JOURNAL OF FUNCTIONAL ANALYSIS 116, 133–157 (1993)

Metric Entropy and the Small Ball Problem for Gaussian Measures

JAMES KUELBS*

Department of Mathematics, University of Wisconsin–Madison, Madison, Wisconsin 53706

AND

WENBO V. LI

Department of Mathematical Sciences, University of Delaware, Newark, Delaware 19716

Communicated by L. Gross

Received October 16, 1992

We establish a precise link between the small ball problem for a Gaussian measure μ on a separable Banach space and the metric entropy of the unit ball of the Hilbert space H_μ generating μ. This link allows us to compute small ball probabilities from metric entropy results, and vice versa.

1. INTRODUCTION

Let μ denote a centered Gaussian measure on a real separable Banach space B with norm $\|\cdot\|$ and dual B^*. If K is the unit ball of the Hilbert space H_μ which generates μ, then it is well known that

$$\lim_{t\to\infty} t^{-2}\log\mu(x:\|x\|\geqslant t) = -(2\sigma^2)^{-1},$$

where

$$\sigma^2 = \sup_{\|f\|_{B^*}\leqslant 1}\ \sup_{x\in K} f^2(x) = \sup_{\|f\|_{B^*}\leqslant 1}\int_B f^2(x)\,d\mu(x),$$

* Supported in part by NSF Grant DMS-9024961.

0022-1236/93 $5.00

and hence the distribution of the norm at infinity is, at the logarithmic level, a simple function of σ^2. The small ball problem studies this distribution near zero, namely, the behavior of

$$\log \mu(x : \|x\| \leqslant \varepsilon) = -\phi(\varepsilon) \qquad \text{as} \quad \varepsilon \to 0,$$

and here the behavior of $\phi(\varepsilon)$ depends on much more than the single parameter σ^2. Indeed, 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 $\varepsilon \to 0$. The point of this paper is to link the behavior of $\phi(\varepsilon)$ to the metric entropy of K. Hence as a parallel to the large ball behavior determined by the simple characteristic of K given by σ^2, the behavior of $\phi(\varepsilon)$ is governed by the more subtle metric entropy. This is a connection which is rather simple, but it links two delicate topics in a useful way. That is, once this link is obtained, then metric entropy results regarding K will yield information regarding $\phi(\varepsilon)$. Conversely, in instances when we know the behavior of $\phi(\varepsilon)$, we can establish some non-trivial and sometimes new results about the metric entropy of the various sets which appear as K. We include a sample of these applications in Section 5, but our primary results are Theorem 1 and 2 below. In Theorem 3, we show the somewhat surprising fact that in many instances randomly centered balls can cover $(1-\delta)K$, $0<\delta<1$, as efficiently as those placed in the best possible fashion required to compute metric entropy. A result related to Theorem 3 appeared earlier in [10].

The small ball problem for the standard Brownian sheet has recently been solved in [23]. Combined with Theorem 1 this yields an interesting result about the set $W_{2,(1,1)}^{(1,1)}$ of [24, Theorems 1.3 and 1.4]. We also mention the interesting partial result in [9] which was one of the starting points of this work.

If μ is a centered Gaussian measure on B, then it is well known that there is a unique Hilbert space $H_\mu \subseteq B$ such that μ is determined by considering the pair (B, H_μ) as an abstract Wiener space (see (12]). For example, if $B = C[0, 1]$ and μ is Wiener measure, then the unit ball of H_μ is

$$K = \left\{ f(t) = \int_0^t f'(s)\, ds,\, 0 \leqslant t \leqslant 1 : \int_0^1 |f'(s)|^2\, ds \leqslant 1 \right\}, \tag{1.1}$$

with the inner product norm given by

$$\|f\|_\mu = \left(\int_0^1 |f'(s)|^2\, ds \right)^{1/2} \qquad f \in H_\mu.$$

In general, H_μ can be described as the completion of the range of the mapping $S: B^* \to B$ defined via the Bochner integral

$$Sf = \int_B x f(x)\, d\mu(x), \qquad f \in B^*,$$

and the completion is in the inner product norm

$$\langle Sf, Sg \rangle_\mu = \int_B f(x)\, g(x)\, d\mu(x) \qquad f, g \in B^*.$$

Lemma 2.1 in [15] presents the details of this construction along with various properties of the relationship between H_μ and B, but the most important for us at this point is that the unit ball K of H_μ is always compact in the B-topology. Hence K has finite metric entropy.

To be precise we recall 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 $H(\varepsilon, A) = \log N(\varepsilon, A)$, where

$$N(\varepsilon, A) = \min \left\{ n \geq 1 : \exists a_1, ..., a_n \in A \text{ such that } \bigcup_{j=1}^{n} B_\varepsilon(a_j) \supseteq A \right\},$$

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

To state our results we use the notation $f(x) \approx g(x)$ as $x \to a$ if

$$0 < \varliminf_{x \to a} f(x)/g(x) \leq \varlimsup_{x \to a} f(x)/g(x) < \infty.$$

and we write $f(x) \preccurlyeq g(x)$ as $x \to a$ if $\varlimsup_{x \to a} f(x)/g(x) < \infty$. Throughout the paper $Lx = \max(\log_e x, 1)$, and all logarithms are natural logarithms.

The authors thank Alex de Acosta for his comments and interest in their results. We also thank Carl de Boor and Yuly Makovoz for their help regarding references to various metric entropy papers.

2. Statement of Theorems

Since the support of a centered Gaussian measure μ on B is the closure of H_μ in B, the behavior of $\mu(x : \|x\| \leq \varepsilon)$ is well understood if $\dim H_\mu < \infty$. This is the case since $\dim H_\mu = d < \infty$ easily implies

$$\mu(x : \|x\| \leq \varepsilon) \approx \varepsilon^d \qquad \text{as} \quad \varepsilon \to 0$$

and, if $N(\varepsilon, K)$ denotes the minimal number of open ε-balls in B which cover K, we also have

$$N(\varepsilon, K) \approx \varepsilon^{-d} \qquad \text{as} \quad \varepsilon \to 0.$$

Hence for the remainder of the paper we assume $\dim H_\mu = \infty$, but the reader should note that both Theorem 1 and Theorem 2 hold trivially when $\dim H_\mu < \infty$. Also, the assumption on $f(\varepsilon)$ in Theorem 3 is seen to require that $\dim H_\mu = \infty$, but the proof of Theorem 3 easily shows that when $\dim H_\mu < \infty$ a similar result can be obtained with $f(\varepsilon) = \log 1/\varepsilon$ and $g(x) = e^x$.

THEOREM 1. *Let μ be a centered Gaussian measure on a real separable Banach space B and let*

$$\log \mu(B_\varepsilon(0)) = -\phi(\varepsilon), \tag{2.1}$$

where $B_\varepsilon(h) = \{x \in B : \|x - h\| < \varepsilon\}$. If $f(1/x)$ is regularly varying at infinity with strictly positive finite constants c_1, c_2 such that

$$c_1 f(\varepsilon) \leqslant \phi(\varepsilon) \leqslant c_2 f(\varepsilon) \tag{2.2}$$

for $\varepsilon > 0$ small and

$$j(\varepsilon) = \varepsilon(4c_2 f(\varepsilon))^{-1/2}, \tag{2.3}$$

then

$$H(\varepsilon, K) \approx f(g(\varepsilon)) \qquad as \quad \varepsilon \to 0 \tag{2.4}$$

provided

$$g(j(\varepsilon)) \approx \varepsilon \qquad as \quad \varepsilon \to 0. \tag{2.5}$$

Remarks. (I) Since $\mu(B_\varepsilon(0))$ is continuous in ε with $\lim_{\varepsilon \to 0} \mu(B_\varepsilon(0)) = 0$, we have from (2.2) that $\lim_{\varepsilon \to 0} f(\varepsilon) = +\infty$. Hence, since $f(1/x)$ is assumed to be regularly varying at infinity, we have as $\varepsilon \to 0$ that

$$f(\varepsilon) = \varepsilon^{-\alpha} J(\varepsilon^{-1}), \tag{2.6}$$

where $\alpha \geqslant 0$ and $J(\cdot)$ is slowly varying at infinity. If α is strictly positive, then by [21, p. 23], $f(\varepsilon)$ can be assumed to be strictly increasing as $\varepsilon \to 0$, and hence $j(\varepsilon)$ is strictly decreasing as $\varepsilon \to 0$. Thus the inverse of $j(\varepsilon)$ exists, and we can take g to be the standard inverse of j. For example, if $f(\varepsilon)$ is as in (2.6) with J slowly varying, monotonic, and such that $J(x) \approx J(x^\rho)$ for each $\rho > 0$ as $x \to \infty$, then

$$j(\varepsilon) = \varepsilon^{(2+\alpha)/2}(4c_2 J(1/\varepsilon))^{-1/2}. \tag{2.7}$$

