Bernoulli **8**(5), 2002, 669–696

Optimal series representation of fractional Brownian sheets

THOMAS KÜHN[1] and WERNER LINDE[2]

[1]*Mathematisches Institut, Universität Leipzig, Augustusplatz 10/11, 04109 Leipzig, Germany. E-mail: kuehn@mathematik.uni-leipzig.de*
[2]*Institut für Stochastik, FSU Jena, Ernst-Abbe-Platz 1–4, 07743 Jena, Germany. E-mail: lindew@minet.uni-jena.de*

For $0 < \gamma < 2$, let B_γ^d be a d-dimensional γ-fractional Brownian sheet with index set $[0, 1]^d$ and let $(\xi_k)_{k\geqslant 1}$ be an independent sequence of standard normal random variables. We prove the existence of continuous functions u_k such that almost surely

$$B_\gamma^d(t) = \sum_{k=1}^{\infty} \xi_k u_k(t), \qquad t \in [0, 1]^d,$$

and

$$\left(\mathbb{E} \sup_{t\in[0,1]^d} \left| \sum_{k=n}^{\infty} \xi_k u_k(t) \right|^2 \right)^{1/2} \approx n^{-\gamma/2}(1 + \log n)^{d(\gamma+1)/2-\gamma/2}.$$

This order is shown to be optimal. We obtain small-ball estimates for B_γ^d, extending former results in the case $\gamma = 1$. Our investigations rest upon basic properties of different kinds of s-numbers of operators.

Keywords: approximation numbers; fractional Brownian motion; Gaussian process; small-ball behaviour

1. Introduction

Let E be a Banach space and let X be an E-valued centred Gaussian random variable. Then the variable X admits an almost surely convergent representation

$$X = \sum_{k=1}^{\infty} \xi_k x_k \tag{1.1}$$

for suitable x_k in E and a standard normal independent sequence $(\xi_k)_{k\geqslant 1}$ (see Lifshits 1995). Since (1.1) is not unique, one may ask for optimal representations, i.e. for those where the tails tend to zero as fast as possible. More precisely, we wish to determine the behaviour of

$$l_n(X) := \inf\left\{ \left(\mathbb{E}\left\| \sum_{k=n}^{\infty} \xi_k x_k \right\|^2 \right)^{1/2} : X = \sum_{k=1}^{\infty} \xi_k x_k \text{ a.s.} \right\} \tag{1.2}$$

as $n \to \infty$, and, if possible, describe those x_k for which the optimal rate is attained. This problem appears in a natural way when simulating Gaussian random processes. Indeed, let $X = (X(t))_{t\in K}$ be a centred Gaussian process over a compact metric space K. If X has a.s. continuous sample paths, by (1.1) there are continuous functions u_k on K such that a.s.

$$X(t) = \sum_{k=1}^{\infty} \xi_k u_k(t), \qquad t \in K. \tag{1.3}$$

An example is

$$B(t) = \xi_0 \cdot t + \sqrt{2} \sum_{k=1}^{\infty} \xi_k \cdot \frac{\sin(\pi k t)}{\pi k}, \qquad 0 \leqslant t \leqslant 1,$$

for the Brownian motion B over [0, 1]; other series representations of B are well known, for example using the Faber–Schauder system instead of the trigonometric.

In order to simulate a Gaussian process X it is necessary to replace the series in (1.3) by a finite sum. To minimize the error one has to choose the u_k such that the average of the tail

$$\left(\mathbb{E} \sup_{t\in K} \left| \sum_{k=n}^{\infty} \xi_k u_k(t) \right|^2 \right)^{1/2}$$

becomes minimal. But this is equivalent to the problem stated above (cf. Proposition 2.1 below).

Our aim is to investigate these questions for the d-dimensional fractional Brownian sheets. Given $\underline{\gamma} = (\gamma_1, \ldots, \gamma_d)$, with $0 < \gamma_j < 2$, there exists a Gaussian process $B_{\underline{\gamma}}^d$ over $[0, 1]^d$ possessing a.s. continuous sample paths and satisfying

$$\mathbb{E} B_{\underline{\gamma}}^d(t) B_{\underline{\gamma}}^d(s) = 2^{-d} \prod_{j=1}^{d} (|t_j|^{\gamma_j} + |s_j|^{\gamma_j} - |t_j - s_j|^{\gamma_j}), \tag{1.4}$$

$t = (t_1, \ldots, t_d)$ and $s = (s_1, \ldots, s_d)$. The process $B_{\underline{\gamma}}^d$ is called a $\underline{\gamma}$-fractional Brownian sheet. For $\underline{\gamma} = (1, \ldots, 1)$ we obtain the ordinary d-dimensional Brownian sheet. The basic result of our paper asserts that

$$l_n(B_{\underline{\gamma}}^d) \approx n^{-\gamma_1/2} (1 + \log n)^{\nu(\gamma_1+1)/2 - \gamma_1/2}, \tag{1.5}$$

where the γ_j are ordered such that

$$0 < \gamma_1 = \ldots = \gamma_\nu < \gamma_{\nu+1} \leqslant \ldots \leqslant \gamma_d < 2.$$

For the Brownian motion B this yields $l_n(B) \approx (n^{-1} \log n)^{1/2}$, which was shown in the theory of so-called average linear widths (Maiorov and Wasilkowski 1996).

As mentioned in Li and Linde (1999), the behaviour of $l_n(X)$ as $n \to \infty$ is tightly related to small-ball estimates for X, i.e. to the behaviour of

$$-\log \mathbb{P}(\|X\| < \varepsilon) \tag{1.6}$$

as $\varepsilon \to 0$. In this context, statement (1.5) is quite surprising because a similar precise assertion for the small-ball behaviour of $B^d_{\underline{\gamma}}$ is known only in some very special cases (Monrad and Rootzén 1995; Shao 1993; Talagrand 1994; Belinsky and Linde 2002). Our results lead to general lower and upper estimates of (1.6) for $\underline{\gamma}$-fractional Brownian sheets which differ from each other by $(\log \varepsilon^{-1})^{1/\gamma_1}$. This extends recent results in Dunker *et al.* (1999) for the Brownian sheet, i.e. $y = (1, \ldots, 1)$.

The paper is organized as follows. In Section 2 we introduce the l-numbers $l_n(T)$ for an operator T from a Hilbert space H into E and state their basic properties. Section 3 is devoted to certain relations between these and ordinary approximation numbers. We strengthen the results in Section 4 for the special case $E = C(K)$ with K metric compact. Section 5 is devoted to the multidimensional case, i.e., more precisely, to the study of tensor products of operators. We investigate, for example, the dependence of $l_n(T \otimes S)$ on $L_n(T)$ and $l_n(S)$, respectively. The main result in this section is Theorem 5.7, where we describe the exact behaviour of $l_n(R^d_{\underline{\alpha}})$. Here $R^d_{\underline{\alpha}}$ denotes the d-fold tensor product of certain Riemann–Liouville operators of fractional integration. In Section 6 we show how the general results about l-numbers of operators lead to estimates for $l_n(X)$ where X is a Gaussian centred Banach space valued random variable. Furthermore, we introduce and describe so-called approximation numbers of X. Finally, in Section 7 we give the proof of (1.5). The basic idea is similar to that in Li and Linde (1998), namely to split $B^d_{\underline{\gamma}}$ into a sum of processes, where one is generated by $R^d_{(\underline{\gamma}+1)/2}$ and the others turn out to be of lower order.

2. Approximation numbers of operators

Let T be a compact operator from a Banach space E into a Banach space F. The behaviour of its approximation numbers

$$a_n(T) := \inf\{\|T - S\| : S \text{ operator from } E \text{ to } F, \operatorname{rank}(S) < n\} \tag{2.1}$$

leads to information about the degree of compactness of T. If T is an operator from a Hilbert space H into a Banach space E, then we have (see Pietsch 1987)

$$a_n(T) := \inf\{\|T - TP\| : P \text{ orthogonal projection in } H, \operatorname{rank}(P) < n\}.$$

For example, if T maps H into itself, then it follows (Carl and Stephani 1990) that

$$a_n(T) = \lambda_n(T)$$

where $\lambda_1(T) \geqslant \lambda_2(T) \geqslant \ldots \geqslant 0$ denotes the sequence of singular numbers of T, i.e. $\lambda_k(T)^2$ is the kth eigenvalue of TT^* in decreasing order.

When investigating Gaussian random processes, another kind of approximation numbers is of importance. Thus suppose we are given a (linear, bounded) operator T from a separable Hilbert space H into a Banach space E. Then its l-norm is defined by

$$l(T) := \sup_{H_0 \subseteq H} \left\{ \left(\int_{H_0} \|Th\|^2 \mathrm{d}\gamma_{H_0}(h) \right)^{1/2} \right\}, \tag{2.2}$$

where the supremum is taken over all finite-dimensional subspaces $H_0 \subseteq H$ and γ_{H_0} denotes the (unique) standard Gaussian measure on H_0.

Let $(\xi_k)_{k=1}^\infty$ be an independent sequence of standard normal random variables. If

$$\sum_{k=1}^\infty \xi_k T e_k \tag{2.3}$$

converges a.s. in E for one (or equivalently each) orthonormal basis (ONB) $(e_k)_{k=1}^\infty$ in H, then we have $l(T) < \infty$ and, moreover,

$$l(T) = \left(\mathbb{E} \left\| \sum_{k=1}^\infty \xi_k T e_k \right\|^2 \right)^{1/2} \tag{2.4}$$

is independent of the choice of ONB.

In order to measure the convergence of (2.3) in E, we use the l-numbers defined by

$$l_n(T) := \inf \left\{ \left(\mathbb{E} \left\| \sum_{k=n}^\infty \xi_k T e_k \right\|^2 \right)^{1/2} : (e_k)_{k=1}^\infty \text{ ONB in } H \right\}. \tag{2.5}$$

It is well known and easy to see (Pietsch 1987; Pisier 1989) that these numbers may also be defined by

(i) $l_n(T) = \inf\{ l(T - S) \colon S$ operator from H to E, $\operatorname{rank}(S) < n\}$ or
(ii) $l_n(T) = \inf\{ l(T - TP) \colon P$ orthogonal projection in H, $\operatorname{rank}(P) < n\}$ or
(iii) $l_n(T) = \inf\{ l(T|_{H_0^\perp}) \colon H_0 \subseteq H, \dim(H_0) < n\}$.

