

## APPROXIMATION, METRIC ENTROPY AND SMALL BALL ESTIMATES FOR GAUSSIAN MEASURES

BY WENBO V. LI<sup>1</sup> AND WERNER LINDE

*University of Delaware and FSU Jena*

A precise link proved by Kuelbs and Li relates the small ball behavior of a Gaussian measure  $\mu$  on a Banach space  $E$  with the metric entropy behavior of  $K_\mu$ , the unit ball of the reproducing kernel Hilbert space of  $\mu$  in  $E$ . We remove the main regularity assumption imposed on the unknown function in the link. This enables the application of tools and results from functional analysis to small ball problems and leads to small ball estimates of general algebraic type as well as to new estimates for concrete Gaussian processes. Moreover, we show that the small ball behavior of a Gaussian process is also tightly connected with the speed of approximation by “finite rank” processes.

**1. Introduction.** Let  $\mu$  denote a centered Gaussian measure on a real separable Banach space  $E$  with norm  $\|\cdot\|$  and dual  $E^*$ . The small ball problem studies the behavior of

$$(1.1) \quad \log \mu(x: \|x\| \leq \varepsilon) = -\phi(\varepsilon)$$

as  $\varepsilon \rightarrow 0$ . The complexity of  $\phi(\varepsilon)$  is well known, and there are only a few Gaussian measures for which  $\phi(\varepsilon)$  has been determined completely as  $\varepsilon \rightarrow 0$ . One of very few general results is a precise link, discovered in Kuelbs and Li (1993), between the function  $\phi(\varepsilon)$  and the metric entropy of the unit ball  $K_\mu$  of the Hilbert space  $H_\mu$  generated by  $\mu$ . In order to give precise statements, we need some notations and definitions.

We recall first that if  $(E, d)$  is any metric space and  $A$  is a compact subset of  $(E, d)$ , then the  $d$ -metric entropy of  $A$  is denoted by  $H(A, \varepsilon) = \log N(A, \varepsilon)$  where the minimum covering number,

$$N(A, \varepsilon) = \min \left\{ n \geq 1: \exists x_1, \dots, x_n \in A \text{ such that } \bigcup_{j=1}^n B_\varepsilon(x_j) \supseteq A \right\}$$

and  $B_\varepsilon(x_0) = \{x: d(x, x_0) < \varepsilon\}$  is the open ball of radius  $\varepsilon$  centered at  $x_0$ .

Next the Hilbert space  $H_\mu$  generated by  $\mu$  can be described as the completion of the range of the mapping  $S: E^* \rightarrow E$  defined via the Bochner

---

Received May 1998; revised February 1999.

<sup>1</sup>Supported in part by the NSF.

AMS 1991 subject classifications. Primary 60G15; secondary 60F99, 47D50, 47G10.

Key words and phrases. Gaussian process, small deviation, metric entropy, approximation number.

integral,

$$Sa = \int_E xa(x) d\mu(x), \quad a \in E^*$$

and the completion is in the inner product norm

$$\langle Sa, Sb \rangle_\mu = \int_E a(x)b(x) d\mu(x), \quad a, b \in E^*.$$

Lemma 2.1 in Kuelbs (1976) presents the details of this construction along with various properties of the relationship between  $H_\mu$  and  $E$ ; see also Ledoux and Talagrand (1991). The most important fact for us is that the unit ball  $K_\mu$  of  $H_\mu$  is always compact. Hence  $K_\mu$  has finite metric entropy. Finally, in order to compare the asymptotic rates, we write  $f(x) \leq g(x)$  as  $x \rightarrow a$  if  $\limsup_{x \rightarrow a} f(x)/g(x) < \infty$ , and  $f(x) \approx g(x)$  as  $x \rightarrow a$  if  $f(x) \leq g(x)$  and  $g(x) \leq f(x)$ .

Now we can state the links between the small ball function  $\phi(\varepsilon)$  given in (1.1) and the metric entropy function  $H(K_\mu, \varepsilon)$ . Note that the original statements of these results are all bilateral and are not as precise as summarized below. The proof of this generalization follows by the same arguments as in Kuelbs and Li (1993), together with basic properties of slowly varying functions.

**THEOREM 1.1** [Kuelbs and Li (1993)]. *Let  $f(1/x)$  and  $g(1/x)$  be regularly varying functions at infinity. We have the following for  $\varepsilon$  small:*

(i) *It holds*

$$H(K_\mu, \varepsilon/\sqrt{2\phi(\varepsilon)}) \geq \phi(2\varepsilon).$$

*In particular, if  $\phi(\varepsilon) \leq \phi(2\varepsilon)$  and  $\phi(\varepsilon) \geq \varepsilon^{-\alpha}J(\varepsilon^{-1})$ ,  $\alpha > 0$  and  $J(x)$  is slowly varying at infinity such that  $J(x) \approx J(x^\rho)$  as  $x \rightarrow \infty$  for each  $\rho > 0$ , then*

$$(1.2) \quad H(K_\mu, \varepsilon) \geq \varepsilon^{-2\alpha/(2+\alpha)}J(1/\varepsilon)^{2/(2+\alpha)}.$$

*Especially, (1.2) holds whenever  $\phi(\varepsilon) \approx \varepsilon^{-\alpha}J(\varepsilon^{-1})$  with  $J$  as before.*

(ii) *If  $\phi(\varepsilon) \leq f(\varepsilon)$ , then*

$$H(K_\mu, \varepsilon/\sqrt{f(\varepsilon)}) \leq f(\varepsilon).$$

*In particular, if  $f(\varepsilon) = \varepsilon^{-\alpha}J(\varepsilon^{-1})$ ,  $\alpha > 0$  and  $J$  is as above, then*

$$H(K_\mu, \varepsilon) \leq \varepsilon^{-2\alpha/(2+\alpha)}J(1/\varepsilon)^{2/(2+\alpha)}.$$

(iii) *If  $H(K_\mu, \varepsilon) \geq g(\varepsilon)$ , then*

$$\phi(\varepsilon) \geq g(\varepsilon/\sqrt{\phi(\varepsilon)}).$$

*In particular, if  $g(\varepsilon) = \varepsilon^{-\alpha}J(1/\varepsilon)$  where  $0 < \alpha < 2$  and  $J$  is as in (i), then*

$$\phi(\varepsilon) \geq \varepsilon^{-2\alpha/(2-\alpha)}(J(1/\varepsilon))^{2/(2-\alpha)}.$$

(iv) *If  $H(K_\mu, \varepsilon) \leq g(\varepsilon)$ , then*

$$\phi(2\varepsilon) \leq g(\varepsilon/\sqrt{\phi(\varepsilon)}).$$

First note that the restriction on  $\alpha$  in part (iii) of the theorem is natural since it is known from Goodman (1990) that  $H(K_\mu, \varepsilon) = o(\varepsilon^{-2})$  regardless of the Gaussian measure  $\mu$ . Second, we see clearly from (i) and (ii) of Theorem 1.1 that in almost all cases of interest, small ball probabilities provide sharp estimates on the metric entropy. This approach has been applied successfully to various problems on estimating metric entropy; see, for example, Kuelbs and Li (1993). On the other hand, the direction from the metric entropy to small ball probabilities has not been fully used because, in order to use (iv) of Theorem 1.1, we need to check a regularity condition such as  $\phi(2\varepsilon) \geq \phi(\varepsilon)$  on the small ball probability. It is easily seen from the proof of (iv) given in Kuelbs and Li (1993) that we only need  $\phi((1 + \delta)\varepsilon) \geq \phi(\varepsilon)$  for some  $\delta > 0$ . Nevertheless, it is an assumption that significantly restricts the ease of application from upper bound on metric entropy to upper bound on  $\phi(\varepsilon)$ .

One of the main result of this paper is to replace (iv) in Theorem 1.1 by Theorem 1.2 below so that we can obtain the desired estimate of small ball probability based only on the upper bound of metric entropy.

**THEOREM 1.2.** *If*

$$H(K_\mu, \varepsilon) \leq \varepsilon^{-\alpha} J(1/\varepsilon),$$

where  $0 < \alpha < 2$  and  $J(x)$  is slowly varying and such that  $J(x) \approx J(x^\rho)$  as  $x \rightarrow \infty$  for each  $\rho > 0$ , then for  $\varepsilon$  small,

$$\phi(\varepsilon) \leq \varepsilon^{-2\alpha/(2-\alpha)} (J(1/\varepsilon))^{2/(2-\alpha)}.$$

As a simple consequence of Theorems 1.1 and 1.2 together, it is easy to see that for  $\alpha > 0$  and  $\beta \in \mathbb{R}$ ,

$$(1.3) \quad \phi(\varepsilon) \approx \varepsilon^{-\alpha} (\log 1/\varepsilon)^\beta \quad \text{iff} \quad H(K_\mu, \varepsilon) \approx \varepsilon^{-2\alpha/(2+\alpha)} (\log 1/\varepsilon)^{2\beta/(2+\alpha)}.$$

Theorem 1.2 not only completes the links in Theorem 1.1 when compared with (iii) of Theorem 1.1, but more important, allows applications of tools and results from functional analysis to attack important problems of estimating the small ball probabilities which are of great interest in probability theory. For example, in a recent paper [cf. Dunker, Kühn, Lifshits and Linde (1998)] a special case of Theorem 1.2 was already used to obtain small ball estimates for Brownian sheets in dimension  $d \geq 2$ . Moreover, in Dunker (1998), our Theorem 1.2 has been successfully applied to get small ball estimates for fractional Brownian sheets and it can also improve the estimates given in Belinsky (1998). We will present a few additional examples of the small ball estimates of fractional integrated (fractional) Brownian motions in the last section. The estimate for the very special case, integrated Brownian motion, was considered recently in Khoshnevisan and Shi (1998) by using local time techniques. As far as we know, our Theorem 1.2, together with basic techniques of estimating entropy numbers as demonstrated in Sections 5 and 6, is one of the most general and powerful among all the existing methods of estimating the small ball lower bound for Gaussian processes. Another

commonly used general lower bound estimate on supremum of Gaussian processes was established in Talagrand (1993) and a nice formulation was given in Ledoux (1996), page 257. Some applications can be found in Li (1998) and Li and Linde (1998). Although it is relatively easy to use, it does not provide sharp estimates for many interesting Gaussian processes such as all the examples considered in Section 6. In fact, this was one of the motivations for us to develop Theorem 1.2. The other was to complete the link between metric entropy and small ball behavior.

