

Entropy of Absolute Convex Hulls in Hilbert Spaces*

FUCHANG GAO
University of Idaho

August 14, 2002

Abstract

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

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

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 $\text{conv}(T)$, knowing the entropy of the set T . The importance of this problem was addressed by Dudley in [6], where some special cases were studied. Dudley's results were improved by Ball and Pajor [2], Carl [3], and extended to Banach spaces by Carl, Kyrezi and Pajor [4]. Recall

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

(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 $\varepsilon \rightarrow 0^+$,

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

The critical case $\alpha = 2$ was later solved by this author in [7]. The best possible estimate is $\log N(\text{conv}(T), \varepsilon) = O(\varepsilon^{-2}(\log(1/\varepsilon))^2)$. This last result has been extended to Banach spaces ([5]).

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 [3] to find a sharp upper bound for $N(\text{conv}(T), \varepsilon)$ when $N(T, \varepsilon)$ has 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 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 $\|x_i\| \leq 1/\sqrt{\log(i+1)}$, then a result of Talagrand says that $\log N(\text{conv}(T), \varepsilon) =$

$O(\varepsilon^{-2}) = \log N(T, \varepsilon)$. However, it was proved in [7] that if one only assumes $\log N(T, \varepsilon) = O(\varepsilon^{-2})$, then the best upper estimate is $\log N(\text{conv}(T), \varepsilon) = O(\varepsilon^{-2} \log^2(1/\varepsilon))$. 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 case of the most interest in applications is 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.

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 $N(\text{conv}(T), \varepsilon)$ using $N(T, \varepsilon)$. To keep the statements of our results relatively simple, we assume a slightly stronger condition

$$\int_0^1 \frac{1}{\varepsilon [N(T, \varepsilon)]^p} d\varepsilon < \infty$$

for some $p > 0$.

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

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

Clearly, $N(\text{conv}(T), \varepsilon) \leq N(\text{abconv}(T), \varepsilon)$ for all $\varepsilon > 0$.

Our main result is the following

Theorem 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*

$$\int_0^1 \frac{1}{\varepsilon [N(T, \varepsilon)]^p} d\varepsilon < \infty \tag{1}$$

for some $p > 0$. Then, there exists a constant C , such that for any $\varepsilon < \text{diam}(T)$,

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

where

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

Remark 1 The exponent 4 in the logarithm term 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 leads to

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

which is sharp. (See Corollary 1)

Proof of the theorem

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

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

For $x \in H$, define

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

By the duality of entropy numbers (c.f. [14]),

$$\log N(\text{abconv}(T), \varepsilon) = O(\log N(B, \|\cdot\|_T, \varepsilon)), \quad (2)$$

provided the regularity of the covering number on the right hand side, which will become clear later in the proof.

It is a striking discovery of Kuelbs and Li [9] that metric entropy numbers are closely related to small ball probabilities. Indeed, for $\lambda > 0$, denote $D_{\lambda\varepsilon} = \{x \in H : \|x\|_T \leq \lambda\varepsilon\}$, and $m = N(B, \|\cdot\|_T, \varepsilon) = N(2\lambda B, \|\cdot\|_T, 2\lambda\varepsilon)$. Then there exist pairwise disjoint sets $a_i + D_{\lambda\varepsilon}$ with $\|a_i\| \leq 2\lambda$, $1 \leq i \leq m$. Let γ be the standard Gaussian measure on H . By Anderson's inequality [1],

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

Thus,

$$N(B, \|\cdot\|_T, \varepsilon) = m \leq e^{2\lambda^2} / \gamma(D_{\lambda\varepsilon}). \quad (3)$$

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

Lemma 1 (*Khatri-Sidak inequality, [8],[12],[13]*)

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

The original proof of Khatri-Sidak inequality is lengthy. A simple proof can be found in [10].

We will also use the following simple estimates.

Lemma 2 *Let γ be the standard Gaussian measure in Hilbert space H , let \mathbf{e} be any unit vector in H . Then,*

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

Fix $\eta < \text{diam}(T)/3$. 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 $\|t - s_{n-1}(t)\| \leq 2^{-n+1}I^{-1}(\eta)$. Denote $S_{n_0} = T_{n_0}$, and $S_n = \{t - s_{n-1}(t) : t \in T_n\}$ for $n > n_0$. Clearly, $\#S_n \leq N(T, 2^{-n}I^{-1}(\eta))$, and for each $y_n \in S_n$, $\|y_n\| \leq 2^{-n+1}I^{-1}(\eta)$.