These numbers enjoy the following properties. Suppose that T and S are operators from H into E. Then, for all $m, n \in \mathbb{N}$, we have

$$l_{n+m-1}(T + S) \leqslant l_n(T) + l_m(S). \tag{2.6}$$

Furthermore, if $T \colon H \to E$ and $S \colon H \to H$, for all $n, m \in \mathbb{N}$ it follows that

$$\begin{aligned} l_{n+m-1}(T \circ S) &\leqslant l_n(T) \cdot a_m(S), \\ l_{n+m-1}(T \circ S) &\leqslant a_n(T) \cdot l_m(S). \end{aligned} \tag{2.7}$$

In special cases the l-numbers may be easily calculated. If T maps H into itself, then the l-numbers of T satisfy

$$l_n(T) = \left(\sum_{k=n}^\infty \lambda_k(T)^2 \right)^{1/2}$$

where as above $\lambda_k(T)$ is the kth singular number of T.

The next result shows (among others) that there exists *one* fixed ONB in H for which *all* nth tails admit the same estimate as $l_n(T)$. Since the proof is standard (see Pietsch 1987, 2.3.8) we omit it. Here and later on we use the following notation: If a_n and b_n are two sequences of non-negative numbers, we write $a_n \preccurlyeq b_n$ whenever there is a constant $c > 0$ such that $a_n \leqslant c \cdot b_n$ for all $n \in \mathbb{N}$, while $a_n \approx b_n$ means $a_n \preccurlyeq b_n$ as well as $b_n \preccurlyeq a_n$.

Proposition 2.1. *Let $\alpha > 0$ and $\beta \in \mathbb{R}$. For $T : H \to E$ the following are equivalent.*

(i) $l_n(T) \preccurlyeq n^{-\alpha}(1 + \log n)^{\beta}$.

(ii) *For some (each) $q > 1$ and $\lambda \in \mathbb{R}$, the operator T may be represented as $T = \sum_{k=1}^{\infty} T_k$, where the T_k map H into E with*

$$\operatorname{rank}(T_k) \preccurlyeq q^k k^{\lambda} \quad \text{and} \quad l(T_k) \preccurlyeq q^{-k\alpha} k^{\beta - \alpha\lambda}.$$

(iii) *For some (each) $q > 1$* and *$\lambda \in \mathbb{R}$ there are orthogonal projections P_k on H of rank less than $c \cdot q^k k^{\lambda}$ such that*

$$T = \sum_{k=1}^{\infty} TP_k \quad \textit{and} \quad l(TP_k) \preccurlyeq q^{-k\alpha} k^{\beta - \alpha\lambda}$$

(iv) *As (iii), with the P_k pairwise orthogonal.*

(v) *There is a fixed ONB $(f_k)_{k=1}^{\infty}$ in H such that, for all $n \in \mathbb{N}$,*

$$\left(\mathbb{E} \left\| \sum_{k=n}^{\infty} \xi_k T f_k \right\|^2 \right)^{1/2} \preccurlyeq n^{-\alpha}(1 + \log n)^{\beta}.$$

3. Relations between approximation and l-numbers

In view of $\|T\| \leqslant l(T)$ for $T : H \to E$, we clearly have

$$a_n(T) \leqslant l_n(T) \tag{3.1}$$

for all $n \in \mathbb{N}$. Our next aim is to improve (3.1). To do so we need a lemma which slightly generalizes a result due to D. R. Lewis (see Pisier 1989, Proposition 1.8).

Lemma 3.1. *Let T be an operator from a Hilbert space H into a Banach space E and let P be an orthogonal projection on H of rank n. Then there are orthonormal elements $f_1, f_2 \ldots$ in $(I - P)(H)$ with*

$$\|Tf_k\| > \tfrac{1}{2} a_{n+k}(T), \qquad k = 1, 2, \ldots . \tag{3.2}$$

Proof. Since $\operatorname{rank}(TP) < n$, it follows that

$$a_{n+1}(T) \leqslant \|T - TP\|$$

by the definition of approximation numbers. Consequently, we find an element $\tilde{f}_1 \in H$, $\|\tilde{f}_1\| = 1$, such that

$$\|T(I-P)\tilde{f}_1\| \geqslant \tfrac{1}{2} a_{n+1}(T).$$

Setting

$$f_1 := \frac{(I-P)\tilde{f}_1}{\|I-P)\tilde{f}_1\|},$$

we obtain a normalized element $f_1 \in (I-P)(H)$ satisfying (3.2) for $k=1$.

Define now an orthogonal projection P_1 on H of rank $n+1$ by

$$P_1 h := Ph + \langle h, f_1 \rangle f_1$$

and apply the above procedure to P_1 and $n+1$. This leads to a normalized element f_2 in $(I-P_1)(H)$ satisfying (3.2) for $k=2$. By the construction of P_1 we have $f_2 \in (I-P)(H)$ as well as $f_1 \perp f_2$. For $\dim(H) = \infty$ we may now proceed in the same way to obtain $f_3, f_4, \ldots$. Otherwise, of course the construction stops at a certain point. This completes the proof. □

The next result is the key to obtaining lower estimates of $l_n(T)$, in particular in the case of $E = C(K)$ with K compact metric.

Proposition 3.2. *Let T be an operator from H into E and let m, n be any natural numbers. Then we have*

$$\sqrt{\log m}\cdot a_{n+m-1}(T) \leqslant c\cdot l_n(T) \tag{3.3}$$

with some constant $c > 0$, independent of T, m and n.

Proof. Let P be an orthogonal projection in H of rank $n-1$ such that

$$l(T - TP) \leqslant 2\cdot l_n(T). \tag{3.4}$$

Next we apply Lemma 3.1 to T and P. Then there exists an orthonormal system $f_1, f_2, \ldots$ in $(I-P)(H)$ with

$$\|Tf_k\| \geqslant \tfrac{1}{2} a_{n+k-1}(T), \qquad k = 1, 2, \ldots . \tag{3.5}$$

By the definition of the l-norm we obtain

$$l(T - TP) \geqslant \left(\mathbb{E} \left\| \sum_{k=1}^{m} \xi_k Tf_k \right\|^2 \right)^{1/2} \tag{3.6}$$

for each $m \in \mathbb{N}$. Recall that the right-hand expectation is increasing with respect to m. To proceed further, we need an inequality due to Pisier (1989, (4.14)) asserting that

$$\mathbb{E} \sup_{1 \leqslant k \leqslant m} |\xi_k| \, \|Tf_k\| \leqslant \mathbb{E} \left\| \sum_{k=1}^{m} \xi_k Tf_k \right\|;$$

hence by (3.5), (3.6) and (3.4) we arrive at

$$\tfrac{1}{2}a_{n+m-1}(T)\cdot\mathbb{E}\sup_{1\leqslant k\leqslant m}|\xi_k| \leqslant l(T-TP) \leqslant 2\cdot l_n(T) \tag{3.7}$$

and the assertion follows from (3.7) in view of

$$c\cdot\sqrt{\log m} \leqslant \mathbb{E}\sup_{1\leqslant k\leqslant m}|\xi_k|$$

(Pisier 1989, Lemma 4.14).

4. Operators with values in spaces of continuous functions

We now specify the Banach space E. Let (K, d) be a compact metric space. As usual, $C(K)$ denotes the Banach space of continuous (real-valued) functions on K endowed with the supremum norm $\|\cdot\|_\infty$. If μ is a finite (Borel) measure on K, let

$$J : C(K) \to L_2(K, \mu)$$

be the natural embedding. Thus, if T is some compact operator from $L_2(K, \mu)$ into $C(K)$, the operator $J \circ T$ maps from a Hilbert space into itself, so that its singular numbers $\lambda_n = \lambda_n(JT)$, with $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant 0$, are well defined. Recall that $\lambda_n(JT) = a_n(JT)$.

With these preliminaries the following is valid.

Proposition 4.1. *For each $n \in \mathbb{N}$, it follows that*

$$\lambda_{3n-2}(JT) \leqslant c\cdot(n\log n)^{-1/2}\cdot l_n(T). \tag{4.1}$$

Before we prove (4.1) let us state some basic facts about Weyl numbers of operators. Let T be an operator acting between arbitrary Banach spaces E and F. Then its nth Weyl number $x_n(T)$ is defined by

$$x_n(T) := \sup\{a_n(T\circ S) : S : l_2 \to E, \|S\| \leqslant 1\}.$$

We shall use the following basic properties of these numbers (see Pietsch 1987; König 1989, pp. 69, 81).

(i) The x_n are a multiplicative s-scale, i.e.

$$x_{n+m-1}(T_1 \circ T_2) \leqslant x_n(T_1)\cdot x_m(T_2)$$

whenever $T_1 \circ T_2$ makes sense.

(ii) For operators T defined on a Hilbert space it follows that

$$a_n(T) = x_n(T).$$

(iii) We have

$$x_n(T) \leqslant n^{-1/2}\pi_2(T)$$

where $\pi_2(T)$ is the 2-absolutly summing norm of T.

