| \tag{6.1}$$
$$\ge \sup_{A\in\mathcal{A}_k} \mathbb{E} \sup_{t\in A} (X(t) - X(t_A)) = \sup_{A\in\mathcal{A}_k} \mathbb{E} \sup_{t\in A} X(t).$$

Using Theorem 1.5 we get

$$\mathbb{E} \sup_{t\in A} X(t) \ge K^{-1} \sup_{\mu_A} \mathcal{M}(\mu_A, \mu_A), \tag{6.2}$$

where the supremum is taken over all measures supported on A. Observe that each probability measure μ on T has the unique representation $\mu = \sum_{A\in\mathcal{A}} \alpha(A)\mu_A$, where $\alpha(A) \ge 0$, $\sum_{A\in\mathcal{A}} \alpha(A) = 1$ and μ_A is supported on A. Consequently, by the property $\sqrt{\log_2(xy)} \le \sqrt{\log_2(x)} + \sqrt{\log_2(y)}$,

$$\mathcal{M}(\mu, \mu, c) = \int_T\int_0^c \sqrt{\log_2(\mu(B(t,\varepsilon))^{-1})}\, d\varepsilon\, \mu(dt)$$
$$\le \sum_{A\in\mathcal{A}} \alpha(A) \int_T\int_0^c \sqrt{\log_2((\alpha(A)\mu_A(B(t,\varepsilon)))^{-1})}\, d\varepsilon\, \mu_A(dt)$$
$$\le \sum_{A\in\mathcal{A}} \Big[\alpha(A) \int_T \int_0^\infty \sqrt{\log_2((\alpha(A)\mu_A(B(t,\varepsilon)))^{-1})}\, d\varepsilon\, \mu_A(dt)$$
$$+ c\alpha(A)\sqrt{\log_2(\alpha(A)^{-1})}\Big].$$

Using the entropy property

$$\sum_{A\in\mathcal{A}} \alpha(A)\sqrt{\log_2(\alpha(A)^{-1})} \le \sqrt{\log_2(N(T,d,\delta))}$$

we deduce that

$$\sup_\mu \mathcal{M}(\mu,\mu,c) \le \sum_{A\in\mathcal{A}} \alpha(A) \sup_{\mu_A} \mathcal{M}(\mu_A,\mu_A) + c\sqrt{\log_2(N(T,d,\delta))}.$$

Consequently, by (6.1) and (6.2),

$$\mathcal{S}(\delta) \ge K^{-1} \sup_{c>0} \Big(\sup_\mu \mathcal{M}(\mu,\mu,c) - c\sqrt{\log_2(N(T,d,\delta))}\Big). \tag{6.3}$$

This is the lower bound in Theorem 1.7.

Corollary 6.1. *If $X(t)$, $t \in T$, is Gaussian and continuous then*

$$\lim_{\delta\to 0} \delta\sqrt{\log_2(N(T,d,\delta))} = 0.$$

Proof. By our main characterization, the process $X(t)$, $t \in T$, is Gaussian and continuous, so $\lim_{\delta\to 0}\mathcal{S}(\delta) = 0$. Consider now a set $F \subset T$ such that $|F| = N(T,d,\delta/2)$ and $\bigcup_{t\in F} B(t,\delta/2) = T$. It is important to observe that F may be chosen in a such way that points in F are $\delta/2$-separated (this can be obtained from any set F with the required properties by a simple induction procedure). We apply (6.3) with $c = \delta/2$ and μ equally distributed on F, i.e.

$$\mu(B(t,c)) = \mu(B(t,\delta/2)) = \mu(\{t\}) = 1/N(T,d,\delta/2).$$

Therefore $\mathcal{M}(\mu,\mu,c) = \mathcal{M}(\mu,\mu,\delta/2) = (\delta/2)\sqrt{\log_2(N(T,d,\delta/2))}$ and hence

$$\mathcal{S}(\delta) \geq (\delta/2)\big(\sqrt{\log_2(N(T,d,\delta/2))} - \sqrt{\log_2(N(T,d,\delta))}\big).$$

Using $\delta = 2^{-k}$ this proves that

$$\lim_{k\to\infty} 2^{-k-1}\big(\sqrt{\log_2(N(T,d,2^{-k-1}))} - \sqrt{\log_2(N(T,d,2^{-k}))}\big) = 0. \tag{6.4}$$

