Bull. London Math. Soc. 36 (2004) 460–468
DOI: 10.1112/S0024609304003121

ENTROPY OF ABSOLUTE CONVEX HULLS IN HILBERT SPACES

FUCHANG GAO

Abstract

The metric entropy of absolute convex hulls of sets in Hilbert spaces is studied for the general case when the metric entropy of the sets is arbitrary. Under some regularity assumptions, the results are sharp.

1. *Introduction*

The Krein–Milman theorem is a powerful tool in analysis. To quantify this theorem, a number of researchers have studied the entropy numbers of the convex hulls of precompact sets in a Banach space or a Hilbert space. The goal is to obtain a sharp upper bound for the entropy of the convex hull $\operatorname{conv}(T)$, knowing the entropy of the set T. The importance of this problem was addressed by Dudley in [**7**], where some special cases were studied. Dudley's results were improved by Ball and Pajor [**2**] and Carl [**4**], and extended to Banach spaces by Carl, Kyrezi and Pajor [**5**]. Recall that

$$N(S, \|\cdot\|, \varepsilon) := \min\left\{ n : \exists\, s_1, s_2, \dots, s_n, \text{ such that } S \subset \bigcup_{k=1}^{n} \{x : \|x - s_k\| < \varepsilon\} \right\}.$$

(When the space is Hilbert, or when there is no confusion, we write $N(S, \varepsilon)$ for short.) With such a notation, their results can be formulated as follows: as $\varepsilon \to 0^+$,

$$\begin{aligned}
&\text{if } N(T,\varepsilon) = O(\varepsilon^{-a}),\ a > 0, && \text{then } \log N(\operatorname{conv}(T),\varepsilon) = O\big(\varepsilon^{-2a/(2+a)}\big);\\
&\text{if } \log N(T,\varepsilon) = O(\varepsilon^{-\alpha}),\ \alpha > 2, && \text{then } \log N(\operatorname{conv}(T),\varepsilon) = O(\varepsilon^{-\alpha});\\
&\text{if } \log N(T,\varepsilon) = O(\varepsilon^{-\alpha}),\ \alpha < 2, && \text{then } \log N(\operatorname{conv}(T),\varepsilon) = O\big(\varepsilon^{-2}(\log(1/\varepsilon))^{1-2/\alpha}\big).
\end{aligned}$$

The critical case $\alpha = 2$ was later solved by Gao [**8**]. The best possible estimate is $\log N(\operatorname{conv}(T),\varepsilon) = O(\varepsilon^{-2}(\log(1/\varepsilon))^2)$. This last result has been extended to Banach spaces [**6**].

Note that the answers are quite different for the case when the growth of $N(T,\varepsilon)$ is of power type, and the case when the growth is of exponential type. A natural question has been asked in [**4**]: to find a sharp upper bound for $N(\operatorname{conv}(T),\varepsilon)$ when $N(T,\varepsilon)$ has an arbitrary rate of growth. When T consists of a sequence of vectors of decreasing length, the question was asked earlier in [**2**]. It should be pointed out that although assuming that T consists of a sequence of vectors of decreasing length can simplify the problem, it does not always give the same upper bound. For example, when T consists of the sequence $\{x_i\}$ with $\|x_i\| \leqslant 1/\sqrt{\log(i+1)}$, then a result of

Received 8 September 2002; revised 7 July 2003.

2000 *Mathematical Subject Classification* 41A46 (primary), 60G15 (secondary).

Research partially supported by NSF grant EPS-0132626, and a seed grant from the University of Idaho.

Talagrand [**16**] says that $\log N(\mathrm{conv}(T),\varepsilon) = O(\varepsilon^{-2}) = \log N(T,\varepsilon)$. However, it was proved in [**8**] that if one assumes only that $\log N(T,\varepsilon) = O(\varepsilon^{-2})$, then the best upper estimate is $\log N(\mathrm{conv}(T),\varepsilon) = O(\varepsilon^{-2}\log^2(1/\varepsilon))$. This is surprising, because such a phenomenon does not appear when ε^{-2} is replaced by $\varepsilon^{-\alpha}$ for $\alpha \neq 2$. Therefore, the case around $\alpha = 2$ seems to be more complicated.