The proof of Theorem 1.2 is carried out in two stages. We show first in Section 2 how other approximation quantities are related to small ball probabilities. These estimates are of independent interest. Based on a totally different argument than the one used to establish (iv) of Theorem 1.1, they provide a preliminary estimate of the small ball probability in terms of metric entropy without any regularity conditions. Then in Section 3, we prove Theorem 1.2 by applying an iteration procedure on the key estimate which was used to establish (iv) of Theorem 1.1 and was obtained from the isoperimetric inequality for Gaussian measures. To make the iteration work, we have to use not only the assumption on metric entropy but also the weaker estimate on the small ball obtained in Section 2. After proving Theorem 1.2, we turn the tables in Section 4. With the help of Theorem 1.2, we obtain more precise estimates on those quantities related to Section 2. In Section 5, we show how simple properties of entropy numbers can give nontrivial algebraic properties of small ball probabilities. In particular, Theorem 5.2 is very general and somewhat surprising. The power of these estimates can be seen from the integral operators discussed in Section 6 together with many applications to various concrete Gaussian processes. Here, in particular, Theorem 6.1 is very simple and useful.

**2.  $l$ -Approximation and measures of small balls.** Let  $H$  be a separable Hilbert space and let  $u$  be an operator from  $H$  into  $E$ . If  $\dim(H) < \infty$ , we may identify  $H$  with  $\mathbb{R}^n$  and define

$$l(u) = \left( \int_{\mathbb{R}^n} \|u(h)\|^2 d\gamma_n(h) \right)^{1/2}$$

where  $\gamma_n$  is the  $n$ -dimensional standard normal law on  $\mathbb{R}^n$ . For  $\dim(H) = \infty$ , we set

$$l(u) = \sup\{l(u|_{H_0}): H_0 \subseteq H, \dim(H_0) < \infty\}.$$

Operators with  $l(u) < \infty$  were first investigated in Linde and Pietsch (1974) and were called  $\gamma$ -summing operators. The mapping  $l$  is a norm on the class of operators  $u$  from  $H$  into  $E$  with  $l(u) < \infty$ . Here are some well-known properties of the  $l$ -norm; more details can be found in Pisier (1989). Assuming  $l(u) < \infty$  for  $u: H \rightarrow E$ , the random sum  $\sum_{j=1}^{\infty} \xi_j u(f_j)$  converges in  $E$  a.s. (or in  $L_2$ -sense) for some (all) complete orthonormal base  $(f_j)_{j=1}^{\infty}$  in  $H$  iff  $u$  is in the

$l$ -closure of finite rank operators from  $H$  into  $E$ . In this case,

$$l(u) = \left( \mathbb{E} \left\| \sum_{j=1}^{\infty} \xi_j u(f_j) \right\|^2 \right)^{1/2}$$

is independent of the ONB  $(f_j)_{j=1}^{\infty}$ , where here and throughout this paper  $\xi_1, \xi_2, \dots$ , denotes an i.i.d. sequence of  $\mathcal{N}(0, 1)$ -distributed random variables. Moreover, for some (each) increasing sequence  $P_n$  of finite rank orthogonal projections in  $H$  with  $P_n h \rightarrow h, h \in H$ , it follows that

$$(2.1) \quad \lim_{n \rightarrow \infty} l(u - uP_n) = 0.$$

To quantify the speed of convergence in (2.1), the  $l$ -approximation numbers [cf. Pisier (1989)] are defined by

$$\begin{aligned} l_n(u) &= \inf\{l(u - uP) : P \text{ orth. proj. in } H, \text{rk}(P) < n\} \\ &= \inf\{l(u|_{H_0}) : \text{codim}(H_0) < n\} \\ &= \inf \left\{ \left( \mathbb{E} \left\| \sum_{j=n}^{\infty} \xi_j u(f_j) \right\|^2 \right)^{1/2} : \dim\{f_n, f_{n+1}, \dots\}^{\perp} = n - 1 \right\}, \end{aligned}$$

with  $f_j$ 's orthonormal. By standard arguments, we have

$$l_n(u) = \inf\{l(u - v) : \text{rk}(v) < n, v : H \rightarrow E\},$$

so  $l_n(u)$  measures the distance (in the  $l$ -norm) of  $u$  to the set of operators with rank less than  $n$ . In fact, the sum  $\sum_{j=1}^{\infty} \xi_j u(f_j)$  converges a.s. iff  $l(u) < \infty$  and, moreover,  $l_n(u) \rightarrow 0$  as  $n \rightarrow \infty$ . It is also well known [cf. Pisier (1989)] that any operator  $u$  with  $l(u) < \infty$  is compact and hence the dual operator  $u^* : E^* \rightarrow H$  is compact as well, thus the metric entropy  $H(u^*(B_{E^*}), \varepsilon)$  where  $B_{E^*} = \{a \in E^* : \|a\|_{E^*} \leq 1\}$ , is finite and provides a measurement for the compactness of  $u^*(B_{E^*})$ . For our purposes it is more convenient to work with a related quantity  $e_n(u^*)$ , which is defined in general by

$$e_n(v) = \inf\{\varepsilon > 0 : N(v(B_E), \varepsilon) \leq 2^{n-1}\}$$

for any compact operator  $v$  from a Banach space  $E$  into another one and it is called the  $n$ th dyadic entropy number of  $v$ . Now we are ready to state two basic results.

LEMMA 2.1 [Pisier (1989), page 141]. *There exist universal constants  $c_1, c_2 > 0$  such that for all operators  $u$  from  $H$  into  $E$  and all  $n \in \mathbb{N}$ , the estimate*

$$(2.2) \quad l_n(u) \leq c_1 \sum_{k \geq c_2 n} e_k(u^*) k^{-1/2} (1 + \log k)$$

*holds. Moreover, if  $E$  is  $K$ -convex (e.g.,  $L_p, 1 < p < \infty$ ), that is,  $E$  does not contain  $l_1^n$ 's uniformly [cf. Pisier (1989), Theorem 2.4], then (2.2) is valid without the log-term on the right-hand side.*

LEMMA 2.2 [Tomczak-Jaegermann (1987)]. *Let  $u: H \rightarrow E$  and assume for some  $\alpha > 0$  and  $\beta \in \mathbb{R}$ ,*

$$e_n(u) \leq n^{-1/\alpha}(1 + \log n)^\beta.$$

*Then we have*

$$e_n(u^*) \leq n^{-1/\alpha}(1 + \log n)^\beta.$$

Note that the above result relates the entropy numbers of an operator with those of its adjoint, provided it is defined on a Hilbert space. The similar relation for operator on Banach space is still an open question [cf. Bourgain, Pajor, Szarek and Tomczak-Jaegermann (1989)]. Next we combine Lemma 2.1 and Lemma 2.2 together to obtain the following.

PROPOSITION 2.1. *Suppose that*

$$(2.3) \quad e_n(u) \leq n^{-1/\alpha}(1 + \log n)^\beta$$

*for  $u: H \rightarrow E$  and some  $\alpha \in (0, 2)$ ,  $\beta \in \mathbb{R}$ . Then we have*

$$(2.4) \quad l_n(u) \leq n^{-(2-\alpha)/(2\alpha)}(1 + \log n)^{\beta+1}.$$

*Moreover, if  $E$  is  $K$ -convex, then (2.4) holds with power  $\beta$  instead of  $\beta + 1$  on the log term.*

Note that (2.3) is equivalent to

$$H(K, \varepsilon) \leq \varepsilon^{-\alpha}(\log 1/\varepsilon)^{\alpha\beta},$$

with  $K = u(B_H) = \{u(h): \|h\|_H \leq 1\}$ .

The main result of this section is to relate  $l$ -approximation numbers with small ball behavior of suitable Gaussian measures. To do so we have to translate the preceding results into the language of Gaussian measures or random variables, respectively. The next well-known result gives us the desired relation between Gaussian random variables, operators and random series.