To finish the argument we observe that by Slepian's lemma [11, Theorem 3.18] we have

$$\sup_{k\geq 0} 2^{-k}\sqrt{\log_2(N(T,d,2^{-k}))} \leq K\mathcal{S} < \infty, \tag{6.5}$$

where $\mathcal{S} = \lim_{\delta\to\infty}\mathcal{S}(\delta)$. For simplicity, denote $N(T,d,2^{-k})$ by a_k for any $k \geq 0$. We have $a_k \leq a_{k+1}$ for $k \geq 0$; moreover, (6.4) can be rewritten as $\lim_{k\to\infty} 2^{-k-1}(a_{k+1} - a_k)$ and (6.5) as $\sup_{k\geq 0} 2^{-k}a_k < K\mathcal{S}$. Suppose there is a subsequence $(k_l)_{l\geq 0}$ such that $\lim_{l\to\infty} 2^{-k_l}a_{k_l} = a > 0$. Then obviously by (6.4), for any $m \geq 0$,

$$\lim_{l\to\infty} 2^{-k_l+m}a_{k_l-m} = 2^m a.$$

Consequently, for large enough m we have a contradiction with the requirement that $\sup_{k\geq 0} 2^{-k}a_k < K\mathcal{S}$. This implies that the only possible limit for subsequences is zero and hence $\lim_{k\to\infty} 2^{-k}a_k = 0$. ∎

Corollary 6.1 and (6.3) imply that if $X(t)$, $t \in T$, is Gaussian and continuous then $\lim_{\delta\to 0}\sup_\mu \mathcal{M}(\mu,\mu,\delta) = 0$.

On the other hand, if F is an $N(T,d,\delta)$-net (i.e. $|F| = N(T,d,\delta)$, $\bigcup_{t\in F} B(t,\delta) = T$) then

$$\mathbb{E}\sup_{s,t\in T}|X(t) - X(s)| \leq \mathbb{E}\sup_{s\in F}\sup_{t\in B(t,2\delta)}|X(t) - X(s)|.$$

By the concentration of measure argument based on Lemma 4.2 we deduce that for a universal $K_1 < \infty$,

$$\mathbb{E}\sup_{s\in F}\sup_{t\in B(t,2\delta)}|X(t) - X(s)| \leq \sup_{s\in F}\mathbb{E}\sup_{t\in B(s,2\delta)}|X(t) - X(s)| + K_1\delta\sqrt{\log_2(N(T,d,\delta))}.$$

Note that

$$\mathbb{E} \sup_{t\in B(s,2\delta)} |X(t)-X(s)| = 2\mathbb{E} \sup_{t\in B(s,2\delta)} X(t)$$

and $\mathbb{E}\sup_{t\in B(s,2\delta)} X(t) \le K_2 \sup_\mu \mathcal{M}(\mu,\mu,2\delta)$ by Theorem 1.5. Hence

$$\mathbb{E}\sup_{s\in F} \sup_{t\in B(t,2\delta)} |X(t)-X(s)| \le K\Big(\sup_\mu \mathcal{M}(\mu,\mu,2\delta) + \delta\sqrt{\log_2(N(T,d,\delta))}\Big).$$

Since the argument used in the proof of Corollary 6.1 implies that

$$\sup_\mu \mathcal{M}(\mu,\mu,2\delta) \ge \delta\sqrt{\log_2(N(T,d,\delta))},$$

we obtain the upper bound in Theorem 1.7. Then $\lim_\delta \sup_\mu \mathcal{M}(\mu,\mu,\delta)=0$ implies that $\lim_{\delta\to 0}\mathcal{S}(\delta)=0$, and therefore the process $X(t)$, $t\in T$, is continuous.

We now prove Theorem 1.6. Assuming the continuity of $X(t)$, $t\in T$, we construct μ_T on T such that $K^{-1}\mathcal{M}(\mu_T,\mu_T)\le \mathbb{E}\sup_{t\in T} X(t)\le K\mathcal{M}(\mu_T,\mu_T)$. Let $(F_n)_{n=0}^\infty$ be any sequence of finite subsets such that $F_n \subset F_{n+1}$ and $\bigcup_{n\ge 0} F_n$ is dense in T. By the compactness of (T,d) the set of cluster points of $(\mu_{F_n})_{n=0}^\infty$ is not empty, and hence going to a subsequence we can assume that μ_T is a weak limit of the sequence. By Theorems 1.5 and 1.4,

$$K^{-1}\mathcal{M}(\mu_T,\mu_T) \le \mathbb{E}\sup_{t\in T} X(t) \le K \limsup_{n\to\infty} \mathcal{M}(\mu_{F_n},\mu_{F_n}),$$

so it suffices to show that

$$\lim_{n\to\infty} \mathcal{M}(\mu_{F_n},\mu_{F_n}) = \mathcal{M}(\mu_T,\mu_T).$$

It is clear that for $\varepsilon>0$ the functionals

$$\Phi_\varepsilon(\nu) = \int_T \int_\varepsilon^\infty \sqrt{\log_2(\nu(B(t,\varepsilon))^{-1}}\, d\varepsilon\, \nu(dt)$$

are continuous on $\mathcal{P}(T,d)$ (space of probability measures with the weak topology). Therefore to get the convergence of $\Phi_0(\mu_{F_n})$ to $\Phi_0(\mu_T)$ we need that

$$\sup_n \int_T\int_0^\varepsilon \sqrt{\log_2(\mu_{F_n}(B(t,\varepsilon))^{-1})}\, \mu_{F_n}(dt) \le \sup_\mu \mathcal{M}(\mu,\mu,\varepsilon)$$

tends to 0 as $\varepsilon\to 0$. Theorem 1.7 implies that the convergence holds whenever $X(t)$, $t\in T$, is continuous. This completes the proof of Theorem 1.6.

