Duality of Metric Entropy *

S. Artstein, V. Milman, S. J. Szarek

Abstract

For two convex bodies K and T in $\mathbb{R}^n$, the covering number of K by T, denoted $N(K,T)$, is defined as the minimal number of translates of T needed to cover K. Let us denote by K° the polar body of K and by D the euclidean unit ball in $\mathbb{R}^n$. We prove that the two functions of t, $N(K,tD)$ and $N(D,tK^\circ)$, are equivalent in the appropriate sense, uniformly over symmetric convex bodies $K \subset \mathbb{R}^n$ and over $n \in \mathbb{N}$. In particular, this verifies the duality conjecture for entropy numbers of linear operators, posed by Pietsch in 1972, in the central case when either the domain or the range of the operator is a Hilbert space.

1 Introduction

For two convex bodies K and T in $\mathbb{R}^n$, the covering number of K by T, denoted $N(K,T)$, is defined as the minimal number of translates of T needed to cover K

$$N(K,T) = \min\{N : \exists\, x_1 \dots x_N \in \mathbb{R}^n,\ K \subset \bigcup_{i \le N} x_i + T\}.$$

We denote by D the euclidean unit ball in $\mathbb{R}^n$. In this paper we prove the following duality result for covering numbers.

Theorem 1 (Main theorem) *There exist two universal constants α and β such that for any dimension n and any convex body $K \subset \mathbb{R}^n$, symmetric with respect to the origin, one has*

$$N(D, \alpha^{-1}K^\circ)^{\frac{1}{\beta}} \le N(K,D) \le N(D, \alpha K^\circ)^{\beta} \qquad (1)$$

where $K^\circ := \{u \in \mathbb{R}^n : \sup_{x \in K} \langle x, u\rangle \le 1\}$ is the polar body of K.

*This research was partially supported by grants from the US-Israel BSF (all authors) and the NSF [U.S.A.] (the third-named author).

The best constant β that our approach yields is $\beta = 2 + \varepsilon$ for any $\varepsilon > 0$, with $\alpha = \alpha(\varepsilon)$.

Our theorem establishes a strong connection between the geometry of a set and its polar or, equivalently, between a normed space and its dual. Notice that since the theorem is true for any K, we can actually infer that for any $t > 0$

$$\beta^{-1} \log N(D, \alpha^{-1} t K^\circ) \le \log N(K, tD) \le \beta \log N(D, \alpha t K^\circ). \qquad (2)$$

(For definiteness, above and in what follows all logarithms are to the base 2.) The quantity $\log N(K, t\,T)$ has a clear information-theoretic interpretation: it is the complexity of K, measured in bits, at the level of resolution t with respect to the metric associated with T (e. g., euclidean if $T = D$). Accordingly, (2) means that the complexity of K in the euclidean sense is controlled by that of the euclidean ball with respect to $\| \cdot \|_{K^\circ}$ (the gauge of K°, i.e., the norm whose unit ball is K°), and vice versa, at *every* level of resolution. While it is clear that these complexities should be related, the universality of the link that we establish is somewhat surprising.

In addition to the immediate information-theoretic ramifications, covering numbers appear in many other areas of mathematics. For example, both quantities $N(K, tD)$ and $N(D, tK^\circ)$ enter the theory of Gaussian processes (see, e.g., [D] and [KL], or the survey [L]) and our results transform some conditional statements into theorems (see, e.g., [LL]).

Theorem 1 resolves an old problem, going back to Pietsch ([P], p. 38) and referred to as the "duality conjecture for entropy numbers," in a special yet most important case. The problem can be stated in terms of covering numbers in the following way (below and in what follows we shall abbreviate "symmetric with respect to the origin" to just "symmetric").

Conjecture 2 (Duality Conjecture) *Do there exist two numerical constants $a, b \ge 1$ such that for any dimension n, and for any two symmetric convex bodies K and T in $\mathbb{R}^n$ one has*

$$\log N(T^\circ, aK^\circ) \le b \log N(K, T), \qquad (3)$$

where A° denotes the polar body of A ?