LEMMA 2.3. *Let  $E$  be a separable Banach space and let  $X$  be an  $E$ -valued random variable. Then the following are equivalent:*

(i)  *$X$  is centered Gaussian.*

(ii) *There exist a separable Hilbert space  $H$  and an operator  $u: H \rightarrow E$  such that  $\sum_{j=1}^\infty \xi_j u(f_j)$  converges a.s. in  $E$  for one (each) ONB  $(f_j)_{j=1}^\infty$  in  $H$  and*

$$(2.5) \quad X =_d \sum_{j=1}^\infty \xi_j u(f_j).$$

(iii) *There are  $x_1, x_2, \dots$ , in  $E$  such that  $\sum_{j=1}^\infty \xi_j x_j$  converges a.s. in  $E$  and*

$$(2.6) \quad X =_d \sum_{j=1}^\infty \xi_j x_j.$$

Let  $\mu = \text{dist}(X)$  be the distribution of a centered Gaussian random variable  $X$  and let  $H_X = H_\mu$  be the Hilbert space mentioned in the Introduction. Then it is easy to see that

$$H_X = u(H) = \left\{ \sum_{j=1}^\infty \alpha_j x_j : \sum_{j=1}^\infty \alpha_j^2 < \infty \right\},$$

where  $u$  and the  $x_j$ 's are related with  $X$  via (2.5) and (2.6), respectively.

Now in view of Lemma 2.3 the quantities  $l$  and  $l_n$  can be carried over from operators to random variables and random sums. Let  $X$  be a centered Gaussian random variable; then we can set

$$l(X) = l(u) = \left( \mathbb{E} \left\| \sum_{j=1}^\infty \xi_j x_j \right\|^2 \right)^{1/2} = (\mathbb{E} \|X\|^2)^{1/2},$$

where  $u$  and the  $x_j$ 's are related to  $X$  via Lemma 2.3. Next the  $n$ th  $l$ -approximation number of  $X$  can be defined by  $l_n(X) = l_n(u)$ , where  $u$  is as in (2.5) and, as can be seen easily,  $l_n(X)$  is independent of the special choice of  $u$  in (2.5). Moreover,

$$(2.7) \quad l_n(X) = \inf \left\{ \left( \mathbb{E} \left\| \sum_{j=n}^\infty \xi_j x_j \right\|^2 \right)^{1/2} : X =_d \sum_{j=1}^\infty \xi_j x_j \right\}.$$

With the above notation, Proposition 2.1 can be formulated as follows.

**PROPOSITION 2.2.** *Let  $X$  be a centered Gaussian random variable with values in  $E$  and let  $K_X$  be the unit ball of  $H_X$ . If for  $0 < \alpha < 2$  and  $\beta \in \mathbb{R}$ ,*

$$H(K_X, \varepsilon) \leq \varepsilon^{-\alpha} \left( \log \frac{1}{\varepsilon} \right)^{\alpha\beta},$$

then

$$(2.8) \quad l_n(X) \leq n^{-(2-\alpha)/(2\alpha)} (1 + \log n)^{\beta+1}.$$

For  $K$ -convex spaces (2.8) holds with power  $\beta$  instead of  $\beta + 1$  on the log-term.

With all the preparation given above, we can finally relate  $l_n(X)$  with the small ball behavior of  $X$ , that is, the behavior of the function

$$\phi(\varepsilon) = -\log \mu(x : \|x\| \leq \varepsilon) = -\log \mathbb{P}(\|X\| \leq \varepsilon).$$

Let us define for  $\varepsilon > 0$  small enough positive integers  $n(\varepsilon)$  by

$$(2.9) \quad n(\varepsilon) = \max\{n : 4 \cdot l_n(X) \geq \varepsilon\}$$

and note that  $l_n(X)$  tends to zero monotonically as  $n \rightarrow \infty$ .

**PROPOSITION 2.3.** *Let  $X$  be a centered Gaussian random variable. Then*

$$\phi(\varepsilon) \leq n(\varepsilon) \log \frac{n(\varepsilon)}{\varepsilon}.$$

PROOF. The basic idea of the proof given here is very simple. It had been used to estimate small ball behavior for  $l_2$ -norm in Hoffmann-Jørgensen, Dudley and Shepp (1979) and for  $l_p$ -norm in Li (1992).

Given  $\varepsilon > 0$ , let  $n = n(\varepsilon)$  be defined in (2.9) which implies  $l_{n+1}(X) < \varepsilon/4$ . Thus by (2.7) we can find  $x_1, x_2, \dots, \in E$  such that

$$(2.10) \quad X =_d \sum_{j=1}^{\infty} \xi_j x_j \quad \text{and} \quad \mathbb{E} \left\| \sum_{j=n+1}^{\infty} \xi_j x_j \right\|^2 \leq \varepsilon^2/16.$$

Then we obtain

$$\begin{aligned} \mathbb{P}(\|X\| \leq \varepsilon) &= \mathbb{P} \left( \left\| \sum_{j=1}^n \xi_j x_j + \sum_{j=n+1}^{\infty} \xi_j x_j \right\| \leq \varepsilon \right) \\ &\geq \mathbb{P} \left( \left\| \sum_{j=1}^n \xi_j x_j \right\| \leq \frac{\varepsilon}{2} \right) \mathbb{P} \left( \left\| \sum_{j=n+1}^{\infty} \xi_j x_j \right\| \leq \frac{\varepsilon}{2} \right). \end{aligned}$$

Now using Chebyshev’s inequality, we have by (2.10)

$$\mathbb{P} \left( \left\| \sum_{j=n+1}^{\infty} \xi_j x_j \right\| \leq \frac{\varepsilon}{2} \right) \geq 1 - 4\varepsilon^{-2} \mathbb{E} \left\| \sum_{j=n+1}^{\infty} \xi_j x_j \right\|^2 \geq \frac{3}{4}.$$

Next note that when  $0 < \delta \leq 1$ , it holds for the standard normal random variable  $\xi$

$$\mathbb{P}(|\xi| \leq \delta) = \frac{2}{\sqrt{2\pi}} \int_0^\delta \exp\left(-\frac{t^2}{2}\right) dt \geq \frac{2\delta}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\right) \geq \frac{\delta}{3}.$$

Hence, for  $\varepsilon > 0$  small, we have by the well-known estimate  $\|x_j\| \leq (E\|X\|^2)^{1/2} = \sigma$  that

$$\begin{aligned} \mathbb{P} \left( \left\| \sum_{j=1}^n \xi_j x_j \right\| \leq \frac{\varepsilon}{2} \right) &\geq \mathbb{P} \left( |\xi_j| \|x_j\| \leq \frac{\varepsilon}{2n}, 1 \leq j \leq n \right) \\ &\geq \prod_{j=1}^n \mathbb{P} \left( |\xi_j| \leq \frac{\varepsilon}{2\sigma n} \right) \geq \left( \frac{\varepsilon}{6\sigma n} \right)^n. \end{aligned}$$

Combining the above estimates together, the assertion follows.  $\square$

Next if we combine the preceding proposition with Proposition 2.2, then the following weaker form of the link corresponding to (iv) of Theorem 1.1 holds.

PROPOSITION 2.4. *Let  $X$  be a centered Gaussian random variable and let  $K = K_X$  and  $\phi$  be as above. If*

$$H(K, \varepsilon) \leq \varepsilon^{-\alpha}$$



for some  $\alpha < 2$ , then

$$(2.11) \quad \phi(\varepsilon) \leq \varepsilon^{-\gamma}$$

for any  $\gamma > 2\alpha/(2 - \alpha)$ .

PROOF. We can pick  $\gamma' < \gamma$  such that

$$\frac{1}{\gamma} < \frac{1}{\gamma'} < \frac{1}{\alpha} - \frac{1}{2}.$$

By Proposition 2.2 we conclude  $l_n(X) \leq c \cdot n^{-1/\gamma'}$  which, of course, implies  $n(\varepsilon) \leq (1/\varepsilon)^{\gamma'}$  with  $n(\varepsilon)$  defined by (2.9). Thus by Proposition 2.3, we obtain

$$\phi(\varepsilon) \leq \varepsilon^{-\gamma'} \log \frac{1}{\varepsilon} \leq \varepsilon^{-\gamma},$$

which completes the proof.  $\square$

Note that the estimate in (2.11) is almost sharp and it is the base to obtain (2.11) with  $\gamma = 2\alpha/(2 - \alpha)$ , as we shall see in the next section.

**3. Proof of Theorem 1.2.** Our next result is essentially Theorem 1.2 under an additional assumption on the lower bound of small ball probability [upper bound on the function  $\phi(\varepsilon)$ ] given in (3.1). The key idea we used here is an iteration procedure.