The only, albeit important, property of this norm (Pietsch 1978, 17.3.8) we shall use later on is

$$\pi_2(J : C(K) \to L_2(K, \mu)) = \mu(K)^{1/2} < \infty. \tag{4.2}$$

Proof of Proposition 4.1. Using properties (i), (ii) and (iii) of the x_n as well as (4.2) we obtain

$$\begin{aligned}\lambda_{3n-2}(JT) = a_{3n-2}(JT) = x_{3n-2}(JT) &\leqslant x_{2n-1}(T)\cdot x_n(J)\\ &\leqslant a_{2n-1}(T)n^{-1/2}\pi_2(J) \leqslant c\cdot n^{-1/2}a_{2n-1}(T).\end{aligned} \tag{4.3}$$

Applying (3.3) with $m = n$ implies

$$\sqrt{\log n}\cdot a_{2n-1}(T) \leqslant c\cdot l_n(T),$$

and plugging this into (4.3) completes the proof. □

Next we present a general tool for obtaining upper estimates for $l_n(T)$ with T having values in $C[0, 1]$. Let u be the piecewise linear function defined by

$$u(t) := t\cdot\mathbf{1}_{[0,1/2]}(t) + (1-t)\cdot\mathbf{1}_{(1/2,1]}(t).$$

For $m = -2, -1, \ldots$ we now set $J_{-2} = J_{-1} = \{0\}$ and

$$J_m := \{0, \ldots, 2^m - 1\}, \qquad m \geqslant 0,$$

and define functions $u_{m,j}$, $m \geqslant -2$, $j \in J_m$, by

$$u_{-2,0} := \mathbf{1}_{[0,1]} \quad \text{and} \quad u_{-1,0}(t) := t$$

and, if $m \geqslant 0$, by

$$u_{m,j}(t) := 2^{-m/2}U(2^m t - j), \qquad 0 \leqslant t \leqslant 1.$$

Of course, the $u_{m,j}$ form the well-known Faber–Schauder system in $C[0, 1]$. Note that $\|u_{m,j}\|_\infty = 2^{-m/2-1}$, $m \geqslant 0$, and that for fixed m the supports of the $u_{m,j}$ are disjoint.

Given $\tau > 0$ and $t \in \mathbb{R}$ with $0 \leqslant t - \tau < t < t + \tau \leqslant 1$, the functional $\Delta_{(t|\tau)} \in C^*[0, 1]$ is defined by

$$\Delta_{(t|\tau)} := 2\delta_t - \delta_{t+\tau} - \delta_{t-\tau},$$

where as usual δ_t denotes Dirac point measure at $t \in \mathbb{R}$. Finally, letting

$$\varphi_{-2,0} := \delta_0, \qquad \varphi_{-1,0} := \delta_1 - \delta_0 \tag{4.4}$$

and, if $m \geqslant 0$,

$$\varphi_{m,j} := 2^{m/2}\Delta_{((2j+1)2^{-(m+1)}|2^{-(m+1)})}, \qquad j \in J_m, \tag{4.5}$$

i.e.

$$\langle f, \varphi_{m,j}\rangle = 2^{m/2}\left\{2f\left(\frac{2j+1}{2^{m+1}}\right) - f\left(\frac{j}{2^m}\right) - f\left(\frac{j+1}{2^m}\right)\right\}, \qquad j \in J_m, \tag{4.6}$$

every $f \in C[0, 1]$ can be written as

$$f = \sum_{m=-2}^{\infty} \sum_{j \in J_m} \langle f, \varphi_{m,j} \rangle \cdot u_{m,j}.$$

Now, if $T : H \to C[0, 1]$, it may be represented as

$$Th = \sum_{m=-2}^{\infty} \sum_{j \in J_m} \langle h, h_{m,j} \rangle \cdot u_{m,j}, \qquad h \in H,$$

with $h_{m,j} := T^* \varphi_{m,j} \in H$.

Proposition 4.2. *Let $T: H \to E$ be an operator such that, for some $\gamma > 0$, $\beta \in \mathbb{R}$ and all $m \in \mathbb{N}$, we have*

$$\sup_{j \in J_m} \| h_{m,j} \|_H = \sup_{j \in J_m} \| T^* \varphi_{m,j} \|_H \leqslant c \cdot m^{\beta} 2^{m(1/2-\gamma)}. \tag{4.7}$$

This implies that

$$l_n(T) \preccurlyeq (1 + \log n)^{\beta+1/2} \cdot n^{-\gamma}. \tag{4.8}$$

Proof. Let $T_m : H \to C[0, 1]$, $m = -2, -1, \ldots,$ be defined by

$$T_m h := \sum_{j \in J_m} \langle h, h_{m,j} \rangle u_{m,j}, \qquad h \in H.$$

Of course, $\operatorname{rank}(T_m) \leqslant 2^{\max\{m,0\}}$, thus in order to prove (4.8), in view of Proposition 2.1, it suffices to verify

$$l(T_m) \preccurlyeq m^{\beta+1/2} 2^{-m\gamma}. \tag{4.9}$$

Let X_m be a standard Gaussian random variable with values in $H_m = \operatorname{span}\{h_{m,j} : j \in J_m\}$. Then

$$\begin{aligned} l(T_m) &= \left(\mathbb{E} \| T_m(X_m) \|_\infty^2 \right)^{1/2} = \left(\mathbb{E} \sup_{j \in J_m} |\eta_j|^2 \| u_{m,j} \|_\infty^2 \right)^{1/2} \\ &= 2^{-m/2-1} \left(\mathbb{E} \sup_{j \in J_m} |\eta_j|^2 \right)^{1/2}, \end{aligned} \tag{4.10}$$

where the centred Gaussian vector $(\eta_j)_{j \in J_m}$ is defined by

$$\eta_j := (\langle X_m, h_{m,j} \rangle, \qquad j \in J_m.$$

Note that the η_j are not necessarily independent. They satisfy

$$\mathbb{E} |\eta_j|^2 = \| h_{m,j} \|_H^2, \tag{4.11}$$

hence a well-known estimate for general centred Gaussian vectors (Pisier 1989, Lemma 4.14), together with (4.7) and (4.11), implies that

$$\left(\mathbb{E}\sup_{j\in J_m}|\eta_j|^2\right)^{1/2} \leqslant c\cdot\sqrt{\log\#(J_m)}\cdot\left(\sup_{j\in J_m}\mathbb{E}|\eta_j|^2\right)^{1/2}$$

$$\leqslant c\cdot m^{1/2}\cdot\sup_{j\in J_m}\|h_{m,j}\|_H \leqslant c\cdot m^{\beta+1/2}2^{m(1/2-\gamma)}. \tag{4.12}$$

Consequently, using (4.10) and (4.12), we finally obtain (4.9) as required. □

Corollary 4.3. *Suppose that, for some* $\gamma > 0$, *we have*

$$\sup_{\|h\|\leqslant 1}|2(Th)(t) - (Th)(t+\tau) - (Th)(t-\tau)| \leqslant c\cdot\tau^\gamma \tag{4.13}$$

for all t, τ *such that* $0 \leqslant t-\tau < t < t+\tau \leqslant 1$. *This implies that*

$$l_n(T) \leqslant \sqrt{\log n}\cdot n^{-\gamma}.$$

In particular, this is satisfied for T *mapping* H *into the space* $C^\gamma[0, 1]$ *of Hölder continuous functions of order* $\gamma, 0 < \gamma \leqslant 1$.

We now give a first example. If $\alpha > 0$, let

$$(R_\alpha f)(s) := \frac{1}{\Gamma(\alpha)}\int_0^s (s-t)^{\alpha-1}f(t)\mathrm{d}t \tag{4.14}$$

be the Riemann–Liouville operator of fractional integration of order α. It is easily checked that R_α maps $L_2[0, 1]$ into $L_p[0, 1]$ provided that $\alpha > [1/2 - 1/p]_+$ where in the case $p = \infty$ the operator R_α maps $L_2[0, 1]$ even into $C[0, 1]$.

Let us state some properties of R_α for later use:

(i) $R_\alpha \circ R_\beta = R_{\alpha+\beta}$
(ii) If $\frac{1}{2} < \alpha \leqslant \frac{3}{2}$, then R_α maps $L_2[0, 1]$ into $C^\gamma[0, 1]$ with $\gamma = \alpha - \frac{1}{2}$.
(iii) If $\alpha > 0$, then the singular numbers of R_α regarded as an operator from $L_2[0, 1]$ into itself behave as

$$\lambda_n(R_\alpha) \approx n^{-\alpha} \tag{4.15}$$

(see Vu and Gorenflo 1996).

Next we wish to describe the asymptotic behaviour of $l_n(R_\alpha)$ for R_α mapping $L_2[0, 1]$ into $C[0, 1]$.

Proposition 4.4. *If* $\alpha > \frac{1}{2}$ *and* $R_\alpha:\ L_2[0, 1] \to C[0, 1]$, *then we have*

$$l_n(R_\alpha) \approx n^{-\alpha+1/2}\cdot(1+\log n)^{1/2}. \tag{4.16}$$

Proof. We start by proving the upper estimate. If $\frac{1}{2} < \alpha \leqslant \frac{3}{2}$, this is a direct consequence of property (ii) above and Corollary 4.3. Now assume $\alpha > \frac{3}{2}$ and choose a natural number k with

$$\tfrac{1}{2} + k < \alpha \leqslant \tfrac{3}{2} + k.$$

Using (i), we write $R_\alpha = R_{\alpha-k} \circ R_k$, where we regard R_k as an operator from $L_2[0, 1]$ into itself. Hence, by the first step and by property (iii) above, we obtain

$$\begin{aligned} l_{2n-1}(R_\alpha) &\leqslant l_n(R_{\alpha-k}) \cdot a_n(R_k : \ L_2 \to L_2) \\ &\leqslant c \cdot n^{-(\alpha-k)+1/2} \cdot (\log n)^{1/2} \cdot n^{-k} = c \cdot n^{-\alpha+1/2} (\log n)^{1/2}, \end{aligned}$$

completing the proof of the upper estimate.

For the lower estimate we use (4.15) and Proposition 4.1. Then

$$n^{-\alpha} \leq c \cdot a_n(R_\alpha : \ L_2 \to L_2) \leqslant c \cdot (n \log n)^{-1/2} l_n(R_\alpha : \ L_2 \to C)$$

completes the proof. □

5. The multidimensional case

For two Banach spaces E and F, let $E \check{\otimes} F$ be their injective tensor product, i.e. the completion of $E \otimes F$ with respect to the norm

$$\left\| \sum_{i=1}^{n} x_i \otimes y_i \right\| := \sup \left\{ \sum_{i=1}^{n} \langle x_i, x^* \rangle \langle y_i, y^* \rangle : \ \|x^*\|_{E^*} \leqslant 1, \|y^*\|_{F^*} \leqslant 1 \right\}.$$

Given two Hilbert spaces H, K, their tensor product $H \otimes K$ has a natural scalar product defined by

$$\langle f_1 \otimes g_1, f_2 \otimes g_2 \rangle := \langle f_1, f_2 \rangle \langle g_1, g_2 \rangle,$$

which may be extended by linearity. The completion is denoted by $H \hat{\otimes}_2 K$ and called the (Hilbert) tensor product of H and K. If $T : H \to E$ and $S : K \to F$ are operators, their tensor product $T \otimes S$ (defined canonically by $(T \otimes S)\,(f \otimes g) := (Tf) \otimes (Sg)$) maps $H \hat{\otimes}_2$ into $E \check{\otimes} F$ and satisfies

$$\|T \otimes S\| \leqslant \|T\|\,\|S\|.$$

The next result is basic for later estimates and may be found in Ledoux and Talagrand (1991, p. 84).