Hence if

$$g(\varepsilon) = \varepsilon^{2/(2+\alpha)} J(1/\varepsilon)^{1/(2+\alpha)}, \tag{2.8}$$

then the conditions imposed on $J(\cdot)$ imply

$$g(j(\varepsilon)) \approx \varepsilon \qquad \text{as} \quad \varepsilon \to 0. \tag{2.9}$$

Thus (2.4) implies

$$H(\varepsilon, K) \approx \varepsilon^{-2\alpha/(2+\alpha)} J(1/\varepsilon)^{2/(2+\alpha)} \qquad \text{as} \quad \varepsilon \to 0. \tag{2.10}$$

(II) The most prevalent form for $f(\varepsilon)$ is

$$f(\varepsilon) = \varepsilon^{-\alpha} (\log 1/\varepsilon)^{\beta}, \tag{2.11}$$

where $\alpha \geqslant 0$ and $\beta \in (-\infty, +\infty)$, and hence from the above as $\varepsilon \to 0$ we have that

$$H(\varepsilon, K) \approx \varepsilon^{-2\alpha/(2+\alpha)} (\log 1/\varepsilon)^{2\beta/(2+\alpha)}. \tag{2.12}$$

When $\beta = 0$ in (2.11), a one-sided estimate of (2.12) was obtained in [9].

(III) Perhaps it should be pointed out that the function $f(\varepsilon)$ is used in Theorem 1 because it is rare that $\phi(\varepsilon)$ is known precisely. Furthermore, if only the upper (lower) bound in (2.2) is known, then the upper (lower) bound result in (2.4) also follows from the proof of the theorem. Now we turn to the converse result.

THEOREM 2. *Let μ be a centered Gaussian measure on a real separable Banach space B, let $\phi(\varepsilon)$ be as in* (2.1), *and let K be the unit ball of the Hilbert space H_μ generating μ. If $g(1/x)$ is regularly varying at infinity and*

$$H(\varepsilon, K) \approx g(\varepsilon) \qquad \textit{as} \quad \varepsilon \to 0, \tag{2.13}$$

then

$$\phi(2\varepsilon) \preccurlyeq g(\varepsilon/(\phi(\varepsilon))^{1/2}) \preccurlyeq \phi(\varepsilon) \qquad \textit{as} \quad \varepsilon \to 0. \tag{2.14}$$

Furthermore, if $\phi(\varepsilon) \preccurlyeq \phi(2\varepsilon)$, then (2.14) *implies*

$$\phi(\varepsilon) \approx g(\varepsilon/(\phi(\varepsilon))^{12}) \qquad \textit{as} \quad \varepsilon \to 0. \tag{2.15}$$

In particular, if (2.15) *holds and if $g(\varepsilon) = \varepsilon^{-\beta} J(1/\varepsilon)$ where $0 < \beta < 2$ and $J(x)$ is slowly varying, monotonic, and such that $J(x) \approx J(x^{\rho})$ as $x \to \infty$ for each $\rho > 0$, then*

$$\phi(\varepsilon) \approx \varepsilon^{-2\beta/(2-\beta)} (J(1/\varepsilon))^{2/(2-\beta)} \qquad \textit{as} \quad \varepsilon \to 0. \tag{2.16}$$

Remarks. (I) The restriction on β in Theorem 2 is natural since it is known from [8] that $H(\varepsilon, K) = o(\varepsilon^{-2})$ regardless of the Gaussian measure μ (and hence the subsequent K). Also, see the remark following (3.15) below.

(II) Putting Theorem 1 and 2 together, it is easy to see that $\phi(\varepsilon) \approx \varepsilon^{-\alpha}$ $(\alpha > 0)$ iff $H(\varepsilon, K) \approx \varepsilon^{-2\alpha/(2+\alpha)}$, provided $\phi(\varepsilon) \preccurlyeq \phi(2\varepsilon)$ as $\varepsilon \to 0$.

(III) The remark (III) following Theorem 1 has a complete analogue for Theorem 2.

(IV) The assumption $\phi(\varepsilon) \preccurlyeq \phi(2\varepsilon)$ as $\varepsilon \to 0$ is easily seen to be equivalent to $\phi(\varepsilon) \preccurlyeq \phi((1+\delta)\varepsilon)$ as $\varepsilon \to 0$ for some $\delta > 0$. Nevertheless, it is an assumption that restricts the ease of application of Theorem 2, and although we cannot eliminate it, we indicate in Section 5 that it holds quite generally for Gaussian measures on Hilbert space.

It is also possible to prove a surprisingly sharp random version of Theorem 1, and we turn to this now. For example, if $X, X_1, ..., X_n$ are independent observations with common law μ and $F_n = \{X_1, ..., X_n\}$, then (2.21) of Theorem 3 asserts that for every $\delta \in (0, 1)$

$$P((1-\delta)\, K \subseteq F_n/(2Ln)^{1/2} + B_{\varepsilon_n}(0) \text{ eventually}) = 1,$$

where $\varepsilon_n \to 0$ is a function of the small ball behavior of μ. Thus the n spheres of radius ε_n centered at the point $X_j/(2Ln)^{1/2}$, $1 \leqslant j \leqslant n$, typically cover $(1-\delta)K$ when n is large, and hence

$$N(\varepsilon_n, (1-\delta)\, K) \leqslant n. \tag{2.17}$$

Solving (2.17) for ε, we obtain an upper bound for $N(\varepsilon, (1-\delta)K)$, and under various circumstances this will be optimal at the logarithmic level. More detailed remarks follow the statement of Theorem 3.

THEOREM 3. *Let μ be a centered Gaussian measure on a real separable Banach space B, $\phi(\varepsilon)$ be as in* (2.1), *and let K be the unit ball of the Hilbert space H_μ generating μ. If $\phi(\varepsilon) \approx f(\varepsilon)$ where $f(1/x)$ is regularly varying at infinity with a strictly positive exponent, then there exists a positive non-decreasing function g such that, as $x \to \infty$, $g(x) \to \infty$ and*

$$f(1/g(x)) \approx x. \tag{2.18}$$

Furthermore, if $X, X_1, X_2, ...$ are i.i.d. Gaussian random vectors with $\mu = \mathscr{L}(X)$, $E_n = \{X_{\eta(n)+1}, ..., X_n\}$ where $\eta(n) \leqslant n^{1/4}$ and $Ln \approx L\eta(n)$ as $n \to \infty$, $\delta \in (0, 1)$, and

$$\varepsilon_n = \sqrt{2}\,(Ln)^{-1/2}\,(g(Ln/\gamma))^{-1} \tag{2.19}$$

where γ is a positive constant, then

$$P((1-\delta)\, K \subseteq E_n/(2Ln)^{1/2} + B_{\varepsilon_n}(0) \textit{ eventually}) = 1 \tag{2.20}$$

for $\gamma > 0$ *sufficiently large, and*

$$P((1-\delta)K \subseteq E_n/(2Ln)^{1/2} + B_{\varepsilon_n}(0) \text{ eventually}) = 0 \tag{2.21}$$

for $\gamma > 0$ *sufficiently small.*

Remarks. (I) If $f(\varepsilon) = \varepsilon^{-\alpha}$ for $\varepsilon > 0$, $\alpha > 0$, then we can take $g(x) = x^{1/\alpha}$, and hence

$$\varepsilon_n = \sqrt{2}\,\gamma^{1/\alpha}(Ln)^{-(2+\alpha)/(2\alpha)}. \tag{2.22}$$

Thus the $n - \eta(n)$ randomly centered balls of radius ε_n cover $(1-\delta)K$ if $\gamma > 0$ is sufficiently large, but not when $\gamma > 0$ is small. Hence

$$H(\varepsilon, (1-\delta)K) \leqslant (\sqrt{2}\,\gamma^{1/\alpha})^{2\alpha/(2+\alpha)}\,\varepsilon^{-2\alpha/(2+\alpha)}, \tag{2.23}$$

and, if $H(\varepsilon, K)$ is sufficiently regular, this implies the same when $(1-\delta)K$ is replaced by K. Moreover, from Theorem 1 we actually know $H(\varepsilon, K)$ in this case, so the randomly placed centers are near best possible. Furthermore, (2.20) and (2.21) combine to demonstrate that this "random metric entropy," denoted by $G(\varepsilon, (1-\delta)K)$, satisfies

$$G(\varepsilon, (1-\delta)K) \approx \varepsilon^{-2\alpha/(2+\alpha)}.$$

Of course, if $\dim H_\mu = \infty$, then $\mu(H_\mu) = 0$, and hence with probability one the random set $E_n/(2Ln)^{1/2}$ is not a subset of K, or even H_μ. Thus the lower bound for $G(\varepsilon, (1-\delta)K)$ does not necessarily translate into a lower bound for $H(\varepsilon, (1-\delta)K)$, but it complements the upper bound nicely, and shows that in some sense the upper bound is not excessive.