PROPOSITION 3.1. *Assume there is a  $\gamma < \infty$  such that*

$$(3.1) \quad \phi(\varepsilon) = -\log \mathbb{P}(\|X\| \leq \varepsilon) \leq \varepsilon^{-\gamma}.$$

*If for  $K = K_X$ ,*

$$(3.2) \quad H(K, \varepsilon) \leq \varepsilon^{-\alpha} J(1/\varepsilon),$$

*where  $0 < \alpha < 2$  and  $J(x)$  is slowly varying and such that  $J(x) \approx J(x^\rho)$  as  $x \rightarrow \infty$  for each  $\rho > 0$ , then*

$$\phi(\varepsilon) \leq \varepsilon^{-2\alpha/(2-\alpha)} (J(1/\varepsilon))^{2/(2-\alpha)}.$$

PROOF. The starting point of our proof is one of the key lemmas in Kuelbs and Li (1993), that states

$$(3.3) \quad H(\lambda K, \varepsilon) - \phi(2\varepsilon) \geq \log \Phi(\lambda + \theta_\varepsilon)$$

for  $\lambda > 0$  where

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t \exp\left(-\frac{u^2}{2}\right) du \quad \text{and} \quad \log \Phi(\theta_\varepsilon) = -\phi(\varepsilon).$$

The proof of (3.3) depends on the isoperimetric inequality for Gaussian measures. Now taking  $\lambda = \sqrt{2\phi(\varepsilon)}$  and observing that  $\theta_\varepsilon \rightarrow -\infty$  as  $\varepsilon \rightarrow 0$ , we have

$$\phi(\varepsilon) = -\log \Phi(\theta_\varepsilon) = -\log(1 - \Phi(-\theta_\varepsilon)) \geq (-\theta_\varepsilon)^2/2,$$

which follows from the well-known inequality,

$$1 - \Phi(x) \leq x^{-1} \exp(-x^2/2) \leq \exp(-x^2/2)$$

for  $x \geq 1$ . Hence  $\lambda + \theta_\varepsilon = \sqrt{2\phi(\varepsilon)} + \theta_\varepsilon \geq 0$  and (3.3) implies

$$H(K, \varepsilon/\sqrt{2\phi(\varepsilon)}) \geq \phi(2\varepsilon) + \log(1/2).$$

Now changing  $\varepsilon$  to  $\varepsilon/2$  and using the assumption (3.2), we see that for  $\varepsilon > 0$  small and some universal constant  $C$ , which may differ from line to line below,

$$(3.4) \quad \begin{aligned} \phi(\varepsilon) &\leq \log 2 + H(K, \varepsilon/\sqrt{8\phi(\varepsilon/2)}) \\ &\leq C \cdot \varepsilon^{-\alpha} \phi(\varepsilon/2)^{\alpha/2} J(\sqrt{8\phi(\varepsilon/2)}/\varepsilon). \end{aligned}$$

In order to estimate the last term on the right-hand side of (3.4) further, we observe by using the assumptions on the function  $J$  and by applying Theorem 2.0.1 in Bingham, Goldie and Teugels (1987), that

$$(3.5) \quad \limsup_{x \rightarrow \infty} \sup_{a \leq \rho \leq b} J(x^\rho)/J(x) < \infty$$

for any  $1 < a < b < \infty$ . If  $0 < a < b < \infty$ , we can write (3.5) (for  $2$  and  $2b/a$ ) as

$$\limsup_{x \rightarrow \infty} \sup_{2 \leq \rho \leq 2b/a} J(x^{a\rho/2})/J(x^{a/2}) < \infty$$

and then it follows by using  $J(x^{a/2}) \approx J(x)$ ,

$$(3.6) \quad \sup_{a \leq \rho \leq b} J(x^\rho) \leq J(x)$$

for all  $0 < a < b < \infty$ . Note that  $1 \leq \phi(\varepsilon) \leq \varepsilon^{-\gamma}$  from (3.1); thus by (3.6) we have

$$(3.7) \quad J(\sqrt{8\phi(\varepsilon/2)}/\varepsilon) \leq C \sup_{1 \leq \rho \leq \gamma'} J(1/\varepsilon^\rho) \leq C \cdot J(1/\varepsilon)$$

with some  $\gamma' > 1 + \gamma/2$ . Thus we get for  $\varepsilon > 0$  small from (3.4) and (3.7) the relation

$$(3.8) \quad \log \phi(\varepsilon) \leq (2^{-1}\alpha) \log \phi(\varepsilon/2) + \log \psi(\varepsilon),$$

with  $\psi(\varepsilon) = C\varepsilon^{-\alpha}J(1/\varepsilon)$ . Now by iterating (3.8), it is clear that for any  $n \geq 1$ ,

$$(3.9) \quad \log \phi(\varepsilon) \leq (2^{-1}\alpha)^n \log \phi(\varepsilon/2^n) + \sum_{j=0}^{n-1} (2^{-1}\alpha)^j \log \psi(2^{-j}\varepsilon).$$

Let  $n \rightarrow \infty$  and use the estimate  $\log \phi(\varepsilon/2^n) \leq \log(C2^{n\gamma}/\varepsilon^\gamma)$  and the fact that  $\alpha < 2$ , we arrive at the estimate

$$(3.10) \quad \begin{aligned} \log \phi(\varepsilon) &\leq \sum_{j=0}^{\infty} (2^{-1}\alpha)^j \log \psi(2^{-j}\varepsilon) \\ &= \frac{2}{2-\alpha} \log \psi(\varepsilon) + \sum_{j=0}^{\infty} (2^{-1}\alpha)^j \log \frac{\psi(2^{-j}\varepsilon)}{\psi(\varepsilon)}. \end{aligned}$$

The term

$$\psi(2^{-j}\varepsilon)/\psi(\varepsilon) = 2^{\alpha j}J(2^j\varepsilon^{-1})/J(\varepsilon^{-1})$$

can be estimated by the Potter bound on the slowly varying function  $J(x)$ ; see Bingham, Goldie and Teugels (1987), page 25. Namely, we have for  $\varepsilon > 0$  small,

$$J(2^j\varepsilon^{-1})/J(\varepsilon^{-1}) \leq 2^{j+1}$$

and thus

$$\sum_{j=0}^{\infty} (2^{-1\alpha})^j \log \frac{\psi(2^{-j}\varepsilon)}{\psi(\varepsilon)} \leq \sum_{j=0}^{\infty} (\alpha/2)^j \log 2^{\alpha j+j+1} < \infty.$$

So by the Lebesgue dominated convergence theorem and the fact that  $J(x)$  is a slowly varying function, the sum on the right hand of (3.10) tends to a finite constant, say  $c > 0$ , as  $\varepsilon \rightarrow 0$ . Hence for  $\varepsilon > 0$  small, we have from (3.10),

$$\log \phi(\varepsilon) \leq \frac{2}{2-\alpha} \log \psi(\varepsilon) + 2c,$$

which implies our result by recalling  $\psi(\varepsilon) = C\varepsilon^{-\alpha}J(1/\varepsilon)$ . And thus the proof is complete.  $\square$

PROOF OF THEOREM 1.2. Combining our previous results, we easily prove Theorem 1.2 as follows. Since for  $0 < \alpha < 2$ ,  $H(K_X, \varepsilon) \asymp \varepsilon^{-\alpha}J(1/\varepsilon)$ , there exists  $\alpha' < 2$  such that  $H(K_X, \varepsilon) \asymp \varepsilon^{-\alpha'}$ . Thus by Proposition 2.4, we see that the extra condition (3.1) in Proposition 3.1 is satisfied. This completes the proof of Theorem 1.2.  $\square$

**4.  $l$ -Approximation numbers and small ball estimates.** The results of Section 2 show the existence of a tight connection between the behaviors of small ball probabilities of a Gaussian random variable  $X$  and of its  $l$ -approximation numbers  $l_n(X)$ . Using Theorem 1.2, this connection can now be made more precise.

PROPOSITION 4.1. *Let  $X$  be a centered Gaussian random variable.*

(a) *If*

$$(4.1) \quad l_n(X) \asymp n^{-1/\alpha}(1 + \log n)^\beta$$

*for some  $\alpha > 0$  and  $\beta \in \mathbb{R}$ , then*

$$(4.2) \quad -\log \mathbb{P}(\|X\| \leq \varepsilon) \asymp \varepsilon^{-\alpha}(\log 1/\varepsilon)^{\alpha\beta}.$$

(b) *Conversely, if (4.2) holds for some  $\alpha > 0$  and  $\beta \in \mathbb{R}$ , this implies*

$$(4.3) \quad l_n(X) \asymp n^{-1/\alpha}(1 + \log n)^{\beta+1}.$$

*Moreover, if  $E$  is  $K$ -convex, then*

$$l_n(X) \asymp n^{-1/\alpha}(1 + \log n)^\beta$$

and thus (4.1) and (4.2) are equivalent in this case.