7. Hilbert–Schmidt ellipsoid. We are ready to discuss the supremum distribution in the setting of the basic toy example for the theory, the Hilbert–Schmidt ellipsoid. Consider the real sequences $x=(x_i)_{i=1}^\infty$ equipped with the Euclidean norm $\|x\| = (\sum_{i=1}^\infty x_i^2)^{1/2}$. The basic object for our purposes is the ℓ_2 space which consists of $x=(x_i)_{i=1}^\infty$ such that $\|x\|<\infty$. We use the notation $x^2 = (x_i^2)_{i=1}^\infty$ and $xy = (x_i y_i)_{i=1}^\infty$ for any sequences $x = (x_i)_{i=1}^\infty$, $y = (y_i)_{i=1}^\infty$. For any sequence $a = (a_i)_{i=1}^\infty$ we set

$$\mathcal{E} = \Big\{ (x_i)_{i=1}^\infty \in \ell_2 : \sum_{i=1}^\infty x_i^2/a_i^2 \le 1 \Big\}.$$

We can require that $a_i \ge a_{i+1} > 0$ for $i \ge 1$. Note that $\mathcal{E}$ is compact if and only if $a_i \to 0$ as $i \to \infty$. Let $g = (g_i)_{i=1}^\infty$, where the g_i are independent standard Gaussian random variables. Let

$$X(x) = \langle x, g \rangle, \qquad x \in \mathcal{E}.$$

The basic question is when $\sup_{x\in\mathcal{E}} X(x) < \infty$ a.s. Note that the process is continuous if it is sample bounded, therefore sample boundedness implies the existence of the supremum distribution (in the sense of the previous section). In the case of $\mathcal{E}$ this implies that the process X is sample bounded if and only if $\sum_{i=1}^\infty a_i^2 < \infty$. Indeed, for any $N < \infty$ define

$$\mathcal{E}_N = \Big\{ (x_i)_{i=1}^\infty : \sum_{i=1}^N x_i^2/a_i^2 \le 1 \text{ and } x_i = 0 \text{ for } i > N \Big\}.$$

Using the Schwarz inequality, we get

$$\mathbb{E} \sup_{x\in\mathcal{E}_N} X(x) = \mathbb{E} \sup_{x\in\mathcal{E}_N} \langle x, g\rangle = \mathbb{E} \sup_{x\in\mathcal{E}_N} \sum_{i=1}^N \frac{x_i}{a_i}(a_i g_i) = \mathbb{E}\Big(\sum_{i=1}^N a_i^2 g_i^2\Big)^{1/2}. \tag{7.1}$$

Note that the supremum $\sup_{x\in\mathcal{E}_N} X(x)$ is attained at $x \in \mathcal{E}_N$ such that $x_i = a_i^2 g_i / (\sum_{i=1}^N a_i^2 g_i^2)^{1/2}$ for $i \le N$ and $x_i = 0$ for $i > N$. Therefore the supremum distribution μ_N on $\mathcal{E}_N$ is the law of $a_i^2 g_i/(\sum_{i=1}^N a_i^2 g_i^2)^{1/2}$ for $i \le N$ and 0 for $i > N$. If $\sum_{i=1}^\infty a_i^2 = \infty$ then the weak limit of μ_N is δ_0 and we have a contradiction with Theorem 1.6. This implies that $\mathbb{E}\sup_{x\in\mathcal{E}} X(x)$ can be finite only if $\|a\|^2 = \sum_{i=1}^\infty a_i^2 < \infty$. Note that in this case the limit of μ_N exists and is equal to μ, the distribution of $a_i^2 g_i/(\sum_{i=1}^\infty a_i^2 g_i^2)^{1/2}$, $i \ge 1$, i.e. the law of the random variable $a^2 g/\|ag\|$ valued in $\mathcal{E}$. In particular under the assumption $\|a\| < \infty$, we have

$$\mathbb{E} \sup_{x\in\mathcal{E}} X(x) = \mathbb{E}\Big(\sum_{i=1}^\infty a_i^2 g_i^2\Big)^{1/2} \le \Big(\sum_{i=1}^\infty a_i^2\Big)^{1/2} = \|a\|. \tag{7.2}$$

It can be proved using the above result and Khinchin's inequality that $\mathbb{E}\sup_{x\in\mathcal{E}} X(x)$ is comparable with $\|a\|$ up to a universal constant. Recall that by Theorems 1.5 and 1.6, $\mathbb{E}\sup_{x\in\mathcal{E}} X(x)$ is comparable with $\mathcal{M}(\mu,\mu)$, and hence $\mathcal{M}(\mu,\mu)$ is comparable with $\|a\|$ up to a universal constant. The measure μ is a bit complicated and that is why we replace it by a more comprehensible distribution based on random signs which shares the same property.