The most interesting case in applications arises when the growth of $N(T,\varepsilon)$ satisfies the Dudley integral condition $\int_0^\infty \log N(T,\varepsilon)\,d\varepsilon < \infty$. Indeed, such a condition guarantees that the Gaussian process X_t, $t \in T$, with $\mathbb{E}\,|X_t - X_s|^2 = \|t-s\|^2$ is sample bounded; that is, $\mathbb{E}\,\sup_{t\in T} X_t < \infty$. Clearly, such a set T must be a bounded set. Therefore the upper limit of the Dudley integral above can be replaced by $\mathrm{diam}(T)$.

Let us also note that the case when T is finite needs to be excluded, because in such a case it is impossible to bound $N(\mathrm{conv}(T),\varepsilon)$ using $N(T,\varepsilon)$. To keep the statements of our results relatively simple, we assume the following slightly stronger condition:

$$\int_0^1 \frac{1}{\varepsilon[N(T,\varepsilon)]^p}\,d\varepsilon < \infty, \quad \text{for some } p > 0.$$

Instead of considering $\mathrm{conv}(T)$, we will consider the absolute convex hull

$$\mathrm{abconv}(T) = \left\{ \sum_{n=1}^\infty a_n t_n : t_n \in T, n \in \mathbb{N}, \sum_{n=1}^\infty |a_n| \leqslant 1 \right\}.$$

The advantage of doing this is that the absolute convex hulls are symmetric, and thus allow us to use duality. Clearly, $N(\mathrm{conv}(T),\varepsilon) \leqslant N(\mathrm{abconv}(T),\varepsilon)$ for all $\varepsilon > 0$.

Our main result is as follows.

THEOREM 1.1. *Let T be a set in a Hilbert space with $\int_0^\infty \sqrt{\log N(T,\varepsilon)}\,d\varepsilon < \infty$. Denote $I(x) = \int_0^x \sqrt{\log N(T,\varepsilon)}\,d\varepsilon$. Suppose that*

$$\int_0^1 \frac{1}{\varepsilon[N(T,\varepsilon)]^p}\,d\varepsilon < \infty, \quad \text{for some } p > 0. \tag{1}$$

Then there exists a constant C such that for any $\varepsilon < \mathrm{diam}(T)$,

$$\log N(\mathrm{abconv}(T),\varepsilon) \leqslant C \inf\left\{ \frac{\eta^2}{\varepsilon^2} + F(\eta) : 0 < \eta < \frac{\mathrm{diam}(T)}{C} \right\},$$

where

$$F(x) = \int_1^{\mathrm{diam}(T)/x} N(T,tx)\frac{\log t}{t}\,dt + \int_1^{x/I^{-1}(x)} N(T,x/t)\exp\left(-\frac{t^2}{\log^4(t+1)}\right)dt.$$

REMARK 1.1. The exponent 4 in the logarithm term above can be replaced by any number larger than 2. However, unless it is near the critical case $\log N(T,\varepsilon) = O(\varepsilon^{-2})$, such a replacement does not lead to any improvement. For example, under the regularity assumption $\log N(T,\varepsilon) \sim c\varepsilon^{-\alpha}\log^{\beta}(1/\varepsilon)$ for some $0<\alpha<2$, and $\beta\in\mathbb{R}$, Theorem 1.1 leads to

$$\log N(\mathrm{abconv}(T),\varepsilon) = O\big(\varepsilon^{-2}(\log(1/\varepsilon))^{1-2/\alpha}(\log\log(1/\varepsilon))^{2\beta/\alpha}\big) \qquad \text{as } \varepsilon\to 0^{+},$$

which is best possible. (See Corollary 2.1 in the next section.)