(c) If

$$-\log \mathbb{P}(\|X\| \leq 2\varepsilon) \geq -\log \mathbb{P}(\|X\| \leq \varepsilon) \geq \varepsilon^{-\alpha} (\log 1/\varepsilon)^{\alpha\beta}$$

for some  $\alpha > 0$  and  $-\infty < \beta < \infty$ , then

$$l_n(X) \geq n^{-1/\alpha} (1 + \log n)^{\beta-1/\alpha}.$$

PROOF. Part (b) is a direct consequence of Proposition 2.2 and Theorem 1.1(ii). Assertion (a) follows from Theorem 1.2 (or Theorem 5.1) by a quite general relation between  $e_n(u)$  and  $l_n(u)$  for operators  $u$  from  $H$  into  $E$ . For the sake of completeness, let us state and prove it here. For  $u$  as before, its  $n$ th Kolmogorov number  $d_n(u)$  is defined by

$$d_n(u) = \inf\{\varepsilon > 0: \exists E_0 \subseteq E \text{ subspace, } \dim(E_0) < n, u(B_H) \subseteq E_0 + \varepsilon B_E\},$$

where  $B_H$  and  $B_E$  are the closed unit balls of  $H$  and  $E$ , respectively. A basic result of Pajor, and Tomczak-Jaegermann (1985) asserts that  $\sup_n n^{1/2} d_n(u) \leq cl(u)$  for some universal  $c > 0$ . Using  $d_n(v) = 0$  for  $\text{rk}(v) < n$ , standard methods [cf. Pisier (1989), page 148] lead to the estimate  $n^{1/2} d_{2n}(u) \leq l_n(u)$ . Now Theorem 1.3 of Carl, Kyrkezi and Pajor (1997) relates  $e_n(u)$  and  $d_n(u)$  as follows. For any increasing sequence  $(b_n)_{n=1}^\infty$  of positive numbers with  $\sup_n b_{2n}/b_n = \kappa < \infty$  it holds  $\sup_n b_n e_n(u) \leq c \sup_n b_n d_n(u)$  with  $c > 0$  only depending on  $\kappa$ . Consequently,

$$(4.4) \quad \sup_n n^{1/2} b_n e_n(u) \leq c \sup_n b_n l_n(u)$$

for  $b_n = n^{1/\alpha} (1 + \log n)^{-\beta}$ . Hence (4.1) and (4.4) yield

$$e_n(u) \leq n^{-(2+\alpha)/(2\alpha)} (1 + \log n)^\beta,$$

which implies (a) by Theorem 1.2 or Theorem 5.1.

Turn to (c): let  $u: H \rightarrow E$  be an operator such that  $X = \sum_j \xi_j u(f_j)$  converges a.s. in  $E$ . Using (i) of Theorem 1.1, we have

$$e_n(u) \geq n^{-(2+\alpha)/(2\alpha)} (1 + \log n)^\beta.$$

Given  $n \in \mathbb{N}$ , there exists a finite rank operator  $v_n: H \rightarrow E$  such that  $\text{rk}(v_n) < n$  and  $l(u - v_n) \leq 2l_n(u)$ . For any  $N \in \mathbb{N}$  we have

$$(4.5) \quad N^{1/2} e_{2N-1}(u) \leq N^{1/2} e_N(u - v_n) + N^{1/2} e_N(v_n).$$

The first term on the right of (4.5) can be estimated by Sudakov's minoration to obtain

$$N^{1/2} e_N(u - v_n) \leq cl(u - v_n) \leq cl_n(u).$$

For the second term on the right of (4.5) we use [cf. Carl and Stephani (1990), 1.3.36]

$$N^{1/2} e_N(v_n) \leq 4N^{1/2} 2^{-(N-1)/n} \|v_n\|.$$

Observe that

$$\|v_n\| \leq l(v_n) \leq l(u - v_n) + l(u) \leq 2l_n(u) + l(u) \leq 3l(u).$$

so  $\sup_n \|v_n\| < \infty$ . Now let  $N = [C \cdot n \log n]$  with the constant  $C$  large enough. Then it is easy to see that  $4N^{1/2} 2^{-(N-1)/n}$  is of smaller order, compared with  $N^{-1/\alpha}(\log N)^\beta \leq N^{1/2} e_{2N-1}(u)$ . Thus we obtain

$$l_n(u) \geq N^{-1/\alpha}(\log N)^\beta \geq n^{-1/\alpha}(1 + \log n)^{\beta-1/\alpha},$$

which completes the proof.  $\square$

Note that  $l_n(u) \leq n^{-1/\alpha}$  iff there are operators  $v_n: H \rightarrow E$  with  $\text{rk}(v_n) \leq 2^n$  and  $l(v_n) \leq 2^{-n/\alpha}$  such that  $u = \sum_{n=1}^\infty v_n$  [cf. Pietsch (1987), 2.8.8]. Thus  $l_n(X) \leq n^{-1/\alpha}$  iff there are Gaussian random variables  $X_n$  which are not necessarily independent, map into an at most  $2^n$ -dimensional subspace of  $E$  with probability 1, such that

$$X = \sum_{n=1}^\infty X_n \quad \text{and} \quad \mathbb{E}\|X_n\|^2 \leq 2^{-2n/\alpha}.$$

There are several natural and important questions left open here in connection with the above proposition. Does  $l_n(X) \geq n^{-1/\alpha}$  imply a lower estimate for  $-\log \mathbb{P}(\|X\| \leq \varepsilon)$  or, equivalently, does  $l_n(u) \geq n^{-1/\alpha}$  imply a lower estimate for  $e_n(u)$ ? Is the power  $\beta + 1$  of the log-term in (4.3) optimal?

**5. Some algebraic properties of small ball probabilities.** We start with some general applications of the relation between entropy and small ball behavior. Let us begin with a reformulation of Theorem 1.2 and Theorem 1.1(ii) in the language of operators.

**THEOREM 5.1.** *Let  $u: H \rightarrow E$  be an operator. Then for  $0 < \alpha < 2$  and  $\beta \in \mathbb{R}$ , the following are equivalent:*

- (i)  $e_n(u) \leq n^{-1/\alpha}(1 + \log n)^\beta,$
- (ii)  $-\log \mathbb{P}\left(\left\|\sum_{j=1}^\infty \xi_j u(f_j)\right\| \leq \varepsilon\right) \leq \varepsilon^{-2\alpha/(2-\alpha)}\left(\log \frac{1}{\varepsilon}\right)^{2\alpha\beta/(2-\alpha)}.$

If we define

$$\mathcal{E}_\alpha(u) = \sup_n n^{1/\alpha} e_n(u),$$

then it is well known [cf. Pietsch (1987), 2.2.5] that  $\mathcal{E}_\alpha$  is a quasi-norm on the set

$$E_\alpha = \{u: H \rightarrow E: \mathcal{E}_\alpha(u) < \infty\}$$

and  $(E_\alpha, \mathcal{E}_\alpha)$  is a complete quasi-normed space. Furthermore, if we set

$$\mathcal{F}_\alpha(u) = \sup_{\varepsilon > 0} \varepsilon \left[ -\log \mathbb{P} \left( \left\| \sum_{j=1}^{\infty} \xi_j u(f_j) \right\| \leq \varepsilon \right) \right]^{(2-\alpha)/2\alpha},$$

then by Theorem 5.1, we get  $\mathcal{E}_\alpha(u) < \infty$  iff  $\mathcal{F}_\alpha(u) < \infty$ ,  $0 < \alpha < 2$ . So the following question is of interest and natural. Do there exist universal constants  $c_1, c_2 > 0$ , which may depend on  $0 < \alpha < 2$ , such that for any  $u$  from  $H$  into  $E$  we always have

$$c_1 \mathcal{F}_\alpha(u) \leq \mathcal{E}_\alpha(u) \leq c_2 \mathcal{F}_\alpha(u)?$$

One possible approach to solve this question would be a closed graph argument. But in order to apply the argument,  $\mathcal{F}_\alpha$  needs to satisfy a generalized triangle inequality. That is, as a first step, we would like to have for operators  $u$  and  $v$ ,

$$(5.1) \quad \mathcal{F}_\alpha(u + v) \leq c(\mathcal{F}_\alpha(u) + \mathcal{F}_\alpha(v)),$$

where  $c > 0$  is a universal constant. The inequality (5.1) can be proved by assuming an affirmative answer to the Gaussian correlation conjecture which states that  $\mu(A \cap B) \geq \mu(A)\mu(B)$  for all symmetric convex sets  $A$  and  $B$ . Other equivalent formulations, early history and recent progress of the conjecture can be found in Schechtman, Schlumprecht and Zinn (1998).

The following weaker form of (5.1) is a direct consequence of Theorem 5.1.

**COROLLARY 5.1.** *Let  $(x_j)_{j=1}^\infty$  and  $(y_j)_{j=1}^\infty$  be elements in  $E$  such that  $\sum_j \xi_j x_j$  and  $\sum_j \xi_j y_j$  exist a.s. in  $E$ . If for some  $\alpha > 0$  and  $\beta \in \mathbb{R}$ ,*

$$-\log \mathbb{P} \left( \left\| \sum_{j=1}^{\infty} \xi_j x_j \right\| \leq \varepsilon \right) \leq c \varepsilon^{-\alpha} \left( \log \frac{1}{\varepsilon} \right)^\beta$$

and

$$-\log \mathbb{P} \left( \left\| \sum_{j=1}^{\infty} \xi_j y_j \right\| \leq \varepsilon \right) \leq c' \varepsilon^{-\alpha} \left( \log \frac{1}{\varepsilon} \right)^\beta,$$

then

$$(5.2) \quad -\log \mathbb{P} \left( \left\| \sum_{j=1}^{\infty} \xi_j (x_j + y_j) \right\| \leq \varepsilon \right) \leq c'' \varepsilon^{-\alpha} \left( \log \frac{1}{\varepsilon} \right)^\beta.$$

**PROOF.** This easily follows from Theorem 5.1 and the fact that

$$e_{n+m-1}(u + v) \leq e_n(u) + e_m(v)$$

for  $m, n \in \mathbb{N}$  and  $u, v: H \rightarrow E$  [cf. Pisier (1989), (5.6)]. Note that by assuming a positive answer to the Gaussian correlation conjecture, a direct proof of (5.2) is standard and the constant  $c''$  can be expressed as a function of  $c$  and  $c'$  along with  $\alpha$  and  $\beta$ .  $\square$

Another useful property of entropy numbers is their multiplicative property in the following sense. If  $u: E \rightarrow E_0$  and  $v: E_0 \rightarrow F$ , then

$$(5.3) \quad e_{n+m-1}(v \circ u) \leq e_n(v)e_m(u)$$

for  $m, n \in \mathbb{N}$  [cf. Pisier (1989), (5.4)]. This leads to the following very useful result.