Proposition 5.1. *Let T and S be as before with tensor product $T \otimes S$ mapping into $E \check{\otimes} F$. Then it follows that*

$$l(T \otimes S) \leqslant l(T)\|S\| + l(S)\|T\|.$$

In particular, since $\|T\| \leqslant l(T)$, this implies

$$l(T \otimes S) \leqslant 2\, l(T) \cdot l(S). \tag{5.1}$$

The preceding estimate yields the following useful result for the l-numbers of tensor products.

Proposition 5.2. *Let T and S be operators on Hilbert spaces H and K with values in Banach spaces E and F, respectively. If*

$$l_n(T) \leqslant n^{-\alpha}(1+\log n)^{\beta},$$

$$l_n(S) \leqslant n^{-\gamma},$$

for some $0<\alpha<\gamma$ and $\beta \in \mathbb{R}$, then for $T \otimes S$ mapping into $E \check{\otimes} F$ it follows that

$$l_n(T \otimes S) \leqslant n^{-\alpha}(1+\log n)^{\beta} \tag{5.2}$$

as well.

Proof. Let us choose integers p and q with

$$1<\frac{p}{q}<\frac{\gamma}{\alpha}. \tag{5.3}$$

By Proposition 2.1 we find for these numbers operators T_k and S_ℓ such that

$$T=\sum_{k=1}^{\infty} T_k, \qquad S=\sum_{\ell=1}^{\infty} S_\ell, \tag{5.4}$$

$$\operatorname{rank}(T_k) \leqslant 2^{kp}, \qquad \operatorname{rank}(S_\ell) \leqslant 2^{\ell q}, \tag{5.5}$$

as well as

$$l(T_k) \leqslant c \cdot 2^{-\alpha kp} \cdot k^{\beta}, \tag{5.6}$$

$$l(S_\ell) \leqslant c \cdot 2^{-\gamma \ell q}, \tag{5.7}$$

for $k, \ell = 1, 2, \ldots$.

Next we define operators V_n on $H \hat{\otimes}_2 K$ with values in $E \check{\otimes} F$ by

$$V_{n+1} := \sum_{k+\ell=n} T_k \otimes S_\ell, \qquad n=1,2,\ldots,$$

which possess the following properties:

(a) By (5.4) we obtain

$$T \otimes S = \sum_{n=1}^{\infty} V_{n+1}.$$

(b) Using (5.5) and $p>q$, hence $p \geqslant q+1$, it follows that

$$\operatorname{rank}(V_{n+1}) \leqslant \sum_{k+\ell=n} \operatorname{rank}(T_k) \cdot \operatorname{rank}(S_\ell) \leqslant \sum_{k+\ell=n} 2^{kp+\ell q}$$

$$= 2^{np} \cdot \sum_{\ell=1}^{n-1} 2^{\ell(q-p)} \leqslant 2^{np}.$$

(c) Finally, because of (5.6) and (5.7), an application of (5.1) leads to

$$\begin{aligned} l(V_{n+1}) &\leqslant \sum_{k+\ell=n} l(T_k \otimes S_\ell) \leqslant 2\cdot \sum_{k+\ell=n} l(T_k)\cdot l(S_\ell) \\ &\leqslant c\cdot \sum_{k+\ell=n} 2^{-\alpha kp-\gamma\ell q}\cdot k^\beta \\ &= c\cdot 2^{-\alpha np} n^\beta \cdot \sum_{\ell=1}^{n-1} 2^{-\ell(\gamma q-\alpha p)}\cdot(1-\ell/n)^\beta \\ &\leqslant c\cdot 2^{-\alpha np}\cdot n^\beta\cdot \sum_{\ell=1}^{\infty} 2^{-\ell(\gamma q-\alpha p)} \\ &= c'\cdot 2^{-\alpha np}\cdot n^\beta, \end{aligned}$$

where we have used $\alpha p < \gamma q$ by the choice of p and q. The above arguments only apply for $\beta \geqslant 0$, yet for $\beta < 0$ a slight modification gives the same estimate.

If we summarize all the properties of V_n, Proposition 2.1 implies the desired estimate

$$l_n(T \otimes S) \leqslant c\cdot n^{-\alpha}(1+\log n)^\beta.$$

□

The preceding proposition no longer applies if the l-numbers of both operators are of the same order. Ideas similar to those above give in this case the following result which, in general, is far from optimal, yet suffices for later purposes.

Proposition 5.3. *Suppose that T and S as above satisfy*

$$l_n(T) \preccurlyeq n^{-\alpha}(1+\log n)^\beta,$$

$$l_n(S) \preccurlyeq n^{-\alpha}(1+\log n)^\delta,$$

for some $\alpha > 0$ and β, $\delta \in \mathbb{R}$. Then we have

$$l_n(T \otimes S) \preccurlyeq n^{-\alpha}(1+\log n)^{\beta+\delta+\alpha+1}.$$

Proof. We choose operators T_k and S_ℓ as above, but this time for $p = q = 1$. Defining V_{n+1} as before, we obtain

$$\operatorname{rank}(V_{n+1}) \leqslant n\cdot 2^n$$

and

$$l(V_{n+1}) \leqslant c\cdot 2^{-\alpha n}\cdot n^{\beta+\delta+1},$$

hence by Proposition 2.1

$$l_n(T \otimes S) \leqslant n^{-\alpha}(1 + \log n)^{\beta+\delta+\alpha+1}$$

as asserted. □

Given $\underline{\alpha} = (\alpha_1, \ldots, \alpha_d)$, we now investigate the d-fold tensor product

$$R^d_{\underline{\alpha}} := R_{\alpha_1} \otimes \cdots \otimes R_{\alpha_d} \tag{5.8}$$

where the components R_{α_j} denote as in (4.14) Riemann–Liouville operators of fractional integration. $R^d_{\underline{\alpha}}$ is bounded from $L_2[0, 1]^d$ into itself if the α_j are positive and it is bounded as an operator from $L_2[0, 1]^d$ into $C[0, 1]^d$ provided that $\alpha_j > \frac{1}{2}$ for $1 \leqslant j \leqslant d$. Furthermore, without loss of generality, we always may order the α_j such that

$$\alpha_1 = \ldots = \alpha_\nu < \alpha_{\nu+1} \leqslant \ldots \leqslant \alpha_d.$$

With this notation the following is valid.

Proposition 5.4. *Suppose $\alpha_1 > 0$ and regard $R^d_{\underline{\alpha}}$ as an operator from $L_2[0, 1]^d$ into itself. Then its singular numbers $\lambda_n(R^d_{\underline{\alpha}})$ satisfy*

$$\lambda_n(R^d_{\underline{\alpha}}) = a_n(R^d_{\underline{\alpha}} :\ L_2[0, 1]^d \to L_2[0, 1]^d) \approx n^{-\alpha_1}(1 + \log n)^{\alpha_1(\nu-1)}. \tag{5.9}$$

Proof. Let $\lambda^j_1 \geqslant \lambda^j_2 \geqslant \ldots > 0$ be the singular numbers of R_{α_j} as an operator in $L_2[0, 1]$, $1 \leqslant j \leqslant d$. Then it is well known and easy to see that $R^d_{\underline{\alpha}}$ has singular numbers

$$\lambda_{n_1,\ldots,n_d}(R^d_{\underline{\alpha}}) := \lambda^1_{n_1} \cdots \lambda^d_{n_d}, \qquad (n_1, \ldots, n_d) \in \mathbb{N}^d.$$

By (4.15) we have

$$\lambda^j_n \approx n^{-\alpha_j}, \tag{5.10}$$

hence there are constants $c_1,\ c_2 > 0$ with

$$c_1 \cdot n_1^{-\alpha_1} \cdots n_d^{-\alpha_d} \leqslant \lambda_{n_1,\ldots,n_d}(R^d_{\underline{\alpha}}) \leqslant c_2 \cdot n_1^{-\alpha_1} \cdots n_d^{-\alpha_d}. \tag{5.11}$$

Let $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant 0$ be the decreasing rearrangement of the singular numbers $\lambda_{n_1,\ldots,n_d}(R^d_{\underline{\alpha}})$. Now, if $\alpha_1 = \ldots = \alpha_d = \alpha$, i.e. $\nu = d$, it is easily checked (Talagrand 1994) that (5.11) implies

$$\lambda_n \approx n^{-\alpha}(1 + \log n)^{(d-1)\alpha}. \tag{5.12}$$

Note that (5.12) is equivalent to

$$\#\{(n_1, \ldots, n_d) :\ \lambda_{n_1,\ldots,n_d}(R^d_{\underline{\alpha}}) \geqslant \varepsilon\} \approx \varepsilon^{-1/\alpha}(\log \varepsilon^{-1})^{d-1},$$

which follows from (5.11) by a d-dimensional integration argument.

Let us now investigate the general case. This may be derived from (5.12) and the following observation. Since $\alpha_1 < \alpha_{\nu+1}$, by (5.10) we have

$$\#\{(m, n) :\ \lambda^1_m \cdot \lambda^{\nu+1}_n \geqslant \varepsilon\} \approx \varepsilon^{-1/\alpha_1}.$$

An iterative application of this reduces the general case to that of $d = \nu$ and completes the proof by (5.12). □

A combination of Propositions 4.1 and 5.4 leads to the following lower estimate of $l_n(R^d_{\underline{\alpha}})$.