(II) Theorem 3 is the sample analogue of Theorem 4 in [10]. In conjunction with Theorem 1 and Theorem 2, it furthers our understanding of the connection between the metric entropy of K and the small ball probabilities.

3. Proof of Theorems 1 and 2

First we present a useful lemma.

Lemma 1. *Let* μ *be a centered Gaussian measure on a real separable Banach space and let* $B_\varepsilon(h) = \{x : \|x-h\| < \varepsilon\}$. *If* $\lambda > 0$, $\varepsilon > 0$, *then*

$$H(2\varepsilon, \lambda K) \leqslant \lambda^2/2 - \log \mu(B_\varepsilon(0)) \tag{3.1}$$

and

$$H(\varepsilon, \lambda K) + \log \mu(B_{2\varepsilon}(0)) \geqslant \log \Phi(\lambda + \alpha_\varepsilon), \tag{3.2}$$

where

$$\Phi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t} e^{-u^2/2}\,du \qquad and \qquad \Phi(\alpha_\varepsilon)=\mu(B_\varepsilon(0)).$$

Proof. For $\lambda>0$, $\varepsilon>0$ define

$$M(\varepsilon,\lambda K)=\max\{n\geqslant 1:\exists h_1,\dots,h_n\in\lambda K,\ \|h_i-h_j\|\geqslant 2\varepsilon \text{ for all } i\neq j\} \tag{3.3}$$

and

$$N(\varepsilon,\lambda K)=\min\left\{k\geqslant 1:\exists h_1,\dots,h_k\in\lambda K,\ \bigcup_{j=1}^{k} B_\varepsilon(h_j)\supseteq\lambda K\right\}. \tag{3.4}$$

Then there exist finite sets $E(\varepsilon,\lambda K)$ and $F(\varepsilon,\lambda K)$, not necessarily unique, such that

$$\text{Card } E(\varepsilon,\lambda K)=M(\varepsilon,\lambda K) \tag{3.5}$$

and for $g, h\in E(\varepsilon,\lambda K)$, $g\neq h$, we have

$$\|h-g\|\geqslant 2\varepsilon, \tag{3.6}$$

and

$$\text{Card } F(\varepsilon,\lambda K)=N(\varepsilon,\lambda K) \tag{3.7}$$

with

$$\bigcup_{h\in F(\varepsilon,\lambda K)} B_\varepsilon(h)\supseteq\lambda K. \tag{3.8}$$

Using the Cameron–Martin formula for Gaussian measures and Jensen's inequality, we have for $h\in\lambda K$ that

$$\mu(B_\varepsilon(h))\geqslant\exp\{-\lambda^2/2\}\,\mu(B_\varepsilon(0)). \tag{3.9}$$

Hence (3.3), (3.5), and (3.6) imply

$$M(\varepsilon,\lambda K)\min_{h\in E(\varepsilon,\lambda K)}\mu(B_\varepsilon(h))\leqslant 1, \tag{3.10}$$

and, applying (3.9), it follows that

$$\log M(\varepsilon,\lambda K)-\lambda^2/2+\log\mu(B_\varepsilon(0))\leqslant 0. \tag{3.11}$$

Now $E(\varepsilon,\lambda K)$ maximal implies

$$\bigcup_{h\in E(\varepsilon,\lambda K)} B_{2\varepsilon}(h)\supseteq\lambda K, \tag{3.12}$$

and hence

$$H(2\varepsilon, \lambda K) \leqslant \log M(\varepsilon, \lambda K). \tag{3.13}$$

Combining (3.11) and (3.13) we see that (3.1) is proved.

Using (3.4), (3.7), and (3.8), we have

$$\bigcup_{h \in F(\varepsilon, \lambda K)} B_{2\varepsilon}(h) \supseteq \lambda K + B_\varepsilon(0), \tag{3.14}$$

and hence

$$N(\varepsilon, \lambda K) \cdot \max_{h \in F(\varepsilon, \lambda K)} \mu(B_{2\varepsilon}(h)) \geqslant \mu(\lambda K + B_\varepsilon(0)) \geqslant \Phi(\lambda + \alpha_\varepsilon) \tag{3.15}$$

by Borell's inequality [1], where $\Phi(\alpha_\varepsilon) = \mu(B_\varepsilon(0))$. Now $\mu(B_{2\varepsilon}(0)) \geqslant \mu(B_{2\varepsilon}(h))$ for all h (see, for example, [13, p. 332]), so (3.15) yields (3.2). Hence Lemma 1 is proved.

Remark. From (3.1) we see

$$\overline{\lim_{\lambda \to \infty}} \lambda^{-2} H(2\varepsilon/\lambda, K) \leqslant 1/2,$$

and hence, if $\delta = 2\varepsilon/\lambda$, then

$$\overline{\lim_{\delta \to 0}} \delta^2 H(\delta, K) \leqslant 2\varepsilon^2.$$

Since $\varepsilon > 0$ was arbitrary, this implies $H(\delta, K) = o(\delta^{-2})$ as was pointed out in [8] via a similar argument.

Proof of Theorem 1. From (2.2) and (3.1) we have for $\lambda > 0$ and $\varepsilon > 0$ sufficiently small that

$$H(2\varepsilon, \lambda K) \leqslant \lambda^2/2 + c_2 f(\varepsilon). \tag{3.16}$$

Since $H(\varepsilon, \lambda K) = H(\varepsilon\lambda^{-1}, K)$, we have by taking $\lambda = 4(c_2 f(\varepsilon))^{1/2}$ in (3.16) that

$$H(\varepsilon(4c_2 f(\varepsilon))^{-1/2}, K) \leqslant 9c_2 f(\varepsilon). \tag{3.17}$$

Letting $\delta = j(\varepsilon)$ as in (2.3) and g be as in (2.5), then (3.17) and $f(1/x)$ regularly varying at infinity imply

$$H(\delta, K) \leqslant c_3 f(g(\delta)) \qquad \text{as} \quad \delta \to 0 \tag{3.18}$$

for some finite positive constant c_3.

From (3.2), we have for $\lambda > 0$ and $\varepsilon > 0$ that

$$H(\varepsilon, \lambda K) \geqslant \phi(2\varepsilon) + \log \Phi(\lambda + \alpha_\varepsilon). \tag{3.19}$$

Taking $\lambda = -\alpha_\varepsilon$, and recalling $\Phi(\alpha_\varepsilon) = \mu(B_\varepsilon(0))$, we have that $\alpha_\varepsilon \to -\infty$ as $\varepsilon \to 0$ with

$$-\alpha_\varepsilon^2/2 \sim \log \mu(B_\varepsilon(0)) = -\phi(\varepsilon). \tag{3.20}$$

Hence $-\alpha_\varepsilon \sim (2\phi(\varepsilon))^{1/2}$ and (3.19) implies

$$H(\varepsilon(-\alpha_\varepsilon)^{-1}, K) \geqslant \phi(2\varepsilon) + \log(1/2). \tag{3.21}$$

Now $H(\delta, K)$ increases as δ decreases, so for $\varepsilon > 0$ sufficiently small, (2.2) and (3.21) imply that

$$H(\varepsilon(4c_2 f(\varepsilon))^{-1/2}, K) \geqslant c_1 f(2\varepsilon)/2. \tag{3.22}$$

Letting $\delta = j(\varepsilon)$ and g be as in (2.5), then (3.22) and $f(1/x)$ regularly varying at infinity implies

$$H(\delta, K) \geqslant c_4 f(g(\delta)) \qquad \text{as} \quad \delta \to 0 \tag{3.23}$$

for some strictly positive constant c_4. Hence the theorem is proved.

Proof of Theorem 2. From (2.13) we have strictly positive finite constants d_1, d_2 such that

$$d_1 g(\varepsilon) \leqslant H(\varepsilon, K) \leqslant d_2 g(\varepsilon) \tag{3.24}$$

as $\varepsilon \to 0$. Hence by (3.2) for $\lambda > 0$ and $\varepsilon > 0$,

$$d_2 g(\varepsilon/\lambda) \geqslant \phi(2\varepsilon) + \log \Phi(\lambda + \alpha_\varepsilon). \tag{3.25}$$

Setting $\lambda = -\alpha_\varepsilon$, we get for all $\varepsilon > 0$ sufficiently small that

$$2d_2 g(\varepsilon(-\alpha_\varepsilon)^{-1}) \geqslant \phi(2\varepsilon). \tag{3.26}$$

Now $\Phi(\alpha_\varepsilon) = \mu(B_\varepsilon(0))$, so $-\alpha_\varepsilon \sim (2\phi(\varepsilon))^{1/2}$, and hence g regularly varying with $g(\varepsilon) \to \infty$ as $\varepsilon \to 0$ implies

$$d_3 g(\varepsilon/(\phi(\varepsilon))^{1/2}) \geqslant \phi(2\varepsilon) \tag{3.27}$$

for some positive finite constant d_3.