First observe that by Theorem 1.5 and (7.2) the quantity $\|a\|$ dominates $\mathcal{M}(\nu,\nu)$ up to a universal constant, where ν is any probability distribution on $\mathcal{E}$. Following this idea let us slightly simplify the setting and consider the process

$$Y(x) = \langle x, \varepsilon\rangle, \qquad x \in \mathcal{E},$$

where $\varepsilon = (\varepsilon_i)_{i=1}^{\infty}$ is a sequence of independent Bernoulli random variables, i.e. $\mathbb{P}(\varepsilon_i = \pm 1) = 1/2$. Clearly $Y(x)$, $x \in \mathcal{E}$, is subgaussian, which means that

$$\|Y(x) - Y(y)\|_{\varphi} \le C\|x-y\|, \quad \text{where} \quad \varphi(x) = 2^{2^x} - 1,$$

and C is a universal constant. Obviously $\mathbb{E}\sup_{x\in\mathcal{E}}\langle x,\varepsilon\rangle = \|a\|$. Therefore by Proposition 2.1, $\|a\|$ is bounded from above by $\mathcal{M}(\nu,\nu)$, where ν is the supremum distribution of $Y(x)$, $x \in \mathcal{E}$. Note that by the same argument as in the Gaussian case the measure ν is the distribution of the random variable $a^2\varepsilon/\|a\|$ valued in $\mathcal{E}$. On the other hand we have learned from the Gaussian case that $\|a\|$ dominates $\mathbb{E}\sup_{x\in\mathcal{E}} X(x)$ and hence also $\mathcal{M}(\nu,\nu)$ up to a universal constant. Therefore $\mathcal{M}(\nu,\nu)$ and $\|a\|$ are comparable, which shows that ν is a good equivalent of μ, and the benefit is that ν has a much simpler structure.

We start to analyze the measure ν. Note that for any $x \in \mathcal{E}$, we have

$$\nu(B(x,\delta)) = \mathbb{P}\big(\big\|a^2\varepsilon - x\|a\|\big\| \le \delta\|a\|\big).$$

The upper bound for this quantity is relatively easy to find, by means of the following construction. We may only consider $\delta \le \|x\|$. For any $y \in \ell_2$ define $y(i) \in \ell_2$ by $y(i)_j = 0$, $j < i$ and $y(i)_j = y_j$ for $j \ge i$. Denote $v_i = \|x(i)\|$; then by the construction $v_1 = \|x\|$, $v_{i+1} \le v_i$ for $i > 1$ and $\lim_{i\to\infty} v_i = 0$, therefore $(v_i)_{i=1}^{\infty}$ forms a partition of $[0, \|x\|]$. For simplicity let $v_0 = a_1 \ge v_1$.

Lemma 7.1. *For each $i \ge 0$ and $\delta > 0$ such that $\delta \le \frac{1}{\sqrt{2}} v_i$,*

$$\nu(B(x,\delta)) \le \exp\Big(-\frac{\|a^2(i)\|^2}{8a_i^4}\Big).$$

Proof. First observe that for $i \ge 0$ and $\delta \le \frac{1}{\sqrt{2}} v_i$, we have

$$\begin{aligned}
\mathbb{P}\big(\big\|a^2\varepsilon - x\|a\|\big\| \le \delta\|a\|\big) &\le \mathbb{P}\big(\big\|a^2(i)\varepsilon - x(i)\|a\|\big\| \le \delta\|a\|\big)\\
&\le \mathbb{P}\big(2\langle a^2(i)\varepsilon, x(i)\rangle\|a\| \ge \|a^2(i)\|^2 + (\|x(i)\|^2 - \delta^2)\|a\|^2\big)\\
&\le \mathbb{P}\big(2\langle a^2(i)\varepsilon, x(i)\rangle\|a\| \ge \|a^2(i)\|^2 + \tfrac{1}{2}\|x(i)\|^2\|a\|^2\big)\\
&\le \mathbb{P}\big(2\langle a^2(i)\varepsilon, x(i)\rangle \ge \|a^2(i)\|\,\|x(i)\|\big) \le \exp\Big(-\frac{\|a^2(i)\|^2\|x(i)\|^2}{8\|a^2(i)x(i)\|^2}\Big)\\
&\le \exp\Big(-\frac{\|a^2(i)\|^2}{8a_i^4}\Big),
\end{aligned}$$

where we have used Hoeffding's inequality and $\|x(i)\|^2-\delta^2 = v_i^2-\delta$, $s^2 \geq \frac{1}{2}v_i^2 = \frac{1}{2}\|x(i)\|^2$ and $u^2+w^2 \geq 2uw$. ∎