Theorem 1 verifies this conjecture in the case where one of the two bodies is a euclidean ball or, more generally, by affine invariance of the problem, when one of the two bodies is an ellipsoid. In the special case where *both* bodies are ellipsoids it is well known and easy to check that there is equality in (3), with $a = b = 1$.

This conjecture originated in operator theory, and so we restate it below in the language of entropy numbers of operators. For two Banach spaces X and Y, with unit balls $B(X)$ and $B(Y)$ respectively, and for a linear operator $u : X \to Y$, the k^{th} entropy number of u is defined by

$$e_k(u) := \inf\{\varepsilon : N(uB(X), \varepsilon B(Y)) \le 2^{k-1}\}.$$

(In fact, above and in what follows k does not need to be an integer.) So, for example, $e_1(u) = \|u\|_{op}$ (the operator norm), and one can easily see that $e_k(u) \to 0$ as $k \to \infty$ if and only if u is a compact operator. Therefore the two sequences $(e_k(u))$ and $(e_k(u^*))$ always begin with the same number $\|u\|_{op} = \|u^*\|_{op}$, and $e_k(u) \to 0$ if and only if $e_k(u^*) \to 0$. Since the sequence $(e_k(u))$ can be thought of as quantifying the compactness of the operator u, it seems natural to ask to what extent do $(e_k(u))$ and $(e_k(u^*))$ behave similarly. This is the context in which the duality conjecture was originally formulated, and it read as follows.

(Duality Conjecture in the language of entropy numbers) *Do there exist numerical constants $a, b \ge 1$, such that for any two Banach spaces X and Y, any linear operator $u : X \to Y$ and any natural number k, one has*

$$e_{bk}(u^*) \le a e_k(u) \ ?$$

The two formulations are equivalent since considering the entropy numbers of the dual operator $u^* : Y^* \to X^*$ means covering $(B(Y))^\circ$ with (translates of) $\varepsilon(B(X))^\circ$, and since one can restrict oneself to bodies which are convex hulls of a finite number of points and thus lie in a finite dimensional space. This is formulated explicitly in Observation 4 of Section 2 below. In other words, Theorem 1 verifies the duality conjecture (when expressed in terms of entropy numbers) for the case in which one of the two spaces, either the domain or the range of the operator u, is a Hilbert space.

Some special cases of the problem have been studied before, and some particular results were established, see, e.g., [A], [AMS1], [BPST], [GKS], [KM], [MS1], [MS2], [PT], [Pi1], [Pi2], [S], [T]. We mention two of the above which have special relevance to our approach: firstly [KM], which shows the duality for entropy numbers when the rank of the operator is (at most) comparable to the logarithm of the covering number, and secondly [T], which demonstrates a form of duality involving some measures of the size of *entire* sequences $(e_k(u)), (e_k(u^*))$.

The proof of the theorem consists of three parts. The first part is based on a fact has already been formulated and proven in the required form in our paper [AMS1], in which we establish duality up to some factor γ depending on the body K. Next, this step is iterated, each time applied to a different body (for example, a multiple of K intersected with a euclidean ball of some radius), and we bound the covering number by a product of covering numbers of polars. In the third and last step we shrink this product to a product of only two or three factors, establishing duality with absolute constants. Since this is a two sided inequality, almost every statement is divided into two parts. However, there is generally no interplay between the two arguments, and the proofs of the two sides of the inequality can be read independently.

We wish to point out that different iterations could be used. One of them is outlined in a short note [AMS2]. We use here the one that yields the best constant β in the exponent and may potentially lead to a result that is optimal in that regard.

The paper is organized as follows. In section 2 we show how duality is established up to constants depending on the diameter of the set. In sections 3 and 4 we first present an iterating scheme which yields a bound for the covering number in the form of a long product, and then a telescoping argument that shrinks the product to a mere product of two terms. This will complete the proof of the main theorem. Section 5 consists of various additions to the proof. First, we show how to improve the constant $\beta = 6$, given by the method described in sections 2-5, to $\beta = 2 + \varepsilon$ (for any $\varepsilon > 0$, and with $\alpha = \alpha(\varepsilon)$). Next, we state a related conjecture and several results associated with that conjecture.