2. *Proofs*

Let T be a set in Hilbert space H, and B_H the unit ball of H. Without loss of generality, we assume that $0\in T$. Let $\mathcal{M}$ be the set of all the signed measures μ on T. Then

$$\mathrm{abconv}(T) = \left\{\int_T t\mu(dt) : \mu\in\mathcal{M}, \|\mu\|\leqslant 1\right\}.$$

For $x\in H$, define

$$\|x\|_T := \sup_{t\in T}|\langle t,x\rangle| = \sup_{t\in \mathrm{abconv}(T)}|\langle t,x\rangle|.$$

By the duality of entropy numbers (see [**17**]),

$$\log N(\mathrm{abconv}(T),\varepsilon) = O(\log N(B_H,\|\cdot\|_T,\varepsilon)), \tag{2}$$

provided that the regularity of the covering number on the right-hand side holds, as will become clear later in the proof.

It is a striking discovery of Kuelbs and Li [**10**] that metric entropy numbers are closely related to small ball probabilities. Indeed, for $\lambda>0$, denote $D_{\lambda\varepsilon}=\{x\in H : \|x\|_T\leqslant\lambda\varepsilon\}$, and $m=N(B_H,\|\cdot\|_T,\varepsilon)=N(2\lambda B_H,\|\cdot\|_T,2\lambda\varepsilon)$. Then there exist pairwise disjoint sets $a_i+D_{\lambda\varepsilon}$ with $\|a_i\|\leqslant 2\lambda$, $1\leqslant i\leqslant m$.

Let γ be the standard Gaussian measure on H (see, for example, [**3**]). By Anderson's inequality [**1**],

$$\gamma(a_i+D_{\lambda\varepsilon}) \geqslant e^{-\|a_i\|^2/2}\gamma(D_{\lambda\varepsilon}) \geqslant e^{-2\lambda^2}\gamma(D_{\lambda\varepsilon}).$$

Thus

$$N(B_H,\|\cdot\|_T,\varepsilon) = m \leqslant e^{2\lambda^2}/\gamma(D_{\lambda\varepsilon}). \tag{3}$$

Now, all we need to do is to obtain a sharp lower bound for $\gamma(D_{\lambda\varepsilon})$, which is achieved by a standard chaining argument, together with the following lemma. This idea has been used by a number of researchers; see, for example, [**11**, **13**].

LEMMA 2.1. (Khatri–Šidák inequality [**9**, **14**, **15**]).

$$\gamma\left(\left\{\sup_{t\in A}|\langle x,t\rangle|\leqslant\varepsilon, |\langle x,t_0\rangle|\leqslant\varepsilon\right\}\right) \geqslant \gamma\left(\{|\langle x,t_0\rangle|\leqslant\varepsilon\}\right)\cdot\gamma\left(\left\{\sup_{t\in A}|\langle x,t\rangle|\leqslant\varepsilon\right\}\right).$$

The original proof of the Khatri–Šidák inequality is lengthy. A simple proof can be found in [**12**].

We will also use the following simple estimates.

LEMMA 2.2. *Let γ be the standard Gaussian measure in Hilbert space H, and let $\mathbf{e}$ be any unit vector in H. Then*

$$\gamma(\{|\langle x, \mathbf{e}\rangle| < u\}) \geqslant \begin{cases} u/3 & 0 < u < 1, \\ \exp\left(-2e^{-u^2/2}\right) & u \geqslant 1. \end{cases}$$

Now we turn to the proof of Theorem 1.1.

Fix $\eta < \operatorname{diam}(T)/3$. Let T_n be a $2^{-n}I^{-1}(\eta)$-net of T with minimum cardinality, and let n_0 be the largest integer such that T_{n_0} is a singleton. For each $t \in T$, and each integer $n > n_0$, there exists $s_{n-1}(t) \in T_{n-1}$, such that $\|t - s_{n-1}(t)\| \leqslant 2^{-n+1}I^{-1}(\eta)$. Denote $S_{n_0} = T_{n_0}$, and $S_n = \{t - s_{n-1}(t) : t \in T_n\}$ for $n > n_0$. Clearly, $\#S_n \leqslant N(T, 2^{-n}I^{-1}(\eta))$, and for each $y_n \in S_n$, $\|y_n\| \leqslant 2^{-n+1}I^{-1}(\eta)$.