Let k be the smallest integer, such that $2^{-k}I^{-1}(\eta) \leq \eta$. Because $I(\eta) > \eta$, we have $I^{-1}(\eta) < \eta < \text{diam}(T)/3$. Thus, $n_0 < k \leq 0$. For each $t \in T$, we can write

$$t = \sum_{n=n_0}^k y_n + \sum_{n=k+1}^{-1} y_n + \sum_{n=0}^{\infty} y_n,$$

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

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

Applying Khatri-Sidak inequality (Lemma 1), we obtain

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

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

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

where \mathbf{e} is any 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 $n \geq 0$. Note that

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

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

$$\begin{aligned} &\prod_{n=0}^{\infty} \gamma \left(\left\{ |\langle x, \mathbf{e} \rangle| \leq \frac{Kc_n 2^{n-1}\eta}{I^{-1}(\eta)} \right\} \right)^{\#S_n} \\ &= \prod_{n=0}^{\infty} \gamma \left(\left\{ |\langle x, \mathbf{e} \rangle| \leq \frac{K}{6} \sqrt{\log N(T, 2^{-n}I^{-1}(\eta))} \right\} \right)^{\#S_n} \\ &\geq \prod_{n=0}^{\infty} \exp \left(-2(\#S_n) \cdot \exp \left(-\frac{K^2}{72} \log N(T, 2^{-n}I^{-1}(\eta)) \right) \right) \end{aligned}$$

$$\begin{aligned}
&\geq \prod_{n=0}^{\infty} \exp\left(-2[N(T, 2^{-n}I^{-1}(\eta))]^{1-K^2/72}\right) \\
&= \exp\left(-2\sum_{n=0}^{\infty}[N(T, 2^{-n}I^{-1}(\eta))]^{1-K^2/72}\right) \\
&\geq \exp\left(-2\int_0^{I^{-1}(\eta)} \frac{1}{t[N(T, t)]^{K^2/72-1}} dt\right) \\
&\geq \frac{1}{2}
\end{aligned} \tag{5}$$

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

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

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

which is no less than 1 for $K \geq 24$. Applying Lemma 2, and using the fact that

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

we obtain

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

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

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

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

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

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

for some $c > 1$.

Combine (4)-(7), and recall the definition of $F(x)$ in the statement of Theorem 1, we obtain that

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

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

Applying (8) to (3), we obtain

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

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

Corollary 1 *If 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

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

where

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

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

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

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

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

Thus,

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

Plugging into the function $F(x)$ in Theorem 1, we have

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

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

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

Hence, $F(x) \leq K_1 G(x)$ for some constant K_1 . The first part of the corollary follows.

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

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

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

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

Therefore

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

finishing the proof.

Remark 2 Though a trivial inequality $G(x) \leq N(I^{-1}(x))$ 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 [3] shows that the estimate obtained is indeed sharp. However, when $N(T, \varepsilon)$ has a slower rate of growth, the trivial inequality $G(x) \leq N(I^{-1}(x))$ no longer gives the sharp estimate, as we will see in Corollary 2.

Corollary 2 *If there exist constants $1 < C_1 < C_2$ such that for all $\varepsilon < \text{diam}(T)$,*

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

Then

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

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

$$\log N(\text{abconv}(T), \varepsilon) \leq KN(T, \varepsilon)^{2/(2+\alpha)}.$$

Proof In view of Corollary 2, all we need to show is that

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

However, this is easy to see, because $N(T, x/t) \leq C_2 t^{\log_2 C_2} N(T, x)$ 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 3 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, 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] Anderson, T. W. 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] Ball, K; Pajor, A. The entropy of convex bodies with "few" extreme points. London Math. Soc. Lecture Note Ser., **158** (Cambridge Univ. Press, 1990.) 25–32.
- [3] Carl, B. Metric entropy of convex hulls in Hilbert spaces. *Bull. London Math. Soc.* **29** (1997), no. 4, 452–458.
- [4] Carl, B; Kyrezi, I; Pajor, A. Metric entropy of convex hulls in Banach spaces. *J. London Math. Soc. (2)* **60** (1999), no. 3, 871–896.

- [5] Creutzig, J; Steinwart, I. Metric entropy of convex hulls in type p spaces—the critical case. *Proc. Amer. Math. Soc.* **130** (2002), no. 3, 733–743. (electronic)
- [6] Dudley, R. M. Universal Donsker classes and metric entropy. *Ann. Probab.* **15** (1987), no. 4, 1306–1326.
- [7] Gao, F. Metric entropy of convex hulls. *Israel J. Math.* **123** (2001), 359–364
- [8] Khatri, C. G. On certain inequalities for normal distributions and their applications to simultaneous confidence bounds. *Ann. Math. Statist.* **38** (1967) 1853–1867.
- [9] Kuelbs, J; Li, W. Metric entropy and the small ball problem for Gaussian measures. *J. Funct. Anal.* **116** (1993), no. 1, 133–157.
- [10] Schechtman, G.; Schlumprecht, Th.; Zinn, J. On the Gaussian measure of the intersection. *Ann. Probab.* **26** (1998), no. 1, 346–357.
- [11] Shao, Q.M. A note on small ball probability of Gaussian processes with stationary increments. *J. Theoret. Probab.* **6** (1993), 595-602.
- [12] Šidák, Z. Rectangular confidence regions for the means of multivariate normal distributions. *J. Amer. Statist. Assoc.* **62** (1967) 626–633.
- [13] Šidák, Z. On multivariate normal probabilities of rectangles: Their dependence on correlations. *Ann. Math. Statist.* **39** (1968) 1425–1434
- [14] Tomczak-Jaegermann, N. Dualité des nombres d’entropie pour des opérateurs à valeurs dans un espace de Hilbert. (French) *C. R. Acad. Sci. Paris Sr. I Math.* **305** (1987), no. 7, 299–301.