Remark on notation: Unless otherwise stated, above and in what follows all constant appearing are universal (notably independent of the dimension and of the particular convex body or the operator that is being considered). If a constant c depends on some parameter θ, we will indicate that by writing $c(\theta)$.

2 A first step toward duality

For a symmetric convex body $K \subset \mathbb{R}^n$, denote by k the logarithm of its covering number (so that $N(K, D) = 2^k$), and define the parameter

$$\gamma(K) := \max\{1, M^*(K \cap D)\sqrt{\frac{n}{k}}\},$$

where, as usual, $M^*(A)$ denotes half the mean width of the set A, that is, $M^*(A) = \int_{S^{n-1}} \sup_{y \in A} \langle u, y \rangle \; d\mu(u)$, with μ the normalized Haar measure on the sphere S^{n-1}.

The first step of the proof of the main theorem is a duality result involving the parameter γ instead of a universal constant α. The following lemma is a combination of two statements, the first of which appeared in [MS1] and the second in [AMS1].

Lemma 3 (First step) *There exists a universal constant* $c_2 > 0$ *and, for every* $\varepsilon > 0$*, a constant* $C_2(\varepsilon) > 0$ *such that, for any dimension* n *and for any symmetric convex body* $K \subset \mathbb{R}^n$*, denoting* $\gamma = \gamma(K)$*, we have*

$$N(K, D) \le N(D, \frac{c_2}{\gamma} K^\circ)^3 \tag{4}$$

and

$$N(D, C_2(\varepsilon)\gamma K^\circ) \le N(K, D)^{1+\varepsilon}. \tag{5}$$

For a general body $K \subset \mathbb{R}^n$ the parameter γ can be as large as $\sqrt{\frac{n}{k}}$. In Observation 4 below we explain why we can restrict our considerations to a certain special class of convex bodies, namely the convex hulls of not too many points. In the rest of the section we will show that in this class there are good bounds for γ.

Observation 4 *For any convex body* K*, any set* $S \subset K \subset S + D$ *of cardinality* $N(K, D)$ *and for any* $\rho > 0$ *we have, denoting* $T = \mathrm{conv}(S)$*,*

$$N(D, (2\rho + 2)K^\circ) \le N(D, \rho T^\circ).$$

Similarly, if $N(K, D) > 1$*, then we can find* $S \subset K$ *of cardinality* $N(K, D)$ *such that* $\mathrm{diam}(S) = \mathrm{diam}(K)$ *and that, denoting* $T = \mathrm{conv}(S)$*, we have*

$$N(K, D) \le N(T, \frac{1}{2} D).$$

Remark: The argument does not require that the original body lies in a finite dimensional space (whereas a convex hull of finitely many points obviously does). In particular, this shows the equivalence of the operator theoretic formulation of the duality conjecture and the finite dimensional analogue with universal constants.

Proof Obviously $T \subset K \subset T + D$. Denoting $N(D, \rho T^\circ) = N$, we can pick a ρT°-net $\{y_i\}_{i=1}^N$ for D, i.e., $D \subset \cup_{i=1}^N y_i + \rho T^\circ$. We want to pass to a net *inside* D, for this notice that $y_i + \rho T^\circ$ intersects D, say at a point z_i, and that $\{z_i\}_{i=1}^N$ is a $2\rho T^\circ$-net for D. We claim that $\{z_i\}_{i=1}^N$ is a $(2\rho + 2)K^\circ$-net of D. Indeed, for every $y \in D$ there exists a z_i in the net such that $y - z_i \in 2\rho T^\circ$, i.e., $\sup_{x \in T}(y - z_i, x) \le 2\rho$. Hence (using that $y - z_i$ is in $2D$) we have $\sup_{x \in T+D}(y - z_i, x) \le 2\rho + 2$, which means precisely that $\|y - z_i\|_{(T+D)^\circ} \le 2\rho + 2$. In particular, since $K \subset T + D$, we see that $\|y - y_i\|_{K^\circ} \le \rho + 2$, as required. We conclude that $N(D, (2\rho + 2)K^\circ) \le N$, and this verifies the first part of the observation.