Let k be the smallest integer such that $2^{-k}I^{-1}(\eta) \leqslant \eta$. Because $I(\eta) > \eta$, we have $I^{-1}(\eta) < \eta < \operatorname{diam}(T)/3$. Thus, $n_0 < k \leqslant 0$.

Let $K > 6$ be a constant whose value will be specified later. It is easy to see that for any non-negative sequence $\{c_n\}$ with $\sum_{n \geqslant n_0} c_n \leqslant 1$, we have

$$\left\{x : \sup_{t \in T} |\langle x, t\rangle| < K\eta\right\} \supset \{x : |\langle x, y_n\rangle| \leqslant Kc_n\eta, y_n \in S_n, n \geqslant n_0\}.$$

Applying the Khatri–Šidák inequality (Lemma 2.1), we obtain

$$\gamma(D_{K\eta}) \geqslant \prod_{n=n_0}^{\infty} \prod_{y_n \in S_n} \gamma\left(\{|\langle x, y_n\rangle| \leqslant Kc_n\eta\}\right).$$

Note that for each $y_n \in S_n$, $\|y_n\| \leqslant 2^{-n+1}I^{-1}(\eta)$, we can further write

$$\gamma(D_{K\eta}) \geqslant \prod_{n=n_0}^{\infty} \left(\gamma\left\{|\langle x, \mathbf{e}\rangle| \leqslant \frac{Kc_n 2^{n-1}\eta}{I^{-1}(\eta)}\right\}\right)^{\#S_n}, \tag{4}$$

where $\mathbf{e}$ is a unit vector in H.

We will split (4) into three products: $\prod_{n=0}^{\infty}$, $\prod_{n=k}^{-1}$ and $\prod_{n=n_0}^{k-1}$.

To study the first product, we choose

$$c_n = \frac{1}{3} \cdot \frac{I^{-1}(\eta)2^{-n}}{\eta} \sqrt{\log N(T, 2^{-n}I^{-1}(\eta))}$$

for $n \geqslant 0$. Note that

$$\begin{aligned} \sum_{n=0}^{\infty} c_n &\leqslant \frac{1}{3} \cdot \frac{I^{-1}(\eta)}{\eta} \int_0^1 \sqrt{\log N(T, sI^{-1}(\eta))}\, ds \\ &= \frac{1}{3} \cdot \frac{1}{\eta} \int_0^{I^{-1}(\eta)} \sqrt{\log N(T, s)}\, ds \\ &= \frac{1}{3}. \end{aligned}$$

Applying Lemma 2.2, and using the fact that $\#S_n \leqslant N(T, 2^{-n}I^{-1}(\eta))$, we obtain

$$\begin{aligned}
&\prod_{n=0}^{\infty} \gamma\left(\left\{|\langle x,e\rangle| \leqslant \frac{Kc_n 2^{n-1}\eta}{I^{-1}(\eta)}\right\}\right)^{\#S_n} \\
&\quad = \prod_{n=0}^{\infty} \gamma\left(\left\{|\langle x,e\rangle| \leqslant \frac{K}{6}\sqrt{\log N(T,2^{-n}I^{-1}(\eta))}\right\}\right)^{\#S_n} \\
&\quad \geqslant \prod_{n=0}^{\infty} \exp\left(-2(\#S_n)\cdot \exp\left(-\frac{K^2}{72}\log N(T,2^{-n}I^{-1}(\eta))\right)\right) \\
&\quad \geqslant \prod_{n=0}^{\infty} \exp\left(-2[N(T,2^{-n}I^{-1}(\eta))]^{1-K^2/72}\right) \\
&\quad = \exp\left(-2\sum_{n=0}^{\infty}[N(T,2^{-n}I^{-1}(\eta))]^{1-K^2/72}\right) \\
&\quad \geqslant \exp\left(-2\int_0^{I^{-1}(\eta)} \frac{1}{t[N(T,t)]^{K^2/72-1}}\,dt\right) \\
&\quad \geqslant \frac{1}{2}
\end{aligned} \tag{5}$$

for K large enough, where in the last inequality we used assumption (1).