Now for $\lambda > 0$ and $\varepsilon > 0$, (3.1) and (3.24) imply

$$d_1 g(2\varepsilon/\lambda) \leqslant \lambda^2/2 + \phi(\varepsilon). \tag{3.28}$$

Taking $\lambda = 2(\phi(\varepsilon))^{1/2}$, (3.28) implies

$$d_1 \, g(\varepsilon/(\phi(\varepsilon))^{1/2}) \leqslant 3\phi(\varepsilon). \tag{3.29}$$

Combining (3.27) and (3.29) now implies

$$\phi(2\varepsilon) \preccurlyeq g(\varepsilon/(\phi(\varepsilon))^{1/2}) \preccurlyeq \phi(\varepsilon) \tag{3.30}$$

as $\varepsilon \to 0$. Thus the first part of the theorem is proved since $\phi(\varepsilon) \preccurlyeq \phi(2\varepsilon)$ and (3.30) together imply (2.15).

If $0 < \beta < 2$ and

$$g(\varepsilon) = \varepsilon^{-\beta} J(\varepsilon^{-1}), \tag{3.31}$$

where $J(\cdot)$ is slowly varying, monotonic, and such $J(x) \approx J(x^\rho)$ for each $\rho > 0$ as $x \to \infty$, then (2.15) implies

$$\phi(\varepsilon)^{(2-\beta)/2} \approx \varepsilon^{-\beta} J((\phi(\varepsilon))^{1/2}/\varepsilon). \tag{3.32}$$

Since $J(x)$ is slowly varying at infinity, we have for each $\delta > 0$ and all x sufficiently large that

$$x^{-\delta} \leqslant J(x) \leqslant x^{\delta}. \tag{3.33}$$

Combining (3.32), (3.33), and that J is monotonic with $J(x) \approx J(x^\rho)$ as $x \to \infty$, we have

$$J((\phi(\varepsilon))^{1/2}/\varepsilon) \approx J(1/\varepsilon). \tag{3.34}$$

Hence (2.16) holds, and the theorem is proved.

4. Proof of Theorem 3

Since $f(1/x)$ is regularly varying with strictly positive index α, then [21, pp. 21, 25] implies that a strictly positive non-decreasing function g satisfying (2.18) exists such that g is also regularly varying with index $1/\alpha$. Hence let ε_n be given as in (2.19). Fix $\delta \in (0, 1)$ and let K_n be a finite subset of $(1-\delta)K$ such that:

(i) Balls centered at points of K_n of B-norm radius $\varepsilon_n/4$ are disjoint.

(ii) K_n is maximal, i.e., if we add a point of $(1-\delta)K$ to K_n we get overlap among the balls of B-norm radius $\varepsilon_n/4$ centered at this larger set.

Thus

$$P(K_n \nsubseteq E_n/(2Ln)^{1/2} + B_{\varepsilon_n/2}(0))$$

$$\leqslant \sum_{f \in K_n} P(\|X_j - (2Ln)^{1/2} f\| > (2Ln)^{1/2} \varepsilon_n/2, \eta(n) \leqslant j \leqslant n)$$

$$\leqslant \sum_{f \in K_n} (1 - \exp\{-\|f\|_\mu^2 Ln\} \mu((2Ln)^{1/2} B_{\varepsilon_n/2}(0))^{n-\eta(n)}$$

$$\leqslant \text{card}(K_n)(1 - \exp\{-(1-\delta)^2 Ln - \phi((2Ln)^{1/2} \varepsilon_n/2)\})^{n/2}. \quad (4.1)$$

Now $\phi(\varepsilon) \approx f(\varepsilon)$ so there is a strictly positive constant a such that $\phi(\varepsilon) \geqslant af(\varepsilon)$ and hence (2.19) and (4.1) combine to imply

$$P(K_n \nsubseteq E_n/(2Ln)^{1/2} + B_{\varepsilon_n/2}(0))$$

$$\leqslant \text{card}(K_n)(1 - \exp\{-(1-\delta)^2 Ln - af(1/g(Ln/\gamma))\})^{n/2}$$

$$\leqslant \text{card}(K_n)(1 - \exp\{-(1-\delta)^2 Ln - \hat{a}Ln/\gamma\})^{n/2}, \quad (4.2)$$

where the strictly positive constant $\hat{a}$ follows from (2.18).

Now $\text{card}(K_n) \leqslant \exp\{+\varepsilon_n^{-2}\}$ by [8], or the remark following (3.15), and since $1 - x \leqslant e^{-x}$ we have from (4.2) that

$$P(K_n \nsubseteq E_n/(2Ln)^{1/2} + B_{\varepsilon_n/2}(0))$$

$$\leqslant \exp\{+\varepsilon_n^{-2} - n/2 \exp\{-(1-\delta)^2 Ln - \hat{a}Ln/\gamma\}$$

$$= \exp\{+g^2(Ln/\gamma) Ln/2 - 2^{-1} n^{1-(1-\delta)^2 - \hat{a}/\gamma}\}. \quad (4.3)$$

Since g is regularly varying at infinity with positive exponent, it follows that if $\gamma > 0$ is taken large enough so that $(1-\delta)^2 + \hat{a}/\gamma < 1$ we have

$$\sum_{n \geqslant 1} P(K_n \nsubseteq E_n/(2Ln)^{1/2} + B_{\varepsilon_n/2}(0)) < \infty. \quad (4.4)$$

Hence by (i) and (ii), (4.4) implies (2.20) for γ sufficiently large.

To prove (2.21) we show that

$$P(\{0\} \subseteq E_n/(2Ln)^{1/2} + B_{\varepsilon_n}(0) \text{ eventually}) = 0 \quad (4.5)$$

if $\gamma > 0$ is sufficiently small.

First we show (4.5) holds if we have for $\gamma > 0$ sufficiently small that

$$P(\varliminf_{n \to \infty} \varepsilon_n^{-1} \|X_n/(2Ln)^{1/2}\| \geqslant \sqrt{2}) = 1. \quad (4.6)$$

That is, if (4.6) holds, then for such $\gamma > 0$ with probability one

$$\begin{aligned}
&\varliminf_{n\to\infty} \inf_{\eta(n)\leqslant k\leqslant n} \varepsilon_n^{-1} \|X_k/(2Ln)^{1/2}\| \\
&= \varliminf_{n\to\infty} \inf_{\eta(n)\leqslant k\leqslant n} \varepsilon_n^{-1} \|X_k/(2Lk)^{1/2}\| \, (Lk/Ln)^{1/2} \\
&\geqslant \varliminf_{n\to\infty} \inf_{\eta(n)\leqslant k\leqslant n} \varepsilon_k^{-1} \|X_k/(2Lk)^{1/2}\| \, (L\eta(n)/Ln)^{1/2} \\
&\geqslant \varliminf_{k\to\infty} \varepsilon_k^{-1} \|X_k/(2Lk)^{1/2}\| \cdot \varliminf_{n} (L\eta(n)/Ln)^{1/2} \\
&\geqslant C\sqrt{2} > 0,
\end{aligned} \tag{4.7}$$

since $L\eta(n) \approx Ln$ and for any $\gamma > 0$ fixed ε_k^{-1} is eventually non-decreasing. Of course, (4.7) implies (4.5), so it remains to verify (4.6).

Let $A_n = \{\varepsilon_n^{-1} \|X_n/(2Ln)^{1/2}\| \leqslant \sqrt{2}\}$. Then

$$\begin{aligned}
P(A_n) &= \mu(x\colon \|x\| \leqslant \sqrt{2}\, \varepsilon_n (2Ln)^{1/2}) \\
&= \mu(x\colon \|x\| \leqslant 2\sqrt{2}\, (g(Ln/\gamma))^{-1}) \\
&\leqslant \exp\{-b \cdot Ln/\gamma\}
\end{aligned} \tag{4.8}$$

for some $b > 0$ uniformly in $\gamma \in (0, 1]$. The inequality in (4.8) follows since $f(1/x)$ is regularly varying at infinity and (2.18) holds. Thus $\sum_{n\geqslant 1} P(A_n) < \infty$ for $\gamma > 0$ sufficiently small, and the Borel–Cantelli lemma thus implies (4.6). Hence (2.21) holds and the theorem is proved.