Consequently,

$$\sqrt{\log_2(\nu(B(x,\delta))^{-1})} \geq \sqrt{\log_2(e)}\,\frac{\|a^2(i)\|}{2\sqrt{2}\,a_i^2} \quad \text{for } \frac{v_{i+1}}{\sqrt{2}} \leq \delta \leq \frac{v_i}{\sqrt{2}}.$$

Therefore

$$\int_{\mathcal{E}}\int_0^{\|x\|} \sqrt{\log(\nu(B(x,\delta))^{-1})}\,d\delta\,\nu(dx) \tag{7.3}$$
$$\geq \sum_{i=1}^{\infty}\int_T \int_{v_{i+1}/\sqrt{2}}^{v_i/\sqrt{2}} \sqrt{\log(\nu(B(x,\delta))^{-1})}\,d\delta\,\nu(dx)$$
$$\geq \sqrt{\log_2(e)}\sum_{i=1}^{\infty}\frac{\|a^2(i)\|}{4a_i^2}\int_T (\|x(i)\|-\|x(i+1)\|)\,\nu(dx).$$

Therefore it suffices to prove the right lower bound on $\int_{\mathcal{E}}(\|x(i)\| - \|x(i+1)\|)\,\nu(dx)$ for $i \geq 0$.

LEMMA 7.2. *We have*

$$\int_{\mathcal{E}} (\|x(i)\| - \|x(i+1)\|)\,\nu(dx) \geq \frac{a_i^4}{2\|a\|\,\|a^2(i)\|}.$$

Proof. First note that

$$\|x(i)\| - \|x(i+1)\| \geq \frac{x_i^2}{2\|x(i)\|}$$

and then observe that

$$\int_{\mathcal{E}} \frac{x_i^2}{2\|x(i)\|}\,\nu(dx) = \mathbb{E}\frac{(\varepsilon_i a_i^2)^2}{2\|a\|\,\|a^2(i)\varepsilon\|} = \frac{a_i^4}{2\|a^2(i)\|\,\|a\|}. \;∎$$

Combining Lemmas 1, 2 and (7.3) implies that

$$\int_{\mathcal{E}}\int_0^{\|x\|} \sqrt{\log_2(\nu(B(x,\delta))^{-1})}\,d\delta\,\nu(dx)$$
$$\geq \sqrt{\log_2(e)}\sum_{i=1}^{\infty}\frac{\|a^2(i)\|}{4a_i^2}\,\frac{a_i^4}{2\|a^2(i)\|^2\|a\|} = \frac{\sqrt{\log_2(e)}}{8}\|a\|.$$

This proves that our upper bound on $\nu(B(x,\delta))$ is of the right order.

We turn to the proof of a suitable lower bound for $\nu(B(x,\delta))$.

LEMMA 7.3. *We have*

$$\int_{\mathcal{E}}\int_0^\infty \sqrt{\log_2(\nu(B(x,\delta))^{-1})}\,d\delta \le 2\sqrt{2}\,\|a\|.$$

Proof. The main trick here is to consider only special $a^2(i)$, namely for $i = 2^k$, $k = 0,1,2,\ldots$. Observe that for points x of the form $a^2\bar{\varepsilon}/\|a\| = (a_i^2\bar{\varepsilon}_i/\|a\|)_{i=1}^\infty$, $\bar{\varepsilon}_i = \pm 1$, on which the measure ν is supported we have

$$\nu(B(x,\delta)) = \mathbb{P}\big(\langle a^2\varepsilon, a^2\rangle \ge \|a^2\|^2 - \tfrac{1}{2}\delta^2\|a\|^2\big).$$

Let $\delta_k = \sqrt{2}\,\|a^2(2^k)\|/\|a\|$. Then

$$\nu(B(x,\delta_k)) = \mathbb{P}\Big(\sum_{i=1}^\infty a_i^4\varepsilon_i \ge \sum_{i=1}^{2^k-1} a_i^4\Big).$$

Now observe that the inequality $\sum_{i=1}^\infty a_i^4\varepsilon_i \ge \sum_{i=1}^{2^k-1} a_i^4$ holds at least on the event where $\varepsilon_1 = \cdots = \varepsilon_{2^k-1} = 1$ and $\sum_{i=2^k}^\infty a_i^4\varepsilon_i \ge 0$, which has probability 2^{-2^k}. Therefore $\nu(B(x,\delta_k)) \ge 2^{-2^k}$, and consequently

$$\sqrt{\log_2(\nu(B(x,\delta_k))^{-1})} \le 2^{k/2}.$$

Clearly

$$\int_0^\infty \sqrt{\log_2(\nu(B(x,\delta))^{-1})}\,d\delta \le \sum_{k=1}^\infty 2^{k/2}(\delta_{k-1}-\delta_k).$$