To bound the second product, we choose $c_n = 1/6(n-k+1)^2$ for $k \leqslant n \leqslant -1$. Then $\sum_{n=k}^{-1} c_n < 1/3$. Note that

$$\frac{Kc_n 2^{n-1}\eta}{I^{-1}(\eta)} \geqslant \frac{K2^{n-k}}{12(n-k+1)^2},$$

which is no less than 1 for $K \geqslant 24$. Applying Lemma 2.2, and using the fact that

$$\#S_n \leqslant N(T,2^{-n}I^{-1}(\eta)) \leqslant N(T,2^{-n+k-1}\eta),$$

we obtain

$$\begin{aligned}
&\prod_{n=k}^{-1} \left(\gamma\left\{|\langle x,\mathbf{e}\rangle| \leqslant \frac{Kc_n 2^{n-1}\eta}{I^{-1}(\eta)}\right\}\right)^{\#S_n} \\
&\quad \geqslant \prod_{n=k}^{-1} \left(\gamma\left\{|\langle x,\mathbf{e}\rangle| \leqslant \frac{K2^{n-k}}{12(n-k+1)^2}\right\}\right)^{\#S_n} \\
&\quad \geqslant \prod_{n=k}^{-1} \exp\left(-2(\#S_n)\cdot \exp\left(\frac{-K^2(2^{n-k})^2}{288(n-k+1)^4}\right)\right) \\
&\quad \geqslant \exp\left(-2\sum_{n=k}^{-1} N(T,2^{-n+k-1}\eta)\cdot \exp\left(\frac{-K^2(2^{n-k})^2}{288(n-k+1)^4}\right)\right) \\
&\quad \geqslant \exp\left(-c\int_1^{\eta/I^{-1}(\eta)} N(T,\eta/t)\exp(-t^2/\log^4(t+1))\,dt\right),
\end{aligned} \tag{6}$$

for some $c > 1$, provided that K is large enough; say, $K > 20$.

To bound the third product, we choose $c_n = 2^{n-k-2}$ for $n_0 \leqslant n \leqslant k-1$. Then $\sum_{n=n_0}^{k-1} c_n \leqslant 1/3$, and

$$\frac{Kc_n 2^{n-1}\eta}{I^{-1}(\eta)} \geqslant Kc_n 2^{n-1} 2^{-k} > 2^{2n-2k}$$

for $K > 8$. Applying Lemma 2.2, and using $\#S_n \leqslant N(T, 2^{-n} I^{-1}(\eta)) \leqslant N(T, 2^{k-n-1}\eta)$, we obtain

$$\begin{aligned}
&\prod_{n=n_0}^{k-1} \gamma\left(\left\{ |\langle x, \mathbf{e}\rangle| \leqslant \frac{Kc_n 2^{n-1}\eta}{I^{-1}(\eta)}\right\}\right)^{\#S_n} \\
&\quad \geqslant \prod_{n=n_0}^{k-1} \gamma(\{|\langle x, \mathbf{e}\rangle| \leqslant 2^{2n-2k}\})^{\#S_n} \\
&\quad \geqslant \prod_{n=n_0}^{k-1} (2^{2n-2k-2})^{N(T, 2^{k-n-1}\eta)} \\
&\quad = \exp\left(-\sum_{n=n_0}^{k-1} 2N(T, 2^{k-n-1}\eta) \log 2^{k-n-1}\right) \\
&\quad \geqslant \exp\left(-c\int_1^{\operatorname{diam}(T)/\eta} N(T, t\eta)\frac{\log t}{t}\, dt\right) \qquad (7)
\end{aligned}$$

for some $c > 1$.