Corollary 5.5. *Suppose that*

$$\tfrac{1}{2} < \alpha_1 = \cdots = \alpha_\nu < \alpha_{\nu+1} \leqslant \cdots \leqslant \alpha_d.$$

Then it follows that

$$l_n(R^d_{\underline{\alpha}}:\ L_2[0,1]^d \to C[0,1]^{\mathrm{d}}) \geqslant n^{-\alpha_1+1/2}(1+\log n)^{\alpha_1(\nu-1)+1/2}. \tag{5.13}$$

We now introduce some more notation. For $k \in \mathbb{N}_0 = \{0, 1, \ldots\}$ put

$$\psi_k(t) := \begin{cases} 1 & \text{if } k = 0, \\ \sqrt{2}\cos(k\pi t) & \text{if } k \geqslant 1, \end{cases}$$

and obtain an ONB in $L_2[0, 1]$. Given $k = (k_1, \ldots, k_d) \in \mathbb{N}_0^d$, we set

$$\vec{\psi}_k := \psi_{k_1} \otimes \cdots \otimes \psi_{k_d}, \tag{5.14}$$

leading to an ONB in $L_2[0, 1]^d$.

The following estimate is crucial for proving the corresponding upper estimate of (5.13).

Lemma 5.6. *Let* $\underline{\alpha} = (\alpha, \ldots, \alpha)$*, for some* $0 < \alpha \leqslant 1$. *Then it follows that*

$$\|R^d_{\underline{\alpha}}\vec{\psi}_k\|_{C[0,1]^d} \leqslant c_{\alpha,d} \cdot \prod_{j=1}^{d} \max\{1, k_j\}^{-\alpha} \tag{5.15}$$

for $k = (k_1, \ldots, k_d) \in \mathbb{N}_0^d$.

Proof. Let us first treat the one-dimensional case. For $k = 0$ or $\alpha = 1$, estimate (5.15) holds by trivial argument. Hence we may suppose $k \geqslant 1$ and $\alpha < 1$. Then we obtain

$$\begin{aligned}
\|R_\alpha \psi_k\|_{C[0,1]} &= c_\alpha \sup_{0\leqslant x\leqslant 1}\left|\int_0^x (x-s)^{\alpha-1}\psi_k(s)\mathrm{d}s\right| \\
&\leqslant c_\alpha\left\{\sup_{0\leqslant x\leqslant 1}\left|\int_0^x t^{\alpha-1}\cos(k\pi t)\mathrm{d}t\right| + \sup_{0\leqslant x\leqslant 1}\left|\int_0^x t^{\alpha-1}\sin(k\pi t)\mathrm{d}t\right|\right\} \\
&= c_\alpha k^{-\alpha}\left\{\sup_{0\leqslant x\leqslant 1}\left|\int_0^{kx} t^{\alpha-1}\cos(\pi t)\mathrm{d}t\right| + \sup_{0\leqslant x\leqslant 1}\left|\int_0^{kx} t^{\alpha-1}\sin(\pi t)\mathrm{d}t\right|\right\} \\
&\leqslant c'_\alpha k^{-\alpha},
\end{aligned}$$

where we have used $-1 < \alpha - 1 < 0$, thus

$$\sup_{0\leqslant y<\infty}\left|\int_0^y t^{\alpha-1}\cos(\pi t)\mathrm{d}t\right| < \infty,$$

and the same is valid for the term with the sine.

Now, if $k = (k_1, \ldots, k_d)$ with $k_j \geqslant 1$, the one-dimensional estimate yields

$$\|R^d_{\underline{\alpha}} \vec{\psi}_k\|_{C[0,1]^d} \leqslant c_{\alpha,d} \cdot (k_1 \cdots k_d)^{-\alpha}, \tag{5.16}$$

and we are done.

If one (or more) of the k_j is zero, we reduce the dimension d to $\#\{j \leqslant d:\ k_j > 0\}$, proceed as before and obtain (5.15) via (5.16). $\square$

We are now in a position to prove the main result of this section.

Theorem 5.7. *Suppose*

$$\tfrac{1}{2} < \alpha_1 = \ldots = \alpha_\nu < \alpha_{\nu+1} \leqslant \ldots \leqslant \alpha_d.$$

Then it follows that

$$l_n(R^d_{\underline{\alpha}}:\ L_2[0,1]^d \to C[0,1]^d) \approx n^{-\alpha_1+1/2}(1+\log n)^{\alpha_1(\nu-1)+1/2}.$$

Proof. Of course, in view of Corollary 5.5 it remains to prove the upper estimate for $l_n(R^d_{\underline{\alpha}})$. Furthermore, a direct application of Propositions 5.2 and 4.4 shows that it suffices to treat the case $\underline{\alpha} = (\alpha \ldots, \alpha)$ for some $\alpha > \frac{1}{2}$.

Step 1. We assume $\frac{1}{2} < \alpha \leqslant 1$. Given $n \in \mathbb{N}_0$, let

$$J_n := \left\{ k \in \mathbb{N}_0^d:\ 2^{n-1} < \prod_{j=1}^d \max\{1, k_j\} \leqslant 2^n \right\}.$$

It may be proved either by induction over d or by an integration argument that

$$\#(J_n) \leqslant c n^{d-1} 2^n. \tag{5.17}$$

Thus, if $S_n:\ L_2[0,1]^d \to C[0,1]^d$ is defined by

$$S_n f := \sum_{k \in J_n} \langle f, \vec{\psi}_k \rangle R^d_{\underline{\alpha}} \vec{\psi}_k,$$

(cf. (5.14) for the definition of the $\vec{\psi}_k$), we obtain

$$\operatorname{rank}(S_n) \leqslant c n^{d-1} 2^n \tag{5.18}$$

and, moreover,

$$R^d_{\underline{\alpha}} = \sum_{n=0}^\infty S_n. \tag{5.19}$$

Thus we have to make a suitable estimate of $l(S_n)$, and to do so, we use a method from Dunker *et al*. (1999). For some $\varepsilon_n > 0$ to be specified later on, let $G_n \subset [0,1]^d$ be an ε_n-net (with respect to the supremum norm in $\mathbb{R}^d$) of cardinality $\#(G_n) \leqslant c \cdot \varepsilon_n^{-d}$.

If $x \in G_n$, we define

$$Q_x := \{t \in [0,1]^d:\ \|x - t\|_\infty < \varepsilon_n\},$$

and by construction

$$[0, 1]^d = \bigcup_{x \in G_n} Q_x.$$

By the equivalence of the moments of a Gaussian vector (see, e.g., Lifshits 1995) there exists some $c > 0$ such that

$$\begin{aligned} c \cdot l(S_n) &\leqslant \mathbb{E} \sup_{t \in [0,1]^d} \left| \sum_{k \in J_n} \xi_k (R^d_{\underline{\alpha}} \vec{\psi}_k)(t) \right| \\ &\leqslant \mathbb{E} \sup_{x \in G_n} \sup_{t \in Q_x} \left| \sum_{k \in J_n} \xi_k [(R^d_{\underline{\alpha}} \vec{\psi}_k)(x) - (R^d_{\underline{\alpha}} \vec{\psi}_k)(t)] \right| + \mathbb{E} \sup_{x \in G_n} \left| \sum_{k \in J_n} \xi_k (R^d_{\underline{\alpha}} \vec{\psi}_k)(x) \right|. \end{aligned} \tag{5.20}$$

Our next aim is to estimate both terms in the last expression separately. Let us start with the first term in (5.20). Recall that R_α maps $L_2[0, 1]$ into the Hölder space $C^{\alpha - 1/2}[0, 1]$, hence, if $t \in Q_x$, for all $k \geqslant 0$ and $1 \leqslant j \leqslant d$ we have

$$|(R_\alpha \psi_k)(x_j) - (R_\alpha \psi_k)(t_j)| \leqslant c \cdot \varepsilon_n^{\alpha - 1/2}. \tag{5.21}$$

Using

$$\left| \prod_{j=1}^d a_j - \prod_{j=1}^d b_j \right| \leqslant \kappa^{d-1} \cdot \sum_{j=1}^d |a_j - b_j|,$$

with $\kappa := \max\{|a_1|, \ldots, |a_d|, |b_1|, \ldots, |b_d|\}$, from (5.21) we derive

$$|(R^d_{\underline{\alpha}} \vec{\psi}_k)(x) - (R^d_{\underline{\alpha}} \vec{\psi}_k)(t)| \leqslant c \cdot \varepsilon_n^{\alpha - 1/2}$$

provided that $t \in Q_x$. Hence for the first term in (5.20) we obtain

$$\begin{aligned} &\mathbb{E} \sup_{x \in G_n} \sup_{t \in Q_x} \left| \sum_{k \in J_n} \xi_k [(R^d_{\underline{\alpha}} \vec{\psi}_k)(x) - (R^d_{\underline{\alpha}} \vec{\psi}_k)(t)] \right| \\ &\qquad \leqslant c \cdot \varepsilon_n^{\alpha - 1/2} \cdot \mathbb{E} \left(\sum_{k \in J_n} |\xi_k| \right) = c' \cdot \#(J_n) \varepsilon_n^{\alpha - 1/2} \leqslant c \cdot n^{d-1} \cdot 2^n \varepsilon_n^{\alpha - 1/2}. \end{aligned} \tag{5.22}$$

To estimate the second term in (5.20) suitably, we use the following result (Pisier 1989, Lemma 4.14). Let $Z_1, \ldots, Z_N$ be a centred Gaussian sequence. Then it follows that

$$\mathbb{E} \sup_{1 \leqslant k \leqslant N} |Z_k| \leqslant c \cdot (1 + \log N)^{1/2} \sup_{1 \leqslant k \leqslant N} (\mathbb{E}|Z_k|^2)^{1/2}.$$