For the second part, we denote this time $N = N(K, D)$ and pick a 1-separated set $\{x_i\}_{i=1}^N$ in K which realizes the diameter. We do this simply by choosing two points, the distance between which is the diameter of K, and completing them to a 1-separated set of cardinality N. Again, this is possible since a maximal separated set has at least as many elements as the minimal covering. Denote $T = \mathrm{conv}\{x_i\}$. Since $\{x_i\}$ were 1-separated, $N(T, \frac{1}{2}D) \geq N$. This completes the demonstration of the observation. $\square$

The following proposition is an estimate for $\gamma(K)$ which is valid whenever K is the convex hull of $\leq 2^k$ points in RD and, in addition, has a covering number $\leq 2^k$. It was established in [MS1]. The general conjecture, which still remains open, is that for this class of bodies the parameter γ is bounded by a universal constant, regardless of the diameter of the body. If this were true, Lemma 3 and Observation 4 would imply the duality of entropy numbers (with $1 + \varepsilon$ in the exponent!). We discuss the conjecture in Section 5; for a more elaborate discussion and related results we refer the reader to [MS1].

Proposition 5 (An $O(\log^3 R)$ estimate for γ) *There exists a universal constant C_0 such that if a set $S \subset RD \subset \mathbb{R}^n$ (for some $R > 1$) consists of 2^k points, and if $N(K, D) \leq 2^k$ for $K = \mathrm{conv}S$, then*

$$M^*(K \cap D) \leq C_0 (\log R)^3 \sqrt{\frac{k}{n}}.$$

Lemma 3, Observation 4 and the above Proposition 5 can be combined as follows. Denote $\psi(x) = 2C_2(C_0 \log^3 x + 1) + 2$, where $C_2 = C_2(1)$ comes from Lemma 3.

Corollary 6 (Duality up to $\psi(R)$) *If $K \subset RD \subset \mathbb{R}^n$ then*

$$N(D, \psi(R)K^\circ) \leq N(K, D)^2 \tag{6}$$

and

$$N(K, D) \leq N(D, (1/\psi(R))K^\circ)^3. \tag{7}$$

3 An iterating scheme

In this section we present an iterating procedure that gives a bound for the covering number. The first lemma is based on a simple geometric iteration procedure (and admits a variant which is valid in the non-euclidean case; see Remark 11).

Lemma 7 (Iterating procedure) *For any symmetric convex body $K \subset \mathbb{R}^n$ and any sequence $R_0 < R_1 < \cdots < R_s$,*

$$N(D, R_0 K^\circ) \leq N(D, R_s K^\circ) \prod_{j=0}^{s-1} N(D, \frac{R_j}{2}(K \cap R_{j+1} D)^\circ), \tag{8}$$

and

$$N(K, R_0 D) \leq N(K, R_s D) \prod_{j=0}^{s-1} N(2K \cap R_{j+1} D, R_j D). \tag{9}$$

Proof For (8) consider the following inequality

$$N(D, R_0 K^\circ) \leq N(D, \frac{R_0}{2}\mathrm{conv}(K^\circ \cup \frac{1}{R_1} D)) N(\frac{R_0}{2}\mathrm{conv}(K^\circ \cup \frac{1}{R_1} D), R_0 K^\circ),$$

which follows from the sub-multiplicativity of covering numbers: for every A, B and C it is true that $N(A, B) \leq N(A, C) N(C, B)$. Rewriting the first term on the right hand side, changing the convex hull in the second term to the Minkowski sum of sets (which is bigger and thus harder to cover) and using the rule $N(A + C, B + C) \leq N(A, B)$ leads to

$$N(D, R_0 K^\circ) \leq N(D, \frac{R_0}{2}(K \cap R_1 D)^\circ) N(D, R_1 K^\circ).$$

Repeating the above argument another $(s-1)$ times yields (8).