Combining (4)–(7), and recalling the definition of $F(x)$ in the statement of Theorem 1.1, we find that

$$\gamma(D_{K\eta}) \geqslant \frac{1}{2}\exp(-cF(\eta)), \qquad (8)$$

for some $c > 1$, provided that K is large enough, depending on assumption (1).

Applying (8) to (3), we obtain

$$\begin{aligned}
\log N(B, \|\cdot\|_T, \varepsilon) &\leqslant \inf_\lambda\{2\lambda^2 + cF(\lambda\varepsilon/K) : \lambda\varepsilon \leqslant \operatorname{diam}(T)/3\} + \log 2 \\
&\leqslant K' \inf\left\{\frac{\lambda^2}{\varepsilon^2} + F(\lambda) : \lambda < \frac{\operatorname{diam}(T)}{K'}\right\}. \qquad (9)
\end{aligned}$$

Theorem 1 follows by applying (9) to (2).

The statement of Theorem 1.1 can be simplified if the growth of the entropy $N(T,\varepsilon)$ is not extremely slow. Indeed, we have the following corollary.

Corollary 2.1. *Let T be a set in a Hilbert space with $\int_0^\infty \sqrt{\log N(T,\varepsilon)}\, d\varepsilon < \infty$. Denote $I(x) = \int_0^x \sqrt{\log N(T,\varepsilon)}\, d\varepsilon$. Suppose that there exists $C > 1$ such that*

$$N(T, \varepsilon/2) > CN(T, \varepsilon)$$

for all small ε; then there exists a constant K such that

$$\log N(\operatorname{abconv}(T), \varepsilon) \leqslant K \inf\left\{\frac{\lambda^2}{\varepsilon^2} + G(\lambda) : 0 < \lambda < \frac{\operatorname{diam}(T)}{K}\right\},$$

where

$$G(x)=\int_1^{x/I^{-1}(x)} N(T,x/t)\exp\left(-\frac{t^2}{\log^4(t+1)}\right)dt \leqslant N(T,I^{-1}(x)).$$

In particular, if $\log N(T,\varepsilon)\sim c\varepsilon^{-\alpha}\log^{\beta}(1/\varepsilon)$ *for some* $0<\alpha<2$, *and* $\beta\in\mathbb{R}$, *then*

$$\log N(\mathrm{abconv}(T),\varepsilon)\leqslant K\varepsilon^{-2}(\log(1/\varepsilon))^{1-2/\alpha}(\log\log(1/\varepsilon))^{2\beta/\alpha}.$$

Proof. For $t\geqslant 1$, let k be the largest integer such that $t\geqslant 2^k$. Then

$$N(T,tx)\leqslant N(T,2^k x)\leqslant C^{-k}N(T,x)\leqslant \frac{C}{t^{\log_2 C}}N(T,x).$$

Thus

$$\int_1^{\mathrm{diam}(T)/x} N(T,tx)\frac{\log t}{t}\,dt\leqslant N(T,\varepsilon)\int_1^{\infty}\frac{C\log t}{t^{\log_2 C+1}}\,dt\leqslant KN(T,x). \tag{10}$$

Plugging this into the function $F(x)$ in Theorem 1.1, we have

$$F(x)\leqslant KN(T,x)+G(x).$$

Note that for small x, $x/I^{-1}(x)>2$; thus

$$\begin{aligned} G(x)&\geqslant\int_1^2 N(T,x/t)\exp\left(-\frac{t^2}{\log^4(t+1)}\right)dt\\ &\geqslant \frac{1}{C}N(T,x)\int_1^2\exp\left(-\frac{t^2}{\log^4(t+1)}\right)dt\\ &\geqslant\frac{1}{25C}N(T,x). \end{aligned}$$

Hence $F(x)\leqslant K_1G(x)$ for some constant K_1. The first part of the corollary follows.