5. Applications

The results of this section are of two types. The first show how metric entropy estimates can provide small ball probabilities in the Hilbert space setting, and then we show how small ball probabilities can sharpen and improve some metric entropy estimates.

A. *Small Ball Probabilities in Hilbert Space*

Here we assume $\mu = \mathscr{L}(X)$ is a centered Gaussian measures on a real separable Hilbert space H. To provide estimates for the small ball probabilities of μ (or X) we apply Theorem 2. However, to do this effectively, we need to know when $\phi(\varepsilon) = -\log P(\|X\| \leqslant \varepsilon)$ satisfies $\phi(\varepsilon) \preccurlyeq \phi(2\varepsilon)$ as $\varepsilon \to 0$. Then Theorem 2 will be quite useful.

Proposition 1. *Let $\mu = \mathscr{L}(X)$ where $X = \sum_{k\geqslant 1} a_k^{1/2} \xi_k e_k$ is a centered Gaussian vector with values in a real separable Hilbert space H, $a_k > 0$ is* *non-increasing,* $\{\xi_k : k \geq 1\}$ *is a sequence of independent* $N(0, 1)$ *random variables, and* $\{e_k : k \geq 1\}$ *is an orthonormal sequence in* H. *Let* $\phi(\varepsilon)$ *be as in* (2.1), $b_k = a_k^{-1}$, *and*

$$f(x) = \sum_{k \geq 1} \frac{1}{b_k + x} \qquad x > 0. \tag{5.1}$$

Thus $f(x)$ *is a finite, decreasing function, and if there exists a* $\theta > 0$ *such that for* x *large*

$$f(\theta x) \geq 4f(x), \tag{5.2}$$

then

$$\overline{\lim_{\varepsilon \to 0}}\ \phi(\varepsilon)/\phi(2\varepsilon) \leq 1 + [\theta^{-1}] < \infty. \tag{5.3}$$

Remark. If $\mu = \mathscr{L}(X)$ is a centered Gaussian measure on a Hilbert space H, then it is well known that X can always be written in the form indicated in Proposition 1. Hence the only assumption of substance here is (5.2).

Proof. Since $b_k = 1/a_k$ with $a_k > 0$, $\sum_{k \geq 1} a_k < \infty$, we see that $f(x) = \sum_{k \geq 1} a_k/(1 + a_k x)$. Hence it is clear that $f(x)$ is a finite, decreasing function for $x > 0$.

Now assume (5.2) holds, say for $x > x_0$. From Theorem 4 in [17] we have as $\varepsilon \to 0$ that

$$\begin{aligned}\phi(\varepsilon) &\sim \int_0^{\gamma_1} \sum_{k \geq 1} \frac{a_k}{1 + 2a_k x}\, dx - \gamma_1 \sum_{k \geq 1} \frac{a_k}{1 + 2a_k \gamma_1} \\ &= \int_0^{\gamma_1} \sum_{k \geq 1} \frac{1}{b_k + 2x}\, dx - \gamma_1 \sum_{k \geq 1} \frac{1}{b_k + 2\gamma_1} \\ &= \int_0^{\gamma_1} \sum_{k \geq 1} \frac{2x}{(b_k + 2x)^2}\, dx,\end{aligned} \tag{5.4}$$

and

$$\begin{aligned}\phi(2\varepsilon) &= -\log P\left(\sum_{k \geq 1} \frac{a_k}{4} \xi_k^2 \leq \varepsilon^2\right) \\ &\sim \int_0^{\gamma_2} \sum_{k \geq 1} \frac{2x}{(4b_k + 2x)^2}\, dx \\ &= \int_0^{\gamma_2/4} \sum_{k \geq 1} \frac{2x}{(b_k + 2x)^2}\, dx,\end{aligned} \tag{5.5}$$

where

$$\varepsilon^2 = \sum_{k \geq 1} \frac{1}{b_n + 2\gamma_1} = f(2\gamma_1), \tag{5.6}$$

$$\varepsilon^2 = \sum_{k \geq 1} \frac{1}{4b_n + 2\gamma_2} = \frac{1}{4} f\left(\frac{\gamma_2}{2}\right). \tag{5.7}$$

Now for $\varepsilon > 0$ sufficiently small we have $\gamma_1 \geq x_0$ and hence (5.2) implies

$$f(2\theta\gamma_1) \geq 4f(2\gamma_1) = f(\gamma_2/2), \tag{5.8}$$

where the equality follows from (5.6) and (5.7). Since $f(x)$ is a decreasing function this implies that

$$\gamma_2 \geq 4\theta\gamma_1. \tag{5.9}$$

Hence

$$\begin{aligned} &\int_0^{\gamma_1} \sum_{k \geq 1} \frac{2x}{(b_k + 2x)^2}\, dx \\ &\leq \int_0^{\gamma_2/4\theta} \sum_{k \geq 1} \frac{2x}{(b_k + 2x)^2}\, dx \\ &\leq \sum_{i=0}^{[\theta^{-1}]} \int_{i\gamma_2/4}^{(i+1)\gamma_2/4} \sum_{k \geq 1} \frac{2x}{(b_k + 2x)^2}\, dx \\ &\leq ([\theta^{-1}] + 1) \int_0^{\gamma_2/4} \sum_{k \geq 1} \frac{2x}{(b_k + 2x)^2}\, dx \end{aligned} \tag{5.10}$$

since $\sum_{k \geq 1} 2x/(b_k + 2x)^2$ is a decreasing function. Combining (5.4), (5.6), and (5.10) we have (5.3). Hence the proposition is proved.

Our next proposition provides conditions to insure that (5.2) holds.

PROPOSITION 2. *Let $\mu = \mathscr{L}(X)$ where $X = \sum_{k \geq 1} a_k^{1/2} \xi_k e_k$ is a centered Gaussian vector as in Proposition 1. Let $b_k = a_k^{-1} = \lambda(k)$ where $\lambda(t)$ is non-decreasing and of the form*

$$\lambda(t) = t^\alpha J(t), \qquad t > 0, \tag{5.11}$$

where $\alpha > 1$ and $J(\cdot)$ is slowly varying at infinity. Then (5.2) holds for some $\theta > 0$.

Proof. Now for x large, say $x \geqslant x_1$, we have

$$f(x) = \sum_{k \geqslant 1} \frac{1}{b_k + x} \leqslant \frac{1}{b_1 + x} + \int_1^{\infty} \frac{1}{\lambda(t) + x}\, dt$$

$$\leqslant (1 + b_2 b_1^{-1}) \int_1^{\infty} \frac{1}{\lambda(t) + x}\, dt. \tag{5.12}$$

Now set $\theta = 2h^{\alpha}$ where $h^{1-\alpha} = 16(1 + b_2 b_1^{-1})$. Since $J(\cdot)$ is slowly varying there exists a t_0 such that $t \geqslant t_0$ implies

$$J(ht) \leqslant 2J(t). \tag{5.13}$$

Also, since $xf(x) \to \infty$ as $x \to \infty$, we have for x large, say $x \geqslant x_2$,

$$(1 + b_2 b_1^{-1}) \cdot \max(h^{-1}, t_0) \cdot x^{-1} \leqslant f(x)/2. \tag{5.14}$$

Hence for $x \geqslant \max(x_1, x_2)$ combining (5.12), (5.13), and (5.14) yields

$$\begin{aligned}
f(\theta x) &= \sum_{k \geqslant 1} \frac{1}{b_k + \theta x} \\
&\geqslant \int_1^{\infty} \frac{1}{\lambda(t) + \theta x}\, dt \\
&= \int_{h^{-1}}^{\infty} \frac{h}{\lambda(ht) + \theta x}\, dt \\
&\geqslant \int_{\max(h^{-1}, t_0)}^{\infty} \frac{h}{h^{\alpha} t^{\alpha} J(ht) + \theta x}\, dt \\
&\geqslant \int_{\max(h^{-1}, t_0)}^{\infty} \frac{h}{2h^{\alpha} t^{\alpha} J(t) + \theta x}\, dt \\
&= \frac{1}{2h^{\alpha - 1}} \int_{\max(h^{-1}, t_0)}^{\infty} \frac{1}{t^{\alpha} J(t) + x}\, dt \\
&= 8(1 + b_2 b_1^{-1}) \left(\int_1^{\infty} \frac{1}{t^{\alpha} J(t) + x}\, dt - \int_1^{\max(h^{-1}, t_0)} \frac{1}{t^{\alpha} J(t) + x}\, dt \right) \\
&\geqslant 8(1 + b_2 b_1^{-1}) \left(\frac{f(x)}{1 + b_2 b_1^{-1}} - \frac{\max(h^{-1}, t_0)}{x} \right) \\
&\geqslant 8(f(x) - f(x)/2) \\
&= 4f(x).
\end{aligned} \tag{5.15}$$

Thus Proposition 2 is proved.