**THEOREM 5.2.** *Let  $X$  be an  $E$ -valued Gaussian random element and suppose that*

$$-\log \mathbb{P}(\|X\| \leq \varepsilon) \leq \varepsilon^{-\alpha} \left( \log \frac{1}{\varepsilon} \right)^\beta$$

for some  $\alpha > 0$  and  $-\infty < \beta < \infty$ . If  $v$  is an operator from  $E$  into another Banach space  $F$  with

$$(5.4) \quad e_n(v) \leq n^{-1/\gamma} (1 + \log n)^\rho,$$

for some  $\gamma > 0$  and  $-\infty < \rho < \infty$ , then

$$(5.5) \quad -\log \mathbb{P}(\|v(X)\| \leq \varepsilon) \leq \varepsilon^{-\alpha\gamma/(\alpha+\gamma)} \left( \log \frac{1}{\varepsilon} \right)^{(\alpha\rho+\beta)\gamma/(\alpha+\gamma)}.$$

**PROOF.** Let  $u: H \rightarrow E$  be such that  $\sum_j \xi_j u(f_j)$  converges a.s. in  $E$  and  $X = \sum_j \xi_j u(f_j)$ . Then  $v(X) = \sum_{j=1}^\infty \xi_j v(uf_j)$  and the result follows from (5.3) and Theorem 5.1.  $\square$

**6. Applications to Gaussian processes.** Let  $(T, d)$  be a compact metric space and let  $X = (X(t))_{t \in T}$  be a centered Gaussian process with a.s. continuous sample paths. Then all our previous results and notations apply to  $X$  and we may regard  $X$  as a  $C(T)$ -valued Gaussian random variable where  $C(T)$  denotes the Banach space of real-valued continuous functions on  $T$ .

We first illustrate how to apply Theorem 5.2 in concrete situations. Let  $Y = (Y(t))_{t \in [0,1]}$  be a centered Gaussian process possessing a.s. continuous sample paths and let  $K: [0, 1]^2 \rightarrow \mathbb{R}$  be a measurable kernel such that  $I_K$  defined by

$$(I_K f)(t) = \int_0^1 K(t, s) f(s) ds$$

is a (bounded) operator from  $C[0, 1]$  into itself. Note that if

$$\lim_{t' \rightarrow t} \int_0^1 |K(t, s) - K(t', s)| ds = 0$$

and

$$\sup_{t \in [0, 1]} \int_0^1 |K(t, s)| ds < \infty,$$

then  $I_K$  is a (bounded) operator from  $C[0, 1]$  into itself.

Next, a continuous centered Gaussian stochastic process  $X = I_K(Y)$  can be constructed by

$$(6.1) \quad X(t) = \int_0^1 K(t, s)Y(s) ds.$$

Then Theorem 5.2 yields the following very general result for the lower bound of small ball probability of  $X$ .

**THEOREM 6.1.** *If for  $\alpha > 0, \gamma > 0$  and  $\beta, \rho \in \mathbb{R}$ ,*

$$-\log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |Y(t)| \leq \varepsilon\right) \leq \varepsilon^{-\alpha} \left(\log \frac{1}{\varepsilon}\right)^\beta$$

and

$$e_n(I_K) \leq n^{-1/\gamma} (1 + \log n)^\rho,$$

then

$$-\log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |X(t)| \leq \varepsilon\right) \leq \varepsilon^{-\alpha\gamma/(\alpha+\gamma)} \left(\log \frac{1}{\varepsilon}\right)^{(\alpha\rho+\beta)\gamma/(\alpha+\gamma)}.$$

In particular, if the kernel  $K$  in (6.1) satisfies the Hölder condition

$$(6.2) \quad \int_0^1 |K(t, s) - K(t', s)| ds \leq c|t - t'|^\lambda, \quad t, t' \in [0, 1],$$

for some  $\lambda \in (0, 1]$  and some  $c > 0$ , then

$$-\log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |X(t)| \leq \varepsilon\right) \leq \varepsilon^{-\alpha/(\alpha\lambda+1)} \left(\log \frac{1}{\varepsilon}\right)^{\beta/(\alpha\lambda+1)}.$$

**PROOF.** The first part follows directly from Theorem 5.2. Now if condition (6.2) holds, then  $I_K$  maps  $C[0, 1]$  into the Hölder space

$$C^\lambda[0, 1] = \left\{ f \in C[0, 1] : \sup_{t \neq t'} \frac{|f(t) - f(t')|}{|t - t'|^\lambda} < \infty \right\}.$$

For the natural embedding  $J_\lambda$  from  $C^\lambda[0, 1]$  into  $C[0, 1]$  one has [cf. Timan (1964)]

$$e_n(J_\lambda) \approx n^{-\lambda},$$

thus

$$e_n(I_K : C \rightarrow C) \leq e_n(J_\lambda) \|I_K : C \rightarrow C^\lambda\| \leq n^{-\lambda},$$

which finishes the proof. It should be noted that a direct probabilistic proof under (6.2) seems very unlikely. On the other hand, the result is almost trivial from the point of view of entropy numbers for operators.  $\square$

Next we will study some specified integral operators and the Gaussian processes generated by them. For  $\beta > 0$  the Riemann–Liouville integral



operator  $R_\beta$  is defined by

$$(R_\beta f)(t) = \frac{1}{\Gamma(\beta)} \int_0^t (t-s)^{\beta-1} f(s) ds, \quad 0 \leq t \leq 1.$$

It is a bounded operator from  $C[0, 1]$  into  $C[0, 1]$  and for  $\beta > 1/2$  it maps  $L_2[0, 1]$  into  $C[0, 1]$  continuously. Note in particular that for any integer  $k \geq 1$ ,

$$\begin{aligned} (R_k f)(t) &= \frac{1}{(k-1)!} \int_0^t (t-s)^{k-1} f(s) ds \\ &= \int_0^t \int_0^{t_1} \cdots \int_0^{t_{k-1}} f(s) ds dt_{k-1} \cdots dt_1 \end{aligned}$$

and  $R_0$  can be interpreted as the identity map. Below are some analytic facts we need.

PROPOSITION 6.1.

- (a)  $R_\alpha \circ R_\beta = R_{\alpha+\beta}$  for  $\alpha, \beta > 0$ .
- (b)  $e_n(R_\beta: C \rightarrow C) \approx n^{-\beta}$  for  $\beta > 0$ .
- (c)  $e_n(R_\beta: L_2 \rightarrow C) \approx n^{-\beta}$  if  $\beta > 1/2$ .

PROOF. Part (a) is well known and immediate to see, so we omit the proof. To verify (b), it is easily checked that for  $0 < \beta \leq 1$  the operator  $R_\beta$  maps  $C[0, 1]$  even into  $C^\beta[0, 1]$ . So the desired upper estimate for  $0 < \beta \leq 1$  follows from the arguments used in the proof of Theorem 6.1. If  $\beta > 1$ , choose  $k \in \mathbb{N}$  with  $\alpha = \beta/k < 1$ . Since

$$R_\beta = \underbrace{R_\alpha \circ \cdots \circ R_\alpha}_{k \text{ times}},$$

the multiplicative property of the entropy numbers yields

$$e_{nk-(k-1)}(R_\beta) \leq e_n(R_\alpha)^k \leq c^k (n^{-\alpha})^k = c' n^{-\beta},$$

which is the desired upper bound.

For the lower bound we first investigate the case  $\beta = k \in \mathbb{N}$ . To do so let  $f$  be an arbitrary  $k$ -times continuously differentiable function on  $\mathbb{R}$  with  $\text{supp}(f) \subset (0, 1)$  and

$$\|f^{(k)}(t)\|_C = \sup_t |f^{(k)}(t)| = 1.$$

Given  $n \in \mathbb{N}$ , we define functions  $f_j, 1 \leq j \leq n$ , on  $[0, 1]$  by

$$f_j(t) = f(nt - (j-1)).$$

Hence,  $\text{supp}(f_j) \subset ((j-1)/n, j/n)$  and  $\|f_j^{(k)}\|_C \leq n^k$ . For  $\delta = (\delta_1, \dots, \delta_n) \in \{0, 1\}^n$ , put

$$g_\delta(t) = \sum_{j=1}^n \delta_j f_j^{(k)}(t).$$

Then we have  $\|g_\delta\|_C \leq n^k$  as well as

$$\|R_k(g_\delta) - R_k(g_{\delta'})\| \geq \|f\|_C > 0$$

for  $\delta \neq \delta'$  which yields

$$e_{n+1}(R_k) \geq (\|f\|_C/2)n^{-k}$$

as asserted.