Applying this to the second term in (5.20) gives

$$\begin{aligned}
&\mathbb{E}\sup_{x\in G_n}\Big|\sum_{k\in J_n}\xi_k(R^d_{\underline{\alpha}}\vec{\psi}_k)(x)\Big|\\
&\leqslant c\cdot(1+\log\#G_n)^{1/2}\sup_{x\in G_n}\left(\mathbb{E}\Big|\sum_{k\in J_n}\xi_k(R^d_{\underline{\alpha}}\vec{\psi}_k)(x)\Big|^2\right)^{1/2}\\
&\leqslant c\cdot(1+|\log\varepsilon_n|)^{1/2}\sup_{x\in G_n}\left(\sum_{k\in J_n}|(R^d_{\underline{\alpha}}\vec{\psi}_k)(x)|^2\right)^{1/2}\\
&\leqslant c\cdot(1+|\log\varepsilon_n|)^{1/2}\cdot\#(J_n)^{1/2}\sup_{k\in J_n}\sup_{x\in[0,1]^d}|(R^d_{\underline{\alpha}}\vec{\psi}_k)(x)|.
\end{aligned}$$

Consequently, by Lemma 5.6 and by the choice of J_n,

$$\mathbb{E}\sup_{x\in G_n}\Big|\sum_{k\in J_n}\xi_k(R^d_{\underline{\alpha}}\vec{\psi}_k)(x)\Big| \leqslant c\cdot(1+|\log\varepsilon_n|)^{1/2}\cdot n^{d/2-1/2}2^{n/2-n\alpha}. \tag{5.23}$$

Now we choose $\varepsilon_n > 0$ such that the first term in (5.20) is of lower order than the second. With $\rho > 0$ satisfying

$$\rho > \frac{2\alpha+1}{2\alpha-1}, \tag{5.24}$$

we set $\varepsilon_n := 2^{-\rho n}$. Then (5.20), (5.22) and (5.23) imply

$$\begin{aligned}
l(S_n) &\leqslant c\cdot\{n^{d-1}2^{-n(\rho(\alpha-1/2)-1)} + n^{d/2}2^{-n(\alpha-1/2)}\}\\
&\leqslant c\cdot n^{d/2}2^{-n(\alpha-1/2)} \qquad (5.25)
\end{aligned}$$

by using (5.24). Summing up, in view of (5.19), (5.18) and (5.25), by Proposition 2.1 we finally obtain

$$l_n(R^d_{\underline{\alpha}}) \preccurlyeq n^{-\alpha+1/2}(1+\log n)^{\alpha(d-1)+1/2}$$

as asserted.

Step 2. We now suppose $\underline{\alpha} = (\alpha, \ldots, \alpha)$ with $\alpha > 1$. Setting $\underline{\beta} := (\alpha-1, \ldots, \alpha-1)$ and $\underline{1} := (1, \ldots, 1)$ we split $R^d_{\underline{\alpha}}$ as

$$[R^d_{\underline{\alpha}}:\ L_2[0,1]^d \to C[0,1]^d] = [R^d_{\underline{1}}:\ L_2 \to C] \circ [R^d_{\underline{\beta}}:\ L_2 \to L_2].$$

Hence by (2.7), by step 1 above and by Proposition 5.4, we obtain

$$\begin{aligned}
l_{2n-1}(R^d_{\underline{\alpha}}:\ L_2 \to C) &\leqslant l_n(R^d_{\underline{1}}:\ L_2 \to C)\cdot a_n(R^d_{\underline{\beta}}:\ L_2 \to L_2)\\
&\leqslant c\cdot n^{-1/2}(1+\log n)^{d-1/2}\cdot n^{-\beta}(1+\log n)^{\beta(d-1)}\\
&= c\cdot n^{-\alpha+1/2}(1+\log n)^{\alpha(d-1)+1/2},
\end{aligned}$$

completing the proof also in this case. □

A consequence of Theorem 5.7 is as follows.

Proposition 5.8. *For* $\underline{\alpha} = (\alpha_1, \ldots, \alpha_d)$ *as before it follows that*

$$a_n(R^d_{\underline{\alpha}}:\ L_2 \to C) \approx n^{-\alpha_1+1/2}(1+\log n)^{\alpha_1(\nu-1)}.$$

Proof. In view of Proposition 3.2, we have

$$a_{2n-1}(R^d_{\underline{\alpha}}:\ L_2 \to C) \leqslant c\cdot(1+\log n)^{-1/2}\cdot l_n(R^d_{\underline{\alpha}}:\ L_2 \to C),$$

so that the upper estimate follows from Theorem 5.7.

From (4.3) we derive

$$a_{2n-1}(R^d_{\underline{\alpha}}:\ L_2 \to L_2) \leqslant n^{-1/2}\cdot a_n(R^d_{\underline{\alpha}}:\ L_2 \to C). \tag{5.26}$$

Hence, the lower estimate is a direct consequence of Proposition 5.4 and (5.26). □

6. Application to Gaussian random variables

Let X be a centred Gaussian random variable with values in a Banach space E, i.e. X attains a.s. values in a separable subspace of E and, moreover, given any $x^* \in E^*$, the dual of E, the random variable $\langle X, x^*\rangle$ is centred normal (possibly degenerated). Each such X can be represented in the form

$$X = \sum_{k=1}^{\infty} \xi_k x_k \text{ a.s.} \tag{6.1}$$

for suitable x_k in E and the ξ_k independent standard normal. Now let H be some separable Hilbert space and $(f_k)_{k\geqslant 1}$ some ONB in H. Then by setting

$$T(f_k) := x_k, \qquad k = 1, 2, \ldots,$$

the operator $T:\ H \to E$ is well defined and bounded. Hence (6.1) now becomes

$$X = \sum_{k=1}^{\infty} \xi_k T f_k \text{ a.s.,} \tag{6.2}$$

and with $l_n(X)$ defined in (1.2) we easily obtain

$$l_n(X) = l_n(T).$$

Consequently, all properties of l-numbers of operators carry over to those of Banach space valued centred Gaussian random variables. Let us state a first application.

We say that a centred Gaussian random variable X has rank n whenever there are linearly independent elements $x_1, \ldots, x_n$ in E such that $X = \sum_{k=1}^{n} \xi_k x_k$ a.s. Now Proposition 2.1 implies the following.

Proposition 6.1. *If $\alpha > 0$ and $\beta \in \mathbb{R}$, the following statements are equivalent:*

(i)

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

(ii) *There exists a representation* (6.1) *of X such that*

$$\left(\mathbb{E}\left\|\sum_{k=n}^{\infty} \xi_k x_k\right\|^2\right)^{1/2} \preccurlyeq n^{-\alpha}(1 + \log n)^{\beta}.$$

(iii) *There are* (*independent*) *E-valued random variables X_k such that*

$$\operatorname{rank}(X_k) < 2^k, \qquad (\mathbb{E}\|X_k\|^2)^{1/2} \leqslant c{\cdot}k^{\beta}2^{-\alpha k}, \qquad X = \sum_{k=1}^{\infty} X_k.$$

In particular, Proposition 6.1 tells us the following. Let $(X(t))_{t\in K}$ be a centred Gaussian process over a compact set K possessing a.s. continuous paths. Then we find continuous functions u_k on K with

$$X(t) = \sum_{k=1}^{\infty} \xi_k u_k(t), \qquad t \in K,$$

and

$$\left(\mathbb{E} \sup_{t\in K}\left|\sum_{k=n}^{\infty} \xi_k u_k(t)\right|^2\right)^{1/2} \preccurlyeq n^{-\alpha}(1 + \log n)^{\beta}$$

if and only if

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

Finally, we treat another problem. Suppose that an operator $T\colon H \to E$ and a random variable X are connected by (6.2). A natural question is now how the approximation numbers of T may be described by properties of X only. Note that neither T nor the ONB $(f_k)_{k\geqslant 1}$ in (6.2) is unique, yet it is not difficult to check that $a_n(T)$ depends only on X. In other words, we may define approximation numbers of X by

$$a_n(X) := a_n(T)$$

provided T and X are related by (6.2). To give an intrinsic description of $a_n(X)$, we first note that

$$a_n(T) = d_n(T^*),$$

where $d_n(T^*)$ denotes the nth Kolmogorov number of the dual operator $T^*\colon E^* \to H$; see Pietsch (1987) for the definition of those numbers. We now choose H and T such that T is connected to X as in (6.2). Suppose that X is defined on a probability space $(\Omega, \mathbb{P})$ and define $H \subseteq L_2(\Omega, \mathbb{P})$ as the closure of

$$H_0 := \{\langle X, x^*\rangle\colon x^* \in E^*\}.$$

Given $f \in H$, we set

$$Tf := \int_\Omega f(\omega)X(\omega)\mathrm{d}\mathbb{P}(\omega),$$

where this integral has to be understood in the sense of Bochner. It is easy to prove that $T\colon H \to E$ is injective and is connected to X. Moreover, its dual T^* satisfies

$$T^*x^* = \langle X, x^*\rangle, \qquad x^* \in E^*. \tag{6.3}$$

Consequently, letting

$$B_X := \{\langle X, x^*\rangle : \|x^*\|_{E^*} \leqslant 1\} \subseteq H,$$

we obtain

$$a_n(X) = \inf\{\varepsilon > 0 : d(B_X, H_{n-1}) < \varepsilon \text{ for some } H_{n-1} \subseteq H, \ \dim(H_{n-1}) = n-1\} \tag{6.4}$$

(as usual by d we denote the distance between two sets, generated by the underlying norm, here in H). Next, recall that H_0 is dense in H and B_X is bounded. Thus, whenever there is a subspace $H_{n-1} \subset H$ as in (6.4) satisfying $d(B_X, H_{n-1}) < \varepsilon$, then we also find a subspace $H^0_{n-1} \subset H_0$ of dimension $n-1$ possessing exactly the same property. Summing up, we obtain the following.

Proposition 6.2 *We have $a_n(X) < \varepsilon$ if and only if there are $x_1^*, \ldots, x_{n-1}^* \in E^*$ such that, for all $x^* \in E^*$, we find $\lambda_1, \ldots, \lambda_{n-1} \in \mathbb{R}$ with*

$$\left(\mathbb{E}\left|\langle X, x^*\rangle - \sum_{j=1}^{n-1} \lambda_j \langle X, x_j^*\rangle\right|^2\right)^{1/2} < \varepsilon \cdot \|x^*\|. \tag{6.5}$$

Let us specify Proposition 6.2 further. To do so, first notice that by Lebesgue's theorem the operator T^* given by (6.3) is $[\sigma(E^*, E), \|\cdot\|_H]$-continuous on bounded subsets of E^*, hence if $A \subset E^*$ denotes the set of extreme points in the unit sphere of E^*, then it suffices that (6.5) is satisfied for all $x^* \in A$.