We now are in position to prove a useful theorem for Hilbert space valued Gaussian random vectors.

THEOREM 4. *Let* $\mu = \mathscr{L}(X)$ *where* $X = \sum_{k \geqslant 1} a_k^{1/2} \xi_k e_k$ *is a Hilbert space valued centered Gaussian vector as in Proposition* 1, *and assume* $\lambda(t) = t^\alpha J(t)$, *where* $\alpha > 1$ *and* $J(\cdot)$ *is slowly varying at infinity, monotonic, and such that* $J(x) \approx J(x^\rho)$ *for each* $\rho > 0$ *as* $x \to \infty$. *If* $a_k^{-1} = \lambda(k)$, *then*

$$\log P(\|X\| \leqslant \varepsilon) \approx -\varepsilon^{-2/(\alpha-1)} (J(1/\varepsilon))^{-1/(\alpha-1)} \qquad as \quad \varepsilon \to 0. \tag{5.16}$$

Remarks. If $a_k = k^{-\alpha} (\log k)^\beta$ with $\alpha > 1$ and β a real number, then $\lambda(t) = t^\alpha (\log t)^{-\beta}$, and Theorem 4 implies

$$\log P(\|X\| \leqslant \varepsilon) \approx -\varepsilon^{-2/(\alpha-1)} (\log(1/\varepsilon))^{\beta/(\alpha-1)}, \qquad \text{as} \quad \varepsilon \to 0. \tag{5.17}$$

The small ball probabilities calculated previously for Gaussian measures on H did not usually involve the logarithmic factors as in (5.17), as they make the estimates which were obtained via a detailed analysis of the Laplace transform of $\|X\|$ rather delicate (see [17] for details and further references). The method applied here is much simpler to use, and for the most part produces the best results.

Proof of Theorem 4. From Proposition 1 and 2 we have $\phi(\varepsilon) = -\log P(\|X\| \leqslant \varepsilon)$ satisfying $\phi(\varepsilon) \preccurlyeq \phi(2\varepsilon)$ as $\varepsilon \to 0$. Hence we can apply Theorem 2 once we show

$$H(\varepsilon, K) \approx g(\varepsilon) \qquad \text{as} \quad \varepsilon \to 0, \tag{5.18}$$

where $g(1/x)$ is slowly varying at infinity. To do this let

$$m(t) = \begin{cases} \sup_{k \geqslant 1} \{k : a_k^{1/2} \geqslant t^{-1}\} & t \geqslant a_1^{-1/2} \\ 0 & t < a_1^{-1/2} \end{cases}$$

and define for $t \geqslant 0$

$$I(t) = \int_0^t x^{-1} m(x)\, dx.$$

Then by Theorem 3 and Corollary 2 in [19], we have

$$H(\varepsilon, K) \approx I(1/\varepsilon) \qquad \text{as} \quad \varepsilon \to 0, \tag{5.19}$$

where $\|\cdot\|$ is the norm on H. Since $a_k^{-1} = \lambda(k)$ it is easy to see that as $t \to \infty$

$$m(t) \approx t^{2/\alpha} J(t)^{-1/\alpha}, \tag{5.20}$$

and hence as $t \to \infty$

$$I(t) \approx t^{2/\alpha} J(t)^{-1/\alpha}. \tag{5.21}$$

Hence Theorem 2 above applies and

$$\phi(\varepsilon) \approx g(\varepsilon/(\phi(\varepsilon))^{1/2}) \qquad \text{as} \quad \varepsilon \to 0, \tag{5.22}$$

where $g(\varepsilon) = I(1/\varepsilon)$ by (5.19). In view of the assumptions on $J(\cdot)$, (5.16) now holds.

B. *Small Ball Probabilities Yield Metric Entropy Results*

Sometimes $\phi(\varepsilon)$ is known very precisely, and then the approach of Theorem 1 can often yield correspondly precise estimates of $H(\varepsilon, K)$ which are better than those in the literature. For examples of this type we consider the following.

B-1. *l^p Entropy Results*

If $X = \sum_{k \geq 1} \lambda_k \xi_k e_k$ where $\{e_k : k \geq 1\}$ is the canonical basis in the l^p spaces, $1 \leq p < \infty$, then $P(X \in l^p) = 1$ iff $\sum_{k \geq 1} |\lambda_k|^p < \infty$ and

$$K = \left\{ x \in l^p : \sum_{k \geq 1} x_k^2/\lambda_k^2 \leq 1 \right\}.$$

If $P(X \in l^p) = 1$, then K is compact in l^p. For $p \neq 2$, the metric entropy of K in the l^p-norm is not so trivial to compute. The basic reason for this is that the volumes of finite-dimensional projections of K do not compare well with the volumes of the same finite-dimensional projection of the unit ball of l^p when $p \neq 2$. However, when $1 \leq p < \infty$ and $\lambda_k = k^{-\alpha/p}$ for $\alpha > 1$ then [18] yields

$$\log P\left(\left(\sum_{k \geq 1} |\lambda_k|^p \, |\xi_k|^p \right)^{1/p} \leq \varepsilon \right) \approx -\varepsilon^{-p/(\alpha - 1)}.$$

Hence by Theorem 1 the corresponding ellipsoids

$$K = \left\{ x \in l^p : \sum_{k \geq 1} k^{2\alpha/p} x_k^2 \leq 1 \right\}$$

have metric entropy in the l^p norm

$$H(\varepsilon, K) \approx \varepsilon^{-2p/(2\alpha + p - 2)}.$$

B-2. *Entropy Results Connected with Brownian Motion*

Another interesting class of examples arises when μ denotes Wiener measure on $C[0, 1]$. In this case, K is given in (1.1) and is a compact subset of $C[0, 1]$ for any of the norms

$$\|f\|_p = \begin{cases} (\int_0^1 |f(s)|^p \, ds)^{1/p} & 1 \leqslant p < \infty \\ \sup_{0 \leqslant s \leqslant 1} |f(s)| & p = \infty. \end{cases}$$

For Wiener measure and $1 \leqslant p \leqslant \infty$, it is known that

$$\log \mu(\|x\|_p \leqslant \varepsilon) \approx -\varepsilon^{-2} \tag{5.23}$$

and it is also known that

$$H(\varepsilon, K, \|\cdot\|_p) \approx \varepsilon^{-1}. \tag{5.24}$$

In fact, more than (5.23) is known for the small ball probabilities, but for $1 \leqslant p < 2$ these results have only been obtained recently [2] and are quite delicate. On the other side, the metric entropy results in (5.24) were obtained in [3] for $p = \infty$, and in [5] for $p = 1$. The remaining cases are then obvious. Of course, in view of our results, (5.23) and (5.24) are in complete duality. Furthermore, since in this case the logarithm of the small ball probabilities are known asymptotically, especially for $p = 2$ and $p = \infty$, we then have correspondly better estimates for (5.24). For example, we have the following proposition.

PROPOSITION 3. *If K is as in* (1.1), *then for each $\delta > 0$ as $\varepsilon \to 0$*

$$(1-\delta)(2-\sqrt{3})/4 \leqslant \varepsilon \cdot H(\varepsilon, K, \|\cdot\|_2) \leqslant 1 + \delta \tag{5.25}$$

and

$$(1-\delta)(2-\sqrt{3})\, \pi/4 \leqslant \varepsilon \cdot H(\varepsilon, K, \|\cdot\|_\infty) \leqslant \pi(1+\delta). \tag{5.26}$$

Remarks. (I) For $p = 2$, (5.25) is more precise than what is given in Theorem XVI of [14], and for $p = \infty$, there are no constant bounds in [3].

(II) The lower estimates for metric entropy are frequently obtained by a volume comparison; i.e., for suitable finite dimensional projections, the total volume of the covering balls is less than the volume of the set being covered. As a result, this method is often too crude to provide reasonable constants.

(III) If $AC[0, 1]$ denotes the absolutely continuous real-valued functions on $[0, 1]$, then the Sobolev space

$$W_2^{(1)} = \{f \in AC[0, 1] : f' \in L^2[0, 1]\} \tag{5.27}$$

is a Banach space in the norm

$$\|f\|_{W_2^{(1)}} = \|f\|_2 + \|f'\|_2, \tag{5.28}$$

where $\|\cdot\|_2$ is the usual norm on $L^2[0, 1]$. Furthermore, if $\tilde{K}$ is the unit ball in this space, then it is easy to see that $H(\varepsilon, \tilde{K}, \|\cdot\|_\infty) \sim H(\varepsilon, K, \|\cdot\|_\infty)$ as $\varepsilon \to 0$, and hence (5.25) and (5.26) imply similar results about $H(\varepsilon, \tilde{K})$.