Moreover using the triangle inequality we get

$$\delta_{k-1}-\delta_k \le \|a\|^{-1}\sqrt{\sum_{i=2^{k-1}}^{2^k-1} a_i^4} \le \|a\|^{-1}2^{(k-1)/2}a_{2^{k-1}}^2.$$

Consequently,

$$\sum_{k=1}^\infty 2^{k/2}(\delta_{k-1}-\delta_k) \le \sqrt{2}\,\|a\|^{-1}\sum_{k=1}^\infty 2^{k-1}a_{2^{k-1}}^2.$$

It remains to observe that

$$\sum_{k=1}^\infty 2^{k-1}a_{k-1}^2 \le 2\|a\|^2.$$

Therefore

$$\int_0^\infty \sqrt{\log_2(\nu(B(x,\delta))^{-1})}\,d\delta \le 2\sqrt{2}\,\|a\|.$$

Integration of the inequality with respect to ν completes the argument. ∎

The theory is also useful to obtain some lower bounds on the small value probability. Recall that for points x of the form $a^2\bar{\varepsilon}/\|a\| = (a_i^2\bar{\varepsilon}_i/\|a\|)_{i=1}^\infty$

with $\bar{\varepsilon}_i = \pm 1$ (the support of ν) the values of $\nu(B(x,\delta))$ are all equal to

$$\nu(B(x,\delta)) = \mathbb{P}\big(\langle a^2\varepsilon, a^2\rangle \geq \|a^2\|^2 - \tfrac{1}{2}\delta^2\|a\|^2\big).$$

On the other hand, the bound $\mathcal{M}(\nu,\nu) \leq K\|a\|$, where K is a universal constant, implies that

$$\delta\sqrt{\log_2(\nu(B(x,\delta))^{-1})} \leq \mathcal{M}(\nu,\nu)K\|a\|$$

for any x of the form $= a^2\bar{\varepsilon}/\|a\| = (a_i^2\bar{\varepsilon}_i/\|a\|)_{i=1}^{\infty}$, $\bar{\varepsilon}_i = \pm 1$, and hence

$$\nu(B(x,\delta)) \geq \exp(-\bar{K}^2\|a\|^2/\delta^2)$$

with $\bar{K} = \sqrt{\log_2(e)}\,K$. In this way we have proved

Corollary 7.4. *For any sequence $(t_i)_{i=1}^{\infty}$ of positive numbers such that $\|t^{1/4}\| < \infty$ we have*

$$\mathbb{P}\Big(\sum_{i=1}^{\infty} t_i\varepsilon_i \geq \sum_{i=1}^{\infty} t_i - \tfrac{1}{2}\delta^2\|t^{1/4}\|^4\Big) \geq \exp(-\bar{K}^2/\delta^2).$$

Note that lower bounds in the problem of small value probability are usually difficult to get.

Similar results can be established for Gaussian variables and more generally for centered random vectors valued in the Euclidean space. Indeed, if X is a random vector valued in $\mathbb{R}^n$ in the isotropic position and A a given $n\times n$ matrix, then for the set $T = A^*B_1^n$, where $B_1^n = B_{\|\cdot\|}(0,1)$ is the Euclidean ball of radius 1 in $\mathbb{R}^n$ centered at 0, we can consider the process $Z(t) = \langle t, X\rangle$, $t \in T$. Again due to the isotropy assumption $\mathbb{E}|Z(t)-Z(s)|^2 = \|t-s\|^2$ for any $s,t \in T$, and obviously

$$\mathbb{E}\sup_{t\in T} Z(t) = \mathbb{E}\|AX\|.$$

Moreover the supremum distribution m is the same as the distribution of the vector $A^*AX/\|AX\|$, and hence for any $x \in A^*B_1^n$,

$$m(B(x,\delta)) = \mathbb{P}\big(\big\|A^*AX - x\|AX\|\big\| \leq \delta\|AX\|\big).$$

Consequently, whenever the presented Gaussian type approach may be used to get a lower bound on $\mathbb{E}\|AX\|$ it should result in lower bounds on the small value probability, i.e. $m(B(x,\delta))$, $\delta > 0$.