When  $\beta > 0$  is arbitrary, let  $k \in \mathbb{N}$  be any integer greater than  $\beta$ . Then  $R_k = R_\beta \circ R_{k-\beta}$  and if we combine the lower estimate for integers with the upper one for any positive number, then we get

$$c(2n - 1)^{-k} \leq e_{2n-1}(R_k) \leq e_n(R_\beta)e_n(R_{k-\beta}) \leq e_n(R_\beta)c' \cdot n^{-(k-\beta)},$$

completing the proof of (b).

To prove (c) we have to regard now  $R_\beta$  as the operator from  $L_2[0, 1]$  into  $C[0, 1]$ . If  $J: C[0, 1] \rightarrow L_2[0, 1]$  is the natural embedding, then

$$e_n(R_\beta: C \rightarrow C) \leq \|J\|e_n(R_\beta: L_2 \rightarrow C);$$

hence  $e_n(R_\beta: L_2 \rightarrow C) \geq n^{-\beta}$  by part (b) of the proposition.

The converse estimate is first proved for  $\beta \geq 1$ . It is well known (and follows also from Theorem 1.1 and the well-known small ball behavior of the Wiener measure in  $C[0, 1]$ ) that  $e_n(R_1: L_2 \rightarrow C) \approx n^{-1}$ . Thus for  $\beta > 1$  the desired upper estimate follows from

$$\begin{aligned} e_{2n-1}(R_\beta: L_2 \rightarrow C) &\leq e_n(R_{\beta-1}: C \rightarrow C)e_n(R_1: L_2 \rightarrow C) \\ &\leq c \cdot n^{-(\beta-1)}n^{-1} = c \cdot n^{-\beta}. \end{aligned}$$

Before we treat the remaining case  $1/2 < \beta < 1$ , we recall some basic facts about fractional Brownian motion. Let  $\{B_\gamma(t): t \geq 0\}$  denote the  $\gamma$ -fractional Brownian motion with  $B_\gamma(0) = 0$  and  $0 < \gamma < 2$ . Then  $\{B_\gamma(t): t \geq 0\}$  is a Gaussian process with mean zero and covariance function

$$\mathbb{E}(B_\gamma(t)B_\gamma(s)) = \frac{1}{2}(|s|^\gamma + |t|^\gamma - |s - t|^\gamma).$$

It is known [cf. Shao (1993) and Monrad and Rootzén (1995)] that

$$(6.3) \quad -\log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |B_\gamma(t)| \leq \varepsilon\right) \approx \varepsilon^{-2/\gamma}.$$

The fractional Brownian motion is related to the Riemann–Liouville integral operator as follows. For  $\gamma \in (0, 2)$  we define an operator

$$Q_\gamma: L_2(-\infty, 0] \rightarrow C[0, 1]$$

by

$$(6.4) \quad (Q_\gamma f)(t) = \frac{1}{\Gamma((\gamma + 1)/2)} \int_{-\infty}^0 ((t - s)^{(\gamma-1)/2} - (-s)^{(\gamma-1)/2})f(s) ds.$$

Then

$$S_\gamma = R_{(\gamma+1)/2} \oplus Q_\gamma: L_2[0, 1] \oplus L_2(-\infty, 0] \rightarrow C[0, 1]$$

generates a multiple of  $B_\gamma$  [cf. Mandelbrot and Van Ness (1968)]. That is, the representation

$$(6.5) \quad B_\gamma(t) = a_\gamma \sum_{j=1}^\infty \xi_j(S_\gamma f_j)(t), \quad t \in [0, 1]$$

holds for one (each) ONB  $(f_j)_{j=1}^\infty$  in  $L_2(-\infty, 1]$  where

$$(6.6) \quad a_\gamma = \Gamma((\gamma + 1)/2) \left( \gamma^{-1} + \int_{-\infty}^0 ((1 - s)^{(\gamma-1)/2} - (-s)^{(\gamma-1)/2})^2 ds \right)^{-1/2}.$$

Now we can finish the proof of the proposition. If  $1/2 < \beta < 1$ , define  $\gamma \in (0, 1)$  by  $\gamma = 2\beta - 1$ . The construction of  $S_\gamma$  easily yields

$$(6.7) \quad e_n(R_{(\gamma+1)/2}) \leq e_n(S_\gamma).$$

On the other hand, Theorem 1.1 and (6.3) imply  $e_n(S_\gamma) \leq n^{-(\gamma+1)/2}$  which completes the proof by combining with (6.7).  $\square$

Another possible approach to part (c) of Proposition 6.1 is as follows. One verifies that  $R_\beta$  is an isometry from  $L_2[0, 1]$  onto the Besov space  $B_{2,2}^\beta(0, 1)$ . Then (c) follows from the known behavior [cf. Edmunds and Triebel (1996)] of the entropy numbers of the embedding of  $B_{2,2}^\beta(0, 1)$  into  $C[0, 1]$ ,  $1/2 < \beta < \infty$ .

Now for  $\beta > 1/2$  and any fixed ONB  $(f_j)_{j=1}^\infty$  in  $L_2[0, 1]$ , we can define a centered Gaussian process  $W_\beta$  by

$$(6.8) \quad W_\beta(t) = \sum_{j=1}^\infty \xi_j(R_\beta f_j)(t), \quad t \in [0, 1].$$

In view of the fact that  $e_n(R_\beta) \approx n^{-\beta}$ ,  $\beta > 1/2$ , it is standard to see [cf. Dudley (1967) and Lemma 2.2] that the right-hand sum of (6.8) converges a.s. in  $C[0, 1]$  and thus  $W_\beta$  possesses a.s. continuous sample paths. Observe that  $W_1$  is the Wiener process  $W$  on  $[0, 1]$  and for  $0 < \gamma < 2$  the process  $W_{(\gamma+1)/2}$  is closely related to the fractional Brownian motion  $B_\gamma$ . The following results provide desired small ball estimates for  $W_\beta$ .

**THEOREM 6.2.**

(a) *If  $\alpha > 0$  and  $\beta > 1/2$ , then*

$$W_{\alpha+\beta}(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} W_\beta(s) ds.$$

*Especially, for  $\beta > 1$  we have*

$$W_\beta(t) = \frac{1}{\Gamma(\beta-1)} \int_0^t (t-s)^{\beta-2} W(s) ds.$$

(b) *It holds*

$$-\log \mathbb{P} \left( \sup_{0 \leq t \leq 1} |W_\beta(t)| \leq \varepsilon \right) \approx \varepsilon^{-2/(2\beta-1)}.$$

PROOF. By definition of  $W_\beta$ , the assertions follow directly from Proposition 6.1(a) and Theorem 5.1 and Proposition 6.1(c).  $\square$

Note that by refined methods [cf. Li and Linde (1998)] one can even show that

$$(6.9) \quad \lim_{\varepsilon \rightarrow 0} \varepsilon^{2/(2\beta-1)} \log \mathbb{P} \left( \sup_{0 \leq t \leq 1} |W_\beta(t)| \leq \varepsilon \right) = -k_\beta$$

with  $0 < k_\beta < \infty$ ,  $\beta > 1/2$ . In particular, when  $\beta = 2$ , then  $W_2$  is the integrated Wiener process and the above result was recently proved by using local time techniques [cf. Khoshnevisan and Shi (1998)].

Next we will study the fractional integrated fractional Brownian motion based on the relation between  $W_\beta$  and the fractional Brownian motion  $B_\gamma$  of order  $\gamma = 2\beta - 1$ . This relation was already used in the proof of Proposition 6.1 and will now be made more precise.

Let  $Q_\gamma: L_2(-\infty, 0] \rightarrow C[0, 1]$  be defined by (6.4) for  $0 < \gamma < 2$  and let  $Z_\gamma$  be the generated centered Gaussian process; that is,

$$Z_\gamma(t) = \sum_{j=1}^{\infty} \xi_j(Q_\gamma f_j)(t), \quad t \in [0, 1],$$

for some fixed ONB  $(f_j)_{j=1}^\infty$  in  $L_2(-\infty, 0]$ . Then the representation (6.5) can be rewritten as

$$(6.10) \quad \begin{aligned} B_\gamma(t) &= a_\gamma \left( \sum_{j=1}^{\infty} \xi'_j(R_{(\gamma+1)/2} f'_j)(t) + \sum_{j=1}^{\infty} \xi_j(Q_\gamma f_j)(t) \right) \\ &= a_\gamma (W_{(\gamma+1)/2}(t) + Z_\gamma(t)) \end{aligned}$$

for some ONB  $(f'_j)_{j=1}^\infty$  in  $L_2[0, 1]$  and  $(\xi'_j)_{j=1}^\infty$  is independent of  $(\xi_j)_{j=1}^\infty$ . It is clear from the above construction that the centered Gaussian process  $Z_\gamma$  is independent of  $W_{(\gamma+1)/2}$ . Now we are in position to obtain the following result.