Proof of Proposition 3. First we recall the classical facts that if μ is a Wiener measure on $C[0, 1]$, then as $\varepsilon \to 0$

$$\log \mu(x: \|x\|_2 \leqslant \varepsilon) \sim -(1/8) \cdot \varepsilon^{-2} \tag{5.29}$$

and

$$\log \mu(x: \|x\|_\infty \leqslant \varepsilon) \sim -(\pi^2/8) \cdot \varepsilon^{-2}. \tag{5.30}$$

The result in (5.30) can be found in [4], and (5.29), as well as further references, can be found in [6, p. 43].

Thus (5.25) and (5.26) will follow from (5.29) and (5.30) if we show for a generic norm $\|\cdot\|$ on $C[0, 1]$ that

$$\log \mu(x: \|x\| \leqslant \varepsilon) \sim -C \cdot \varepsilon^{-2} \tag{5.31}$$

as $\varepsilon \to 0$ implies that for all $\delta > 0$ and $\varepsilon \to 0$ that

$$(1-\delta)(\sqrt{2} - \sqrt{3/2})\sqrt{C} \leqslant \varepsilon \cdot H(\varepsilon, K, \|\cdot\|) \leqslant 2\sqrt{2C}\,(1+\delta). \tag{5.32}$$

Hence assume (5.31) and then by (3.1) we have with $\lambda = \sqrt{2C}\,\varepsilon^{-1}$ that as $\varepsilon \to 0$

$$H(2\varepsilon, \sqrt{2C}\,\varepsilon^{-1}K, \|\cdot\|) \leqslant 2C(1+\delta)\,\varepsilon^{-2} \tag{5.33}$$

for every $\delta > 0$. Thus we get the right side of (5.32) by rescaling.

Now taking $\lambda = \sqrt{3/2}\sqrt{C}\,\varepsilon^{-1}$ in (3.2) we have for each $\delta > 0$ that as $\varepsilon \to 0$

$$H(\varepsilon, \sqrt{3/2}\sqrt{C}\,\varepsilon^{-1}K, \|\cdot\|) \geqslant (1-\delta)\,C(2\varepsilon)^{-2} + \log \Phi(\lambda + \alpha_\varepsilon). \tag{5.34}$$

From (5.31) and (3.2) we have $\alpha_\varepsilon \to -\infty$ and $\alpha_\varepsilon \sim -\sqrt{2C}\,\varepsilon^{-1}$. Hence for $\lambda = \sqrt{3/2}\sqrt{C}\,\varepsilon^{-1}$, we have $\lambda + \alpha_\varepsilon \to -\infty$ as $\varepsilon \to 0$ and thus

$$\log \Phi(\lambda + \alpha_\varepsilon) \sim -(\lambda + \alpha_\varepsilon)^2/2 \sim -(1 - \sqrt{3}/2)^2\,C\varepsilon^{-2}. \tag{5.35}$$

Hence (5.34) and (5.35) combine to imply that as $\varepsilon \to 0$

$$H(\varepsilon, \sqrt{3/2}\sqrt{C}\,\varepsilon^{-1}K, \|\cdot\|) \geqslant (1-2\delta)\,C(2\varepsilon)^{-2}\,(1 - (2-\sqrt{3})^2).$$

This implies the left hand side of (5.32) by rescaling. Thus (5.32) holds, and setting $C = -\pi^2/8$ in (5.32) yields (5.26). Also, if $C = -1/8$ in (5.32), we have (5.25). Hence the proposition is proved.

B-3. *Entropy Results Conneced to the Brownian Sheet*

If μ is the centered Gaussian measure on $C(Q)$, $Q = [0, 1] \times [0, 1]$ induced by the standard Brownian sheet, then the unit ball of the generating Hilbert space for μ is

$$K = \left\{ f \in C(Q) : f(s, t) = \int_0^s \int_0^t g(u, v)\, du\, dv, \iint_Q g^2(u, v)\, du\, dv \leqslant 1 \right\} \tag{5.36}$$

If $\|\cdot\|_p$ is the usual L^p norm on $C(Q)$, $1 \leqslant p \leqslant \infty$, then [17] an [7] establish that as $\varepsilon \to 0$

$$\log \mu(f \in C(Q) : \|f\|_2 \leqslant \varepsilon) \sim -(8\pi)^{-2} \cdot \varepsilon^{-2} (\log 1/\varepsilon^2)^2. \tag{5.37}$$

Hence by Theorem 1 we have as $\varepsilon \to 0$

$$H(\varepsilon, K, \|\cdot\|_2) \approx \varepsilon^{-1} \log 1/\varepsilon. \tag{5.38}$$

This coincides with the metric entropy result obtained in [22]. Furthermore, since the L^p-norms are increasing in p, we have

$$H(\varepsilon, K, \|\cdot\|_p) \preccurlyeq \varepsilon^{-1} \log 1/\varepsilon \tag{5.39}$$

for $1 \leqslant p \leqslant 2$, and also

$$\log \mu(f \in C(Q) : \|f\|_p \leqslant \varepsilon) \succcurlyeq -\varepsilon^{-2} (\log 1/\varepsilon)^2. \tag{5.40}$$

We also know from [24, Theorem 1.4] that as $\varepsilon \to 0$

$$H(\varepsilon, K, \|\cdot\|_1) \succcurlyeq \varepsilon^{-1} \log 1/\varepsilon. \tag{5.41}$$

To see why (5.41) follows from [24] we observe that (5.41) does hold for the ball $W^{(1,1)}_{2,(1,1)}$ defined in [24]. Now one can show that $f \in W^{(1,1)}_{2,(1,1)}$ iff

$$f(s, t) = \int_0^s \int_0^t g(u, v)\, du\, dv - s \int_0^t \int_0^1 g(\lambda, v)\, d\lambda\, dv$$

$$- t \int_0^s \int_0^1 g(u, \lambda)\, d\lambda\, du + st \int_0^1 \int_0^1 g(u, v)\, du\, dv + \text{constant},$$

where $\int_0^1 \int_0^1 g^2(u, v)\, du\, dv \leqslant 1$. Now $\int_0^1 (\int_0^1 g(u, \lambda)\, d\lambda)^2\, dv \leqslant 1$ by Jensen's inequality, and the metric entropy of (1.1) in the sup-norm is of order $1/\varepsilon$, hence the metric entropy of K in the $\|\cdot\|_1$-norm is of the same order as that of $W^{(1,1)}_{2,(1,1)}$. Thus (5.41) holds.

Since (5.41) holds, by setting $\lambda = 2(\phi(\varepsilon))^{1/2}$ we have from (3.1) that as $\varepsilon \to 0$

$$\varepsilon^{-1} \log(1/\varepsilon) \preccurlyeq \phi(\varepsilon)^{1/2}, \tag{5.42}$$

where $\phi(\varepsilon) = -\log \mu(f \in C(Q) : \|f\|_1 \leqslant \varepsilon)$. Hence as $\varepsilon \to 0$

$$\phi(\varepsilon) \succcurlyeq \varepsilon^{-2}(\log 1/\varepsilon)^2, \tag{5.43}$$

and by combining (5.40) and (5.43) we have

$$\phi(\varepsilon) \approx \varepsilon^{-2}(\log 1/\varepsilon)^2 \tag{5.44}$$

for the norm $\|\cdot\|_1$ as well. Of course by Theorem 1, (5.44) now easily implies that the following has been proved.

PROPOSITION 4. *Let μ be the centered Gaussian measure on $C(Q)$ induced by the standard Brownian sheet and let K be given as in* (5.36). *If $\|\cdot\|_p$ denotes the usual L^p norm on $C(Q)$ and*

$$\phi(\varepsilon) = -\log \mu(f \in C(Q) : \|f\|_p \leqslant \varepsilon), \tag{5.45}$$

then for $1 \leqslant p \leqslant 2$, as $\varepsilon \to 0$

$$H(\varepsilon, K, \|\cdot\|_p) \approx \varepsilon^{-1} \log 1/\varepsilon, \tag{5.46}$$

and

$$\phi(\varepsilon) \approx \varepsilon^{-2}(\log 1/\varepsilon)^2. \tag{5.47}$$

Remarks. (I) A recent result of Talagrand [23] yields that as $\varepsilon \to 0$

$$\phi(\varepsilon) \approx \varepsilon^{-2}(\log 1/\varepsilon)^3 \tag{5.48}$$

when $p = \infty$ in (5.45). Combined with Theorem 1 this implies

$$H(\varepsilon, K, \|\cdot\|_\infty) \approx \varepsilon^{-1}(\log 1/\varepsilon)^{3/2}. \tag{5.49}$$

Hence in the two-variable setting, the metric entropy of K changes from (5.47) when $1 \leqslant p \leqslant 2$ to (5.49) when $p = \infty$. This contrasts sharply with the one-variable results of (5.24); also see (5.23). Of course, if K is given by (5.36), then by the argument used to verify Proposition 4, and (5.24), we have for $1 \leqslant p \leqslant \infty$ that as $\varepsilon \to 0$

$$H(\varepsilon, W^{(1,1)}_{2,(1,1)}, \|\cdot\|_p) \approx H(\varepsilon, K, \|\cdot\|_p).$$

Hence (5.49) yields

$$H(\varepsilon, W^{(1,1)}_{2,(1,1)}, \|\cdot\|_\infty) \approx \varepsilon^{-1}(\log 1/\varepsilon)^{3/2}$$

as $\varepsilon \to \infty$. This last result seems to settle an interesting case of the work of [24].