8. The duality principle. Let φ be a Young function. In this section we consider general processes $X(t)$, $t \in T$, on (T,ρ) under the increment condition (2.1). By the result of [3], for φ that satisfy

$$\varphi(2x) \geq 2C\varphi(x) \quad \text{for some } C > 1 \tag{8.1}$$

and small enough $x \geq 0$ we have

$$\mathcal{S} = \sup_X \mathbb{E}\sup_{s,t\in T}|X(t)-X(s)| \geq K^{-1}\sup_{\mu}\mathcal{M}_{\rho,\varphi}(\mu,\mu),$$

where the supremum is taken over all processes $X(t)$, $t \in T$, that satisfy (2.1). On the other hand, by (2.2),

$$\mathcal{S} \le K \sup_{\mu} \mathcal{M}_{\rho,\varphi}(\mu,\mu), \tag{8.2}$$

and therefore $\mathcal{S} = \sup_X \mathbb{E} \sup_{s,t\in T} |X(t) - X(s)|$ is comparable with $\sup_\mu \mathcal{M}_{\rho,\varphi}(\mu,\mu)$. By the general result on majorizing measures [2] we have

$$\mathcal{S} \le K \inf_{\mu} \sup_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t),$$

and hence

$$K^{-1} \inf_{\mu} \sup_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t) \le \sup_{\mu} \mathcal{M}_{\rho,\varphi}(\mu,\mu) \le K \inf_{\mu} \sup_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t) \tag{8.3}$$

for a large class of φ and ρ on T. The objective of this section is to show that there is another quantity comparable with $\sup_\mu \mathcal{M}_{\rho,\varphi}(\mu,\mu)$, namely $\sup_\mu \inf_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t)$. We now state the main result for this section, which is an extension of Theorem 1.8.

THEOREM 8.1. *Assuming that φ satisfies* (8.1) *there exists a universal constant $K < \infty$ such that*

$$K^{-1} \sup_{\mu} \inf_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t) \le \sup_{\mu} \mathcal{M}_{\rho,\varphi}(\mu,\mu) \le K \sup_{\mu} \inf_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t).$$

Proof. In the proof we need that $\varphi^{-1}(1/x)$ is comparable up to a universal constant with a convex function and $\varphi(1) = 1$. Namely by Lemma 2.1 in [15] there exists a convex function ξ such that

$$\tfrac{1}{2}\varphi^{-1}(1/x) \le \xi(x) \le \varphi^{-1}(1/x) \quad \text{for } x > 0. \tag{8.4}$$

Consequently, it suffices to prove the result for a convex function $\xi(x)$ comparable up to a universal constant with $\varphi^{-1}(1/x)$, which can also be normed so that $\xi(1) = 1$. For simplicity we keep the notation $\varphi^{-1}(1/x)$ even when $\xi(x)$ is used, but we stress that by $\varphi^{-1}(1/x)$ we always mean the convex equivalent.

Clearly $\sup_\mu \mathcal{M}_{\rho,\varphi}(\mu,\mu) \ge \sup_\mu \inf_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t)$, which implies that

$$\mathcal{S} \ge K^{-1} \sup_{\mu} \inf_{t\in T} \mathcal{M}_{\rho,\varphi}(\mu,\delta_t). \tag{8.5}$$

By (8.2) and (8.3), for any measure ν on T,

$$\mathcal{S} \le K \sup_{t\in T} \mathcal{M}_{\rho,\varphi}(\nu,\delta_t). \tag{8.6}$$

We show that on each finite subset $F \subset T$ there exists an equality measure ν_F on F such that $\sigma_{\nu_F,\rho,\varphi}(t)$ are equal on each $t \in F$ and finite. Indeed, let $F = \{t_1,\dots,t_m\}$, and note that each probability measure μ on F can be treated as a point $(\alpha(1),\dots,\alpha(m))$ in the simplex $\triangle_m = \{(\alpha(1),\dots,\alpha(m)) : \alpha(i) \ge 0, \sum_{i=1}^m \alpha(i) = 1\}$, namely we set $\alpha(i) = \mu(t_i)$. We define a mapping

$\Phi : \triangle_m \to \mathbb{R}^m$, $\Phi = (\Phi_1, \dots, \Phi_m)$, by

$$\Phi_i(\mu) = \sigma_{\mu,\rho,\varphi}(t_i) = \int_0^{\mathrm{Diam}_\rho(T)} \varphi^{-1}\bigg(\frac{1}{\mu(B(t_i,\varepsilon))}\bigg)\, d\varepsilon.$$

LEMMA 8.2. *There exists a unique measure ν_F on F such that $\Phi_i(\nu_F)$ are all equal and finite for $1 \le i \le m$.*

Proof. Note that by the convexity of $\varphi^{-1}(1/x)$ (the discussion of (8.4)), Φ is convex and continuous on $\triangle_m$. Moreover $\Phi_i(\delta_{t_i}) = \mathrm{Diam}_\rho(T)$ and $\Phi_i(\delta_{t_j}) = \infty$ if $i \ne j$. Therefore Φ is simplicial in the sense that each facet of $\triangle_m$, say

$$\triangle_I = \{(\alpha(1), \dots, \alpha(m)) : \alpha(i) = 0,\ i \in I;\ \alpha(i) > 0,\ i \notin I\}$$

for some $I \subset \{1, \dots, m\}$, is mapped onto $\bar{\triangle}_I$, where

$$\bar{\triangle}_I = \{(\Phi_1(\mu), \dots, \Phi_m(\mu)) : \Phi_i(\mu) = \infty,\ i \in I;\ \Phi_i(\mu) < \infty,\ i \notin I\}.$$