The inequality $G(x)<N(T,I^{-1}(x))$ is almost trivial. In fact,

$$\begin{aligned} G(x)&=\int_1^{x/I^{-1}(x)} N(T,x/t)\exp\left(-\frac{t^2}{\log^4(t+1)}\right)dt\\ &\leqslant N(T,I^{-1}(x))\int_1^{\infty}\exp\left(-\frac{t^2}{\log^4(t+1)}\right)dt\\ &\leqslant N(T,I^{-1}(x)). \end{aligned}$$

When $\log N(T,\varepsilon)\sim c\varepsilon^{-\alpha}(\log(1/\varepsilon))^{\beta}$ for some $0<\alpha<2$ and $\beta\in\mathbb{R}$, we have

$$I(x)\sim\sqrt{c}\int_0^x \varepsilon^{-\alpha/2}(\log(1/\varepsilon))^{\beta/2}\sim c'x^{1-\alpha/2}(\log(1/x))^{\beta/2}$$

for $x<1$. Thus

$$I^{-1}(x)\sim c''x^{2/(2-\alpha)}[\log(1/x)]^{-\beta/(2-\alpha)}.$$

Hence

$$G(x)\leqslant N(T,I^{-1}(x))\sim\exp\left(Cx^{-2\alpha/(2-\alpha)}[\log(1/x)]^{2\beta/(2-\alpha)}\right),$$

and therefore

$$\begin{aligned} \inf\left\{\frac{\lambda^2}{\varepsilon^2}+G(\lambda)\right\}&\leqslant\inf\left\{\frac{\lambda^2}{\varepsilon^2}+\exp\left(C\lambda^{-2\alpha/(2-\alpha)}[\log(1/\lambda)]^{2\beta/(2-\alpha)}\right)\right\}\\ &\leqslant K\varepsilon^{-2}(\log(1/\varepsilon))^{1-2/\alpha}(\log\log(1/\varepsilon))^{2\beta/\alpha}, \end{aligned}$$

finishing the proof. ∎

REMARK 2.1. Though the trivial inequality $G(x) \leqslant N(T, I^{-1}(x))$ is used to obtain the estimate for the case $\log N(T,\varepsilon) \sim \varepsilon^{-\alpha} \log^{\beta}(1/\varepsilon)$, a simple modification of the example in [**4**] shows that the estimate obtained is indeed sharp. However, when $N(T,\varepsilon)$ has a slow rate of growth, the trivial inequality $G(x) \leqslant N(I^{-1}(x))$ no longer gives the sharp estimate, as we see in Corollary 2.2.

COROLLARY 2.2. *Let T be a set in a Hilbert space with $\int_0^\infty \sqrt{\log N(T,\varepsilon)}\, d\varepsilon < \infty$. Denote $I(x) = \int_0^x \sqrt{\log N(T,\varepsilon)}\, d\varepsilon$. Suppose that there exist constants $1 < C_1 < C_2$ such that for all $\varepsilon < \operatorname{diam}(T)$,*

$$C_1 N(T,\varepsilon) \leqslant N(T,\varepsilon/2) \leqslant C_2 N(T,\varepsilon).$$

Then

$$\log N(\operatorname{abconv}(T),\varepsilon) \leqslant K \inf\left\{\lambda^2 + N(T,\lambda\varepsilon) : 0 < \lambda < \frac{\operatorname{diam}(T)}{\varepsilon}\right\}.$$

In particular, if $N(T,\varepsilon) \sim \varepsilon^{-\alpha}\log^{\beta}(1/\varepsilon)$ for some $\alpha > 0$ and $\beta \in \mathbb{R}$, then

$$\log N(\operatorname{abconv}(T),\varepsilon) \leqslant K N(T,\varepsilon)^{2/(2+\alpha)}.$$

Proof. In view of Corollary 2.1, all we need to show is that

$$G(x) = \int_1^{x/I^{-1}(x)} N(T,x/t) \exp\left(-\frac{t^2}{\log^4(t+1)}\right) dt \leqslant K N(T,x).$$

However, this is easy to see because $N(T,x/t) \leqslant C_2 t^{\log_2 C_2} N(T,x)$ for $t > 1$, and the integral

$$\int_1^\infty t^{\log_2 C_2} \exp\left(-\frac{t^2}{\log^4(t+1)}\right) dt$$

converges.