An interesting application is as follows. Suppose that $X = (X(t))_{t\in K}$ is a centred Gaussian process over the compact metric space K with a.s. continuous sample paths. Then X may be regarded as a $C(K)$-valued Gaussian random variable. In this case we have $a_n(X) < \varepsilon$ if and only if there are (signed) measures $u_1, \ldots, u_{n-1}$ over K such that for each $s \in K$ we find reals $\lambda_1, \ldots, \lambda_{n-1}$ such that

$$\left(\mathbb{E}\left|X(s) - \sum_{j=1}^{n-1} \lambda_j \int_K X(t)\mathrm{d}\mu_j(t)\right|^2\right)^{1/2} < \varepsilon.$$

In other words, up to an error of at most ε, the values of X can be linearly reconstructed from $n-1$ functionals of the process.

7. Fractional Brownian sheets

We start with an application of Proposition 4.2 (or Corollary 4.3) to Gaussian processes over [0, 1]. This implies a quite general upper estimate of $l_n(X)$ for those processes. Similar assertions for the small-ball behaviour can be found in Csörgő and Shao (1994).

Proposition 7.1. *Let X be a centred Gaussian process over* [0, 1] *such that, for some* $\gamma > 0$ *and* $\beta \in \mathbb{R}$,

$$\mathbb{E}|2X(t) - X(t+\tau) - X(t-\tau)|^2 \leqslant c \cdot \tau^{2\gamma} \left(\log \frac{1}{\tau} \right)^{2\beta} \tag{7.1}$$

for $0 \leqslant t - \tau < t < t + \tau \leqslant 1$. *Then this implies that*

$$l_n(X) \preccurlyeq n^{-\gamma}(1 + \log n)^{\beta+1/2}. \tag{7.2}$$

In particular, (7.1) *is satisfied if*

$$\mathbb{E}|X(t) - X(t')|^2 \leqslant c \cdot |t - t'|^{2\gamma} \left(\log \frac{1}{|t - t'|} \right)^{2\beta}$$

for $0 \leqslant t \neq t' \leqslant 1$.

Remark. It is worth mentioning that a good approximation of X is obtained by

$$X_n(t) := \sum_{m=-2}^{n} \sum_{j \in J_m} \langle X, \varphi_{m,j} \rangle u_{m,j}(t)$$

(we use the notation of Section 4). We have $\operatorname{rank}(X_n) \leqslant 2^{n+2}$ and

$$\left(\mathbb{E} \sup_{0 \leqslant t \leqslant 1} |X(t) - X_n(t)|^2 \right)^{1/2} \preccurlyeq n^{\beta+1/2} 2^{-\gamma n},$$

provided (7.1) is satisfied. By Theorem 7.4 below, this order is optimal for fractional Brownian motion.

Our next objective is to apply the former results to fractional Brownian sheets defined in (1.4). First, we have to find a Hilbert space H and suitable operators $T_{\underline{\gamma}}$ from H into $C[0, 1]^d$ related to $B^d_{\underline{\gamma}}$ in the sense of (6.2). Let us start with the one-dimensional case. If $0 < \gamma < 2$, let $H = L_2[0, 1] \oplus L_2(-\infty, 0]$ and define T_γ from H to $C[0, 1]$ by

$$T_\gamma := \kappa_\gamma (R_{(\gamma+1)/2} \oplus Q_\gamma), \tag{7.3}$$

where

$$(Q_\gamma f)(t) = \frac{1}{\Gamma((\gamma+1)/2)} \int_{-\infty}^{0} \left((t-s)^{(\gamma-1)/2} - (-s)^{(\gamma-1)/2} \right) f(s) \mathrm{d}s, \tag{7.4}$$

the operator $R_{(\gamma+1)/2} \colon L_2[0, 1] \to C[0, 1]$ is the Riemann–Liouville integration operator of order $(\gamma+1)/2$ defined in (4.14) and

$$\kappa_\gamma = \Gamma\left(\frac{\gamma+1}{2}\right)\left(\gamma^{-1} + \int_{-\infty}^{0}\left((1-s)^{(\gamma-1)/2} - (-s)^{(\gamma-1)/2}\right)^2 \mathrm{d}s\right)^{-1/2}. \tag{7.5}$$

Since, for $t, s \in [0, 1]$,

$$\langle T_\gamma^* \delta_t, T_\gamma^* \delta_s \rangle = \tfrac{1}{2}(|s|^\gamma + |t|^\gamma - |s-t|^\gamma)$$

(Mandelbrot and Van Ness 1968), the following is valid.

Proposition 7.2. *For any ONB* $(f_k)_{k=1}^\infty$ *in* $H = L_2[0, 1] \oplus L_2(-\infty, 0]$, *the series*

$$\sum_{k=1}^{\infty} \xi_k T_\gamma f_k$$

converges a.s. in $C[0, 1]$ *and defines a* γ*-fractional Brownian motion.*

Now let $\underline{\gamma} = (\gamma_1, \ldots, \gamma_d)$ be a multi-index with $0 < \gamma_j < 2$. We define on operator $T_{\underline{\gamma}}^d$ from $H^{\otimes d}$ to $C[0, 1]^d$ as

$$T_{\underline{\gamma}}^d := T_{\gamma_1} \otimes \cdots \otimes T_{\gamma_d} \tag{7.6}$$

and obtain the following d-dimensional version of Proposition 7.2.

Proposition 7.3. *The operator* $T_{\underline{\gamma}}^d$ *is related to* $B_{\underline{\gamma}}^d$, *i.e. for each ONB* $(f_k)_{k=1}^\infty$ *in* $H^{\otimes d}$ (*here H is in Proposition* 7.2) *the series*

$$\sum_{k=1}^{\infty} \xi_k T_{\underline{\gamma}}^d f_k$$

converges a.s. in $C[0, 1]^d$ *and defines a* γ*-fractional Brownian sheet.*

For later purposes let us state the following properties of $T_{\underline{\gamma}}^d$.

(a) Note that the approximation numbers, and hence also the l-numbers, of the operator Q_γ defined in (7.4) tend to zero exponentially (Belinsky and Linde 2002). In particular, it follows that

$$l_n(Q_\gamma) \preccurlyeq n^{-\rho} \tag{7.7}$$

for any $\rho > 0$. Li and Linde (1998) give an estimate for the entropy of Q_γ implying $l_n(Q_\gamma) \preccurlyeq n^{-1}(1 + \log n)^2$. This would suffice for later purposes as well.

(b) Representation (7.3) implies

$$\begin{aligned} T_{\underline{\gamma}}^d &= \left(\prod_{j=1}^{d} \kappa_{\gamma_j}\right)\left[\left(R_{(\gamma_1+1)/2} \otimes \cdots \otimes R_{(\gamma_d+1)/2}\right) + \sum_{i=1}^{2^d-1} V_i\right] \\ &= \kappa\left(R_{(\underline{\gamma}+1)/2}^d + \sum_{i=1}^{2^d-1} V_i\right) \end{aligned} \tag{7.8}$$

with the κ_{γ_j} defined by (7.5), $\kappa = \prod_{j=1}^{d} \kappa_{\gamma_j}$ and the V_i given by

$$V_i = S_1 \otimes \cdots \otimes S_d,$$

where S_j is either $R_{(\gamma_j+1)/2}$ or Q_{γ_j} and, moreover, the latter case appears at least for one j.

(c) Denote by $P\colon\ H^{\otimes d} \to L_2[0,1]^d$ the orthogonal projection defined by

$$Pf := f \cdot 1_{[0,1]^d}, \qquad f \in H^{\otimes d}.$$

By (7.3) and (7.6), we have

$$T^d_{\underline{\gamma}} \circ P = \kappa \cdot R^d_{(\underline{\gamma}+1)/2}$$

with $\kappa > 0$ as in (7.8), hence

$$l_n(R^d_{(\underline{\gamma}+1)/2}) \leqslant \kappa^{-1} \cdot l_n(T^d_{\underline{\gamma}}). \tag{7.9}$$

Thus prepared, we can now prove the main result of the paper.

Theorem 7.4. *Let* $\underline{\gamma} = (\gamma_1, \ldots, \gamma_d)$ *and suppose that*

$$0 < \gamma_1 = \ldots = \gamma_\nu < \gamma_{\nu+1} \leqslant \ldots \leqslant \gamma_d < 2.$$

Then it follows that

$$l_n(B^d_{\underline{\gamma}}) \approx n^{-\gamma_1/2}(1+\log n)^{\nu(\gamma_1+1)/2-\gamma_1/2}.$$

Proof. We start with the upper estimate for $\nu = d$, i.e. $\underline{\gamma} = (\gamma, \ldots, \gamma)$ for a certain $\gamma \in (0,2)$. By (7.8) we have

$$T^d_{\underline{\gamma}} = \kappa\left(R^d_{(\underline{\gamma}+1)/2} + \sum_{i=1}^{2^d-1} V_i \right) \tag{7.10}$$

with the V_i as defined there. In our situation (perhaps after some permutation of the coordinates) each operator V_i admits the representation

$$V_i = Y_1 \otimes Y_2, \tag{7.11}$$

where Y_1 is the tensor product of a certain number, say $\tilde{d}$, of operators $R_{(\gamma+1)/2}$ while Y_2 is the $(d-\tilde{d})$-fold tensor power of Q_γ. Note that necessarily $0 \leqslant \tilde{d} < d$. Using (7.7), an application of Proposition 5.3 implies

$$l_n(Y_2) \preccurlyeq n^{-\rho} \tag{7.12}$$

for all $\rho > 0$. On the other hand, from Theorem 5.7 we derive

$$l_n(Y_1) \preccurlyeq n^{-\gamma/2}(1+\log n)^{\tilde{d}(\gamma+1)/2-\gamma/2}. \tag{7.13}$$

If we combine (7.11), (7.12) and (7.13), Proposition 5.2 yields

$$l_n(V_i) \leqslant n^{-\gamma/2}(1+\log n)^{\beta}, \tag{7.14}$$

where

$$\beta = \frac{\tilde{d}(\gamma+1)}{2} - \frac{\gamma}{2} < \frac{d(\gamma+1)}{2} - \frac{\gamma}{2}. \tag{7.15}$$

Of course, by (7.12) estimate (7.14) also holds for $Y_1 = 0$, i.e. $\tilde{d} = 0$. If we now compare (7.14) and (7.15) with

$$l_n(R^d_{(\underline{\gamma}+1)/2}) \approx n^{-\gamma/2}(1+\log n)^{d(\gamma+1)/2-\gamma/2}, \tag{7.16}$$

by (2.6) we derive, using (7.10),

$$l_n(T^d_{\underline{\gamma}}) \leqslant l_n(R^d_{(\underline{\gamma}+1)/2}).$$

Consequently, (7.16) gives the desired upper estimate for $l_n(T^d_{\underline{\gamma}})$ in the case $\nu = d$.