Consequently, $\triangle_{[m]}$, where $[m] = \{1, \dots, m\}$, must be mapped onto the convex surface in $\mathbb{R}^m$ that connects the points $x_i = (x_i(1), \dots, x_i(m))$, $1 \le i \le m$, where $x_i(i) = \mathrm{Diam}_\rho(T)$ and $x_i(j) = \infty$ if $i \ne j$. This implies that there exists exactly one point of intersection of the surface with $y \mapsto (y, \dots, y)$, $y \in \mathbb{R}$. Therefore there exists exactly one probability measure ν such that $\Phi_i(\nu)$ are all equal and finite for $1 \le i \le m$. ■

Consequently, by (8.6) and Lemma 8.2 we obtain

$$\mathbb{E} \sup_{s,t \in F} |X(t) - X(s)| \le K \inf_{t \in F} \mathcal{M}_{\rho,\varphi}(\nu_F, \delta_t),$$

and therefore

$$\mathcal{S} \le K \sup_{\nu} \inf_{t \in T} \mathcal{M}_{\rho,\varphi}(\nu, \delta_t). \tag{8.7}$$

Clearly (8.5) and (8.7) complete the proof. ■

This proves the duality principle.

COROLLARY 8.3. *The following quantities are comparable up to a universal constant:*

$$\inf_{\mu} \sup_{t \in T} \mathcal{M}_{\rho,\varphi}(\mu, \delta_t) \quad \textit{and} \quad \sup_{\mu} \inf_{t \in T} \mathcal{M}_{\rho,\varphi}(\mu, \delta_t).$$

That is, either we can search for the optimal measure μ that works for all $t \in T$, or for all measures we have to find the worst possible point $t \in T$.

As we have pointed out, the result has an application to the extension of the Dvoretzky theorem to metric spaces [12].

Acknowledgements. This research was partly supported by MNiSW (grant no. N N201 397437).

REFERENCES

[1] R. Adler, *An Introduction to Continuity, Extrema, and Related Topics for General Gaussian Processes*, Inst. Math. Statist., Hayward, CA, 1990.

[2] W. Bednorz, *A theorem on majorizing measures*, Ann. Probab. 34 (2006), 1771–1781.

[3] W. Bednorz, *Majorizing measures on metric spaces*, C. R. Math. Acad. Sci. Paris 348 (2010), 75–78.

[4] W. Bednorz, *On the convergence of orthogonal series*, C. R. Math. Acad. Sci. Paris 349 (2011), 455–458.

[5] W. Bednorz, *On the complete characterization of the sample boundedness of orthogonal processes*, Ann. Probab. 41 (2013), 1055–1071.

[6] R. M. Dudley, *Sample functions of the Gaussian process*, Ann. Probab. 1 (1973), 66–103.

[7] X. Fernique, *Regularité des trajectoires des fonctions aléatoires gaussiennes*, in: École d'Été de Probabilités de Saint-Flour, IV-1974, Lecture Notes in Math. 480, Springer, Berlin, 1975, 1–96.

[8] X. Fernique, *Caractérisation de processus à trajectoires majorées ou continues*, in: Séminaire de Probabilités XII, Lecture Notes in Math. 649, Springer, Berlin, 1978, 691–706.

[9] J. Kim and D. Pollard, *Cube root asymptotics*, Ann. Statist. 18 (1990), 191–219.

[10] M. Ledoux, *The Concentration of Measure Phenomenon*, Math. Surveys Monogr. 89, Amer. Math. Soc., 2001.

[11] M. Ledoux and M. Talagrand, *Probability in Banach Spaces. Isoperimetry and Processes*, Ergeb. Math. Grenzgeb. 23, Springer, Berlin, 1991.

[12] M. Mendel and A. Naor, *Ultrametric skeletons*, Proc. Nat. Acad. Sci. USA 110 (2013), 19256–19262.

[13] V. N. Sudakov and B. S. Tsirelson, *Extremal properties of half-spaces for spherically invariant measures*, II, Zap. Nauch. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 41 (1974), 14–24 (in Russian).

[14] M. Talagrand, *Regularity of Gaussian processes*, Acta Math. 159 (1987), 99–149.

[15] M. Talagrand, *Sample boundedness of stochastic processes under increment conditions*, Ann. Probab. 18 (1990), 1–49.

[16] M. Talagrand, *A simple proof of the majorizing measure theorem*, Geom. Funct. Anal. 2 (1992), 118–125.

[17] M. Talagrand, *The Generic Chaining*, Springer, 2005.

Witold Bednorz
Institute of Mathematics
University of Warsaw
Banacha 2
02-097 Warszawa, Poland
E-mail: wbednorz@mimuw.edu.pl

Received 30 October 2013;
revised 24 October 2014 (6061)