The estimate for the special case when $N(T,\varepsilon) \sim \varepsilon^{-\alpha}\log^{\beta}(1/\varepsilon)$ is just a straightforward calculation. □

REMARK 2.2. It is not difficult to modify the example in [**2**] to show that the estimate for the case $N(T,\varepsilon) \sim \varepsilon^{-\alpha}\log^{\beta}(1/\varepsilon)$ is also sharp. In general, to study the sharpness of Theorem 1.1, one needs to construct examples when $N(T,\varepsilon)$ is arbitrary. In principle, this can be done if some regularity on $N(T,\varepsilon)$ is assumed.

References

1. T. W. ANDERSON, 'The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities', *Proc. Amer. Math. Soc.* 6 (1955) 170–176.

2. K. BALL and A. PAJOR, *The entropy of convex bodies with 'few' extreme points*, London Mathematical Society Lecture Note Series 158 (Cambridge University Press, 1990) 25–32.

3. V. I. BOGACHEV, *Gaussian measures*, Mathematical Surveys and Monographs 62 (American Mathematical Society, Providence, RI, 1998).

4. B. CARL, 'Metric entropy of convex hulls in Hilbert spaces', *Bull. London Math. Soc.* 29 (1997) 452–458.

5. B. CARL, I. KYREZI and A. Pajor, 'Metric entropy of convex hulls in Banach spaces', *J. London Math. Soc.* (2) 60 (1999) 871–896.

6. J. CREUTZIG and I. STEINWART, 'Metric entropy of convex hulls in type p spaces – the critical case', *Proc. Amer. Math. Soc.* 130 (2002) 733–743 (electronic).

7. R. M. DUDLEY, 'Universal Donsker classes and metric entropy', *Ann. Probab.* 15 (1987) 1306–1326.

8. F. Gao, 'Metric entropy of convex hulls', *Israel J. Math.* 123 (2001) 359–364.
9. C. G. Khatri, 'On certain inequalities for normal distributions and their applications to simultaneous confidence bounds', *Ann. Math. Statist.* 38 (1967) 1853–1867.
10. J. Kuelbs and W. Li, 'Metric entropy and the small ball problem for Gaussian measures', *J. Funct. Anal.* 116 (1993) 133–157.
11. M. Ledoux, 'Isoperimetry and Gaussian analysis', *Lectures on probability theory and statistics* (Saint-Flour, 1994), Lecture Notes in Mathematics 1648 (Springer, Berlin, 1996) 165–294.
12. G. Schechtman, T. Schlumprecht and J. Zinn, 'On the Gaussian measure of the intersection', *Ann. Probab.* 26 (1998) 346–357.
13. Q. M. Shao, 'A note on small ball probability of Gaussian processes with stationary increments', *J. Theoret. Probab.* 6 (1993) 595–602.
14. Z. Šidák, 'Rectangular confidence regions for the means of multivariate normal distributions', *J. Amer. Statist. Assoc.* 62 (1967) 626–633.
15. Z. Šidák, 'On multivariate normal probabilities of rectangles: Their dependence on correlations', *Ann. Math. Statist.* 39 (1968) 1425–1434.
16. M. Talagrand, 'Regularity of Gaussian processes', *Acta Math.* 159 (1987) 99–149.
17. N. Tomczak-Jaegermann, 'Dualité des nombres d'entropie pour des opérateurs à valeurs dans un espace de Hilbert', *C. R. Acad. Sci. Paris Sér. I Math.* 305 (1987) 299–301 (in French).

Fuchang Gao
Department of Mathematics
PO Box 441103
University of Idaho
Moscow
ID 83844-1103
USA

fuchang@uidaho.edu