(II) If $Q=[0,1]^d$, $d \geqslant 1$ is an integer, and μ is the centered Gaussian measure on $C(Q)$ induced by the standard d-parameter Brownian sheet with $\phi(\varepsilon)$ as in (5.45) and $p=2$, then [7] shows that as $\varepsilon \to 0$

$$\phi(\varepsilon) \approx \varepsilon^{-2}(\log 1/\varepsilon)^{2d-2}. \tag{5.50}$$

Hence the same argument as used above implies that if

$$K=\left\{f \in C(Q) : f(s_1, ..., s_d)=\int_0^{s_1} \cdots \int_0^{s_d} g(u_1, ..., u_d)\, du_1 \cdots du_d,\right.$$
$$\left.\iint_Q g^2(u_1, ..., du_d)\, du_1 \cdots du_d \leqslant 1\right\}, \tag{5.51}$$

then for $1 \leqslant p \leqslant 2$ and $\varepsilon \to 0$

$$H(\varepsilon, K, \|\cdot\|_p) \approx \varepsilon^{-1}(\log 1/\varepsilon)^{d-1}. \tag{5.52}$$

This, of course, agrees with the previous lower bound in [24] for the W-ball analogue of K in d-variables, If $p=\infty$, we are not aware of an analogue of (5.50) when $d>2$. Perhaps Talagrand's approach will apply.

6. Some Final Remarks

Recently asymptotics for the sup-norm small ball probability, at the logarithmic level, were established in [20] for fractional Brownian motions, $0<\alpha<2$. The corresponding unit balls K_α, $0<\alpha<2$, are given by

$$K=\left\{f(t)=T_\alpha g(t) : 0 \leqslant t \leqslant 1, \int_{R^1} g^2(u)\, du \leqslant 1\right\}, \tag{6.1}$$

where

$$T_\alpha g(t)=\int_0^t (t-u)^{(\alpha-1)/2} g(u)\, du$$
$$+\int_{-\infty}^0 ((t-u)^{(\alpha-1)/2}-(-u)^{(\alpha-1)/2})\, g(u)\, du. \tag{6.2}$$

See, for example [11, p. 66], which also indicates some further references. If μ_α is the centered Gaussian measure on $C[0,1]$ induced by the α-fractional Brownian motion, then [20, Corollary 2.2] implies

$$-\log \mu_\alpha(f \in C[0,1] : \|f\|_\infty \leqslant \varepsilon) \approx \varepsilon^{-2/\alpha} \tag{6.3}$$

as $\varepsilon \to 0$. Hence Theorem 1 implies

$$H(\varepsilon, K_\alpha, \|f\|_\infty) \approx \varepsilon^{-2/(\alpha+1)} \tag{6.4}$$

as $\varepsilon \to 0$. Of course, when $\alpha = 1$, α-fractional Brownian motion is standard Brownian motion, and these results coincide with premvious results. It is interesting to try to compute (6.4) directly from (6.1) and the concept of ε-entropy. We have tried, but it seems difficult.

If $q_\alpha(\cdot)$, $0 < \alpha < 1/2$, is the α-Hölder norm defined in [16] and K is given by (5.51) with $d = 2$, then Theorem 2 of [16] and Theorem 1 above imply the metric entropy of K in the $q_\alpha(\cdot)$ norm satisfies

$$H(\varepsilon, K, q_\alpha(\cdot)) \approx \varepsilon^{-1/(1-\alpha)}(\log 1/\varepsilon)^{(1/2-\alpha)/(1-\alpha)}$$

as $\varepsilon \to 0$. If $d = 1$, and hence K is given as in (1.1), and $\|\cdot\|_\alpha$, $0 < \alpha < 1/2$, is the usual α-Hölder norm for functions of a single variable, then Theorem 1 in [16] and Theorem 1 above imply

$$H(\varepsilon, K, \|\cdot\|_\alpha) \approx \varepsilon^{-1/(1-\alpha)}$$

as $\varepsilon \to 0$. To the best of our knowledge these are new metric entropy results.

References

1. C. Borell, The Brunn–Minkowski inequality, *Invent. Math.* **30** (1975), 207–216.
2. A. A. Borovkov and A. A. Mogulskii, On probabilities of small deviations for stochastic processes, *Siberian Adv. Math.* **1** (1991), 39–63.
3. M. S. Birman and M. Z. Solomjak, Piecewise polynomial approximation of the functions of the classes W_p^α, *Mat. Sb.* **73** (1967), 331–355; *Math. USSR-Sb.* **2** (1967), 295–317 [Engl. transl.].
4. R. H. Cameron and W. T. Martin, The Wiener measure of Hilbert neighborhoods in the space of real continuous functions, *J. Math. Phys.* **23** (1944), 195–209.
5. G. F. Clements, Entropies of several sets of real valued functions, *Pacific J. Math.* **13** (1963), 1085–1095.
6. M. Csörgö and P. Révész, "Strong Approximations in Probability and Statistics," Academic Press, New York, 1981.
7. E. Csaki, On small values of the square integral of a multiparameter Wiener process, *in* "Proc. of the 3rd Pannonian Symp. on Math. Stat.," Reidel, Dordrecht, 1982.
8. V. Goodman, Characteristics of normal samples, *Ann. Probab.* **16** (1988), 1281–1290.
9. V. Goodman, "Some Probability and Entropy Estimates for Gaussian Measures," Progress in Probability, Vol. 20, Birkhauser, Berlin, 1990.
10. V. Goodman and J. Kuelbs, Rates of convergence for increments of Brownian motion, *J. Theoret. Probab.* **1** (1988), 27–64.
11. V. Goodman and J. Kuelbs, Rates of clustering for some Gaussian self-similar processes, *Probab. Theory Related Fields* **88** (1991), 47–75.
12. L. Gross, "Lectures in Modern Analysis and Applications. II," Lecture Notes in Math., Vol. 140, Springer-Verlag, Berlin, 1970.

13. J. Hoffmann-Jørgensen, R. M. Dudley, and L. A. Shepp, On the lower tail of Gaussian seminorms, *Ann. Probab.* **7** (1979), 319–342.
14. A. N. Kolmogorov and V. M. Tihomirov, ε-entropy and ε-capacity of sets in function spaces, *Uspekhi Mat. Nauk.* **14** (1959), 3–86; *Amer. Math. Soc. Transl.* **17** (1961), 277–364 [English transl.].
15. J. Kuelbs, A strong convergence theorem for Banach space valued random variables, *Ann. Probab.* **4** (1976), 744–771.
16. J. Kuelbs and W. V. Li, Small ball estimates for Brownian motion and the Brownian sheet, *J. Theoret. Probab.* (1992), to appear.
17. W. V. Li, Comparison results for the lower tail of Gaussian seminorms, *J. Theoret. Probab.* **5** (1992), 1–31.
18. W. V. Li, On the lower tail of Gaussian measures on l^p, *in* "Progress in Probability," Vol. 30, Birkhäuser, Berlin, 1992.
19. B. S. Mityagin, The approximate dimension and bases in nuclear spaces, *Uspekhi Mat. Nauk* **16** (1961), 63–132; *Russian Math. Surveys* **16** (1961), 59–127 [Engl. transl.].
20. D. Monrad and H. Rootzén, Small values of fractional Brownian motion and locally nondeterministic Gaussian processes, preprint, 1992.
21. E. Seneta, "Regularly Varying Functions," Lecture Notes in Mathematics, Vol. 508, Springer-Verlag, Berlin, 1976.
22. S. A. Smolyak, ε-Entropy of the classes $E_s^{\alpha,k}(B)$ and $W_s^{\alpha}(B)$ in the metric of L^2, *Soviet Math. Dokl.* **1** (1960), 192–196.
23. M. Talagrand, private communication.
24. V. N. Temlyakov, On estimates of ε-entropy and widths of classes of functions with bounded mixed derivative or difference, *Soviet Math. Dokl.* **38** (1989), 84–89.