In the general situation, i.e. $\underline{\gamma} = (\gamma_1, \ldots, \gamma_d)$, we write

$$T^d_{\underline{\gamma}} = T^{\nu}_{\underline{\gamma_1}} \otimes T_{\gamma_{\nu+1}} \otimes \cdots \otimes T_{\gamma_d} \tag{7.17}$$

with $\underline{\gamma_1} := (\gamma_1, \ldots, \gamma_\nu)$ Recalling $\gamma_1 = \ldots = \gamma_\nu$, from the previous paragraph it follows that

$$l_n(T^{\nu}_{\underline{\gamma_1}}) \leqslant n^{-\gamma_1/2}(1+\log n)^{\nu(\gamma_1+1)/2-\gamma_1/2}, \tag{7.18}$$

while Proposition 4.4 gives

$$l_n(T_{\gamma_j}) \approx n^{-\gamma_j/2}(1+\log n)^{1/2}$$

for $j = \nu+1, \ldots, d$. Since $\gamma_j > \gamma_1$ for those j, an iterative application of Proposition 5.2 to (7.17) shows that $l_n(T^d_{\underline{\gamma}})$ admits the same upper estimate as $l_n(T^d_{\underline{\gamma_1}})$ in (7.18). This completes the proof of the upper estimate.

The lower estimate of $l_n(T^d_{\underline{\gamma}})$ is a direct consequence of (7.9) and Theorem 5.7. This concludes the proof of Theorem 7.4. □

Let us state a first application of Theorem 7.4.

Corollary 7.5. *For the Gaussian process $B^d_{\underline{\gamma}}$ we have*

$$\begin{aligned} \varepsilon^{-2/\gamma_1}(\log \varepsilon^{-1})^{(\nu-1)(1+1/\gamma_1)} &\leqslant -\log \mathbb{P}\left(\sup_{t\in[0,1]^d} |B^d_{\underline{\gamma}}(t)| < \varepsilon \right) \\ &\leqslant \varepsilon^{-2/\gamma_1}(\log \varepsilon^{-1})^{(\nu-1)(1+1/\gamma_1)+1/\gamma_1}. \end{aligned} \tag{7.19}$$

Proof. The upper estimate in (7.19) follows directly from the preceding theorem and Proposition 4.1 in Li and Linde (1999).

For the lower estimate let $J_d : C[0, 1]^d \to L_2[0, 1]^d$ be the canonical embedding and define

$$X^d_{\underline{\gamma}} := J_d \circ B^d_{\underline{\gamma}}.$$

Then

$$-\log \mathbb{P}(\|X^d_{\underline{\gamma}}\|_{L_2} < \varepsilon) \leqslant -\log \mathbb{P}\left(\sup_{t\in[0,1]^d} |B^d_{\underline{\gamma}}(t)| < \varepsilon \right).$$

But it is known – or follows from Theorem 1.2 in Li and Linde (1999) and properties of entropy numbers of operators acting between Hilbert spaces – that

$$-\log \mathbb{P}(\|X^d_{\underline{\gamma}}\|_{L_2} < \varepsilon) \approx \varepsilon^{-2/\gamma_1}(\log \varepsilon^{-1})^{(\nu-1)(1+1/\gamma_1)}.$$

This completes the proof. □

Remark. For $d = 1$ the left-hand estimate in (7.19) is known to give the precise order (Shao 1993; Monrad and Rootzén 1995), while for $d = 2$ and $\underline{\gamma} = (\gamma, \gamma)$ the right-hand expression is the correct one (Talagrand 1994; Belinsky and Linde 2002). If $d > 2$, the exact small-ball behaviour of $B^d_{\underline{\gamma}}$ is an open problem. Estimate (7.19) was proved in Dunker et al. (1999) for the Brownian sheet.

Another corollary of Theorem 7.4 leads to the following information about the approximation numbers of $B^d_{\underline{\gamma}}$ (with respect to the supremum norm).

Corollary 7.6. *It follows that*

$$a_n(B^d_{\underline{\gamma}}) \approx n^{-\gamma_1/2}(1 + \log n)^{(\gamma_1+1)(\nu-1)/2}.$$

Proof. The upper estimate follows directly from Theorem 7.4 and Proposition 3.2, while the lower estimate is an immediate consequence of (7.9) and Theorem 5.7. □

Final remark. Given a Gaussian random variable X, we say that a representation $X = \sum_{k=1}^{\infty} \xi_k x_k$ is optimal provided that

$$\left(\mathbb{E}\left\| \sum_{k=n}^{\infty} \xi_k x_k \right\|^2\right)^{1/2} \approx l_n(X).$$

A careful inspection of the proofs in Theorems 5.7 and 7.4 gives some information about optimal representations of $B^d_{\underline{\gamma}}$ at least if $\gamma_1 \leqslant 1$. Let us explain this for $d = 1$. Here we have

$$B_\gamma(t) = \sum_{k=1}^{\infty} \xi_k u_k(t) + \sum_{l=1}^{\infty} \xi'_l v_l(t) \tag{7.20}$$

with $(\xi_k)_{k\geqslant 1}$ and $(\xi'_l)_{l\geqslant 1}$ independent,

$$u_k = \kappa_\gamma \cdot R_{(\gamma+1)/2}\psi_k \quad \text{and} \quad v_l = \kappa_\gamma \cdot Q_\gamma f_l$$

for a certain ONB $(f_l)_{l\geqslant 1}$ in $L_2(-\infty, 0]$. If $\gamma \leqslant 1$ then, by Theorem 7.4, we have

$$\left(\mathbb{E}\left\|\sum_{k=n}^{\infty}\xi_k u_k\right\|_{\infty}^{2}\right)^{1/2} \approx l_n(B_\gamma),$$

while by the results in Belinsky and Linde (2002) the f_k may be chosen such that

$$\left(\mathbb{E}\left\|\sum_{l=n}^{\infty}\xi_l' v_l\right\|^{2}\right)^{1/2} \preccurlyeq n^{-\rho}$$

for any $\rho > 0$. Thus, after suitably combining the two sums in (7.20), this leads to an optimal representation of B_γ. We do not know if a similar assertion is true for $1 < \gamma < 2$.

Acknowledgement

The authors are very grateful to M.A. Lifshits for several very helpful discussions on the subject of the paper. In particular, Proposition 3.2 goes back to his ideas. Furthermore, we thank the referee for pointing out an error in an earlier version of the proof of Theorem 5.7.

References

Belinsky, E. and Linde, W. (2002) Small ball probabilities of fractional Brownian sheets via fractional integration operators. *J. Theoret. Probab.* To appear.

Carl, B. and Stephani, I. (1990) *Entropy, Compactness and Approximation of Operators*. Cambridge: Cambridge University Press.

Csörgő, M. and Shao, Q.M. (1994) On almost sure limit inferior for *B*-valued stochastic processes and applications. *Probab. Theory Related Fields*, **99**, 29–54.

Dunker, T., Kühn, T., Lifshits, M.A. and Linde, W. (1999) Metric entropy of integration operators and small ball probabilities for the Brownian sheet. *J. Approx. Theory*, **101**, 63–77.

König, H. (1989) *Eigenvalue Distribution of Compact Operators*. Boston: Birkhäuser.

Ledoux, M. and Talagrand, M. (1991) *Probability in a Banach Space*. Berlin: Springer-Verlag.

Li, W.V. and Linde, W. (1998) Existence of small ball constants for fractional Brownian motions. *C. R. Acad. Sci. Paris Sér. I. Math.*, **326**, 1329–1334.

Li, W.V. and Linde, W. (1999) Approximation, metric entropy and small ball estimates for Gaussian measures. *Ann. Probab.*, **27**, 1556–1578.

Lifshits, M.A. (1995) *Gaussian Random Functions*. Dordrecht: Kluwer.

Maiorov, V.E. and Wasilkowski, G.W. (1996) Probabilistic and average linear widths in L_∞-norm with respect to *r*-fold Wiener measure. *J. Approx. Theory*, **84**, 31–40.

Mandelbrot, B.B. and Van Ness, J.W. (1968) Fractional Brownian motions, fractional noises and applications. *SIAM Rev.*, **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.

Pietsch, A. (1978) *Operator Ideals*. Berlin: Verlag der Wissenschaften.

Pietsch, A. (1987) *Eigenvalues and s-Numbers*. Cambridge: Cambridge University Press.

Pisier, G. (1989) *The Volume of Convex Bodies and Banach Space Geometry*. Cambridge: Cambridge University Press.

Shao, Q.M. (1993) A note on small ball probability of Gaussian processes with stationary increments. *J. Theoret. Probab.*, **6**, 595–602.

Talagrand, M. (1994) The small ball problem for the Brownian sheet. *Ann. Probab.*, **22**, 1331–1354.

Vu, K.T. and Gorenflo, R. (1996) Asymptotics of singular values of Volterra integral operators. *Numer. Funct. Anal. Optim.*, **17**, 453–461.

Received May 2001 and revised April 2002