**THEOREM 6.3.** For  $0 < \gamma < 2$  and  $\alpha > 0$ , let  $B_{\gamma, \alpha}$  be defined as

$$B_{\gamma, \alpha}(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} B_\gamma(s) ds, \quad t \in [0, 1].$$

Then

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{2/(\gamma+2\alpha)} \log \mathbb{P} \left( \sup_{0 \leq t \leq 1} |B_{\gamma, \alpha}(t)| \leq \varepsilon \right) = -a_\gamma^{2/(\gamma+2\alpha)} k_{\alpha+(\gamma+1)/2}$$

where  $a_\gamma$  is given in (6.6) and  $k_{\alpha+(\gamma+1)/2}$  is the constant in (6.9).

PROOF. Note that (6.10) implies

$$B_{\gamma, \alpha} = a_\gamma (W_{\alpha+(\gamma+1)/2} + R_\alpha(Z_\gamma))$$

and  $R_\alpha(Z_\gamma)$  is independent of  $W_{\alpha+(\gamma+1)/2}$ . Using Anderson's inequality [see

Anderson (1955)], it follows that

$$\mathbb{P}\left(\sup_{0 \leq t \leq 1} |B_{\gamma, \alpha}(t)| \leq \varepsilon\right) \leq \mathbb{P}\left(\sup_{0 \leq t \leq 1} |a_\gamma W_{\alpha+(\gamma+1)/2}(t)| \leq \varepsilon\right),$$

and thus by (6.9),

$$\limsup_{\varepsilon \rightarrow 0} \varepsilon^{2/(\gamma+2\alpha)} \log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |B_{\gamma, \alpha}(t)| \leq \varepsilon\right) \leq -a_\gamma^{2/(\gamma+2\alpha)} k_{\alpha+(\gamma+1)/2}.$$

On the other hand, from the proof of Lemma 3.2 in Li and Linde (1998) we have

$$-\log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |Z_\gamma(t)| \leq \varepsilon\right) \leq \varepsilon^{-1} \log(1/\varepsilon)$$

and thus by Theorem 6.1,

$$-\log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |R_\alpha(Z_\gamma)(t)| \leq \varepsilon\right) \leq \varepsilon^{-1/(\alpha+1)} (\log 1/\varepsilon)^{1/(\alpha+1)},$$

which implies

$$\lim_{\varepsilon \rightarrow 0} \varepsilon^{2/(\gamma+2\alpha)} \log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |R_\alpha(Z_\gamma)(t)| \leq \varepsilon\right) = 0.$$

To obtain the lower bound of Theorem 6.3, we have for any  $0 < \delta < 1$ ,

$$\begin{aligned} &\mathbb{P}\left(\sup_{0 \leq t \leq 1} |B_{\gamma, \alpha}(t)| \leq \varepsilon\right) \\ &\geq \mathbb{P}\left(\sup_{0 \leq t \leq 1} |W_{\alpha+(\gamma+1)/2}(t)| \leq (1 - \delta)\varepsilon/\alpha_\gamma\right) \\ &\quad \times \mathbb{P}\left(\sup_{0 \leq t \leq 1} |R_\alpha(Z_\gamma)(t)| \leq \delta\varepsilon/\alpha_\gamma\right) \end{aligned}$$

since  $W_{\alpha+(\gamma+1)/2}(t)$  and  $R_\alpha(Z_\gamma)(t)$  are independent of each other. Thus

$$\begin{aligned} &\liminf_{\varepsilon \rightarrow 0} \varepsilon^{2/(\gamma+2\alpha)} \log \mathbb{P}\left(\sup_{0 \leq t \leq 1} |B_{\gamma, \alpha}(t)| \leq \varepsilon\right) \\ &\geq -k_{\alpha+(\gamma+1)/2} (1 - \delta)^{-2/(\gamma+2\alpha)} \alpha_\gamma^{2/(\gamma+2\alpha)}. \end{aligned}$$

So we obtain the desired lower bound by taking  $\delta \rightarrow 0$ . This completes the proof.  $\square$

**Acknowledgments.** The authors are grateful for the support of the RiP program in Oberwolfach in which this work originated. We thank Professor M. Lifshits for careful reading of the paper and providing many helpful suggestions.

### REFERENCES

ANDERSON, T. W. (1955). The integral of symmetric unimodular functions over a symmetric convex set and some probability inequalities. *Proc. Amer. Math. Soc.* **6** 170–176.

- BELINSKY, E. S. (1998). Estimates of entropy numbers and Gaussian measures for classes of functions with bounded mixed derivative. *J. Approx. Theory* **93** 114–127.
- BINGHAM, N. H., GOLDIE, C. M. and TEUGELS, J. L. (1987). *Regular Variation*. Cambridge Univ. Press.
- BOURGAIN, J., PAJOR, A., SZAREK, S. and TOMCZAK-JAEGERMANN, N. (1989). On the duality problem for entropy numbers of operators. *Geometric Aspects of Functional Analysis. Lecture Notes in Math.* **1376** 50–63. Springer, Berlin.
- CARL, B., KYREZI, I. and PAJOR, A. (1997). Metric entropy of convex hulls in Banach spaces. *J. London Math. Soc.* To appear.
- CARL, B. and STEPHANI, I. (1990). *Entropy, Compactness and Approximation of Operators*. Cambridge Univ. Press.
- DUDLEY, R. M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Funct. Anal.* **1** 290–330.
- DUNKER, T. (1998). Estimates for the small ball probabilities of the fractional Brownian sheet. *J. Theoret. Probab.* To appear.
- DUNKER, T., KÜHN, T., LIFSHITS, M. A. and LINDE, W. (1998). Metric entropy of integration operator and small ball probabilities of the Brownian sheet. *C. R. Acad. Sci. Paris* **326** 347–352.
- EDMUNDS, D. E. and TRIEBEL, H. (1996). *Function Spaces, Entropy Numbers and Differential Operators*. Cambridge Univ. Press.
- GOODMAN, V. (1990). Some probability and entropy estimates for Gaussian measures. In *Progr. Probab.* **20** 150–156.
- HOFFMANN-JØRGENSEN, J., DUDLEY, R. M. and SHEPP, L. A. (1979). On the lower tail of Gaussian seminorms. *Ann. Probab.* **7** 319–342.
- KHOSHNEVISAN, D. and SHI, Z. (1998). Chung's law for integrated Brownian motion. *Trans. Amer. Math. Soc.* **350** 4253–4264.
- KUELBS, J. (1976). A strong convergence theorem for Banach space valued random variables. *Ann. Probab.* **4** 744–771.
- KUELBS, J. and LI, W. V. (1993). Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* **116** 133–157.
- LEDOUX, M. (1996). Isoperimetry and Gaussian analysis. *Lectures on Probability Theory and Statistics. Lecture Notes in Math.* **1648** 165–294. Springer, Berlin.
- LEDOUX, M. and TALAGRAND, M. (1991). *Probability in a Banach Space*. Springer, Berlin.
- LI, W. V. (1992). On the lower tail of Gaussian measures on  $l^p$ . *Progr. Probab.* **30** 106–117.
- LI, W. V. (1998). Small deviations for Gaussian Markov processes under the sup-norm. *J. Theoret. Probab.* To appear.
- LI, W. V. and LINDE, W. (1998). Existence of small ball constants for fractional Brownian motions. *C. R. Acad. Sci. Paris* **326** 1329–1334.
- LINDE, W. and PIETSCH, A. (1974). Mappings of Gaussian measures in Banach spaces. *Theoret. Probab. Appl.* **19** 445–460.
- MANDELBROT, B. B. and VAN NESS, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review* **10** 422–437.
- MONRAD, D. and ROOTZÉN, H. (1995). Small values of Gaussian processes and functional laws of the iterated logarithm. *Probab. Theory Related Fields* **101** 173–192.
- PAJOR, A. and TOMCZAK-JAEGERMANN, N. (1985). Remarques sur les nombres d'entropie d'un opérateur et de son transposé. *C. R. Acad. Sci. Paris* **301** 743–746.
- PIETSCH, A. (1987). *Eigenvalues and s-Numbers*. Cambridge Univ. Press.
- PISIER, G. (1989). *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge Univ. Press.
- SCHECHTMAN, G., SCHLUMPRECHT, T. and ZINN, J. (1998). On the Gaussian measure of the intersection. *Ann. Probab.* **26** 346–357.
- SHAO, Q. M. (1993). A note on small ball probability of Gaussian processes with stationary increments. *J. Theoret. Probab.* **6** 595–602.
- TALAGRAND, M. (1993). New Gaussian estimates for enlarged balls. *Geom. Funct. Anal.* **3** 502–526.

- TIMAN, A. F. (1964). On the order of growth of the  $\varepsilon$ -entropy of spaces of real continuous functionals defined on a connected compactum. *Uspehi Mat. Nauk* **19** 173–177. (In Russian.)
- TOMCZAK-JAEGERMANN, N. (1987). Dualité des nombres d'entropie pour des opérateurs à valeurs dans un espace de Hilbert. *C. R. Acad. Sci. Paris* **305** 299–301.

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF DELAWARE  
NEWARK, DELAWARE 19711  
E-MAIL: wli@math.udel.edu

FRIEDRICH SCHILLER UNIVERSITÄT JENA  
INSTITUT FÜR STOCHASTIK  
ERNST ABBE PLATZ 1-4  
07743 JENA  
GERMANY  
E-MAIL: lindew@minet.uni-jena.de