To show (9) we first notice that

$$N(K, R_0 D) \leq N(K, R_1 D) N(R_1 D \cap 2K, R_0 D),$$

where we use the fact that $N(K, R_1 D \cap 2K) = N(K, R_1 D)$, since the centers of a covering by euclidean balls may always be assumed to lie inside K, and also use sub-multiplicativity of covering numbers. Iterating this inequality gives (9). The proof of Lemma 7 is thus complete. □

Now is the time to choose the sequence (R_j). In fact, we will choose two different sequences, each corresponding to a different inequality in the main theorem. There is much freedom in this choice, and we do not suggest that our choice is optimal.

For the first sequence, let R_0 be a large constant to be specified later. Define R_{j+1} by the formula

$$\frac{\sqrt{R_j}}{2} = \psi\left(\frac{R_{j+1}}{\sqrt{R_j}}\right).$$

Remembering that $\psi(x) = 2C_2(C_0(\log x)^3 + 1) + 2$, the above means that

$$R_{j+1} = \sqrt{R_j}\ \exp\left(((\sqrt{R_j} - 4 - 4C_2)/4C_2C_0)^{1/3}\right).$$

In particular, if R_0 is large enough then this sequence increases to ∞. (This is needed since we will later use the fact that $N(D, R_jK^\circ) = 1$ for j large enough.) Corollary 6 together with Lemma 7 imply now the following

Corollary 8 *With the above choice of the sequence* (R_j) *we have, for every symmetric convex body* K*,*

$$N(D, R_0K^\circ) \le N(D, R_sK^\circ)\prod_{j=0}^{s-1} N(K \cap R_{j+1}D, \sqrt{R_j}D)^2. \qquad (10)$$

Proof To deduce Corollary 8 from Lemma 7 we only need to explain the inequality

$$N(D, \frac{R_j}{2}(K \cap R_{j+1}D)^\circ) \le N(K \cap R_{j+1}D, \sqrt{R_j}D)^2.$$

To this end, rewrite

$$\begin{aligned}
N\left(D, \frac{R_j}{2}(K \cap R_{j+1}D)^\circ\right) &= N\left(D, \frac{\sqrt{R_j}}{2}\left(\frac{K}{\sqrt{R_j}} \cap \frac{R_{j+1}}{\sqrt{R_j}}D\right)^\circ\right)\\
&= N\left(D, \psi\left(\frac{R_{j+1}}{\sqrt{R_j}}\right)\left(\frac{K}{\sqrt{R_j}} \cap \frac{R_{j+1}}{\sqrt{R_j}}D\right)^\circ\right)\\
&\le N\left(\frac{K}{\sqrt{R_j}} \cap \frac{R_{j+1}}{\sqrt{R_j}}D, D\right)^2\\
&= N(K \cap R_{j+1}D, \sqrt{R_j}D)^2,
\end{aligned}$$

where for the inequality we used (6) of Corollary 6. □

For the proof of the other side of the inequality in the main theorem we have a different condition on the sequence to make this type of argument work. Again, let R'_0 be a big constant to be specified later. Define R'_{j+1} by

$$\psi\left(\frac{R'_{j+1}}{R'_j}\right) = \frac{\sqrt{R'_j}}{2},$$

which can be rewritten as $R'_{j+1} = R'_j\ \exp\left(\left(\left(\sqrt{R'_j} - 4 - 4C_2\right)/4C_2C_0\right)^{1/3}\right)$. Again, it is clear that this sequence is increasing to ∞.

Corollary 9 *With the above choice of a sequence R'_j we have, for every convex symmetric body K,*

$$N(K, R'_0 D) \leq N(K, R'_s D) \prod_{j=0}^{s-1} N(D, \sqrt{R'_j}(K \cap \frac{R'_{j+1}}{2} D)^\circ)^3. \quad (11)$$

Proof Again, we will use Lemma 7 together with Corollary 6. Here we should explain the inequality

$$N(2K \cap R'_{j+1} D, R'_j D) \leq N(D, \sqrt{R'_j}(K \cap \frac{R'_{j+1}}{2} D)^\circ)^3.$$

This is even simpler since

$$\begin{aligned} N(2K \cap R'_{j+1} D, R'_j D) &= N\left(\frac{2}{R'_j} K \cap \frac{R'_{j+1}}{R'_j} D, D\right) \\ &\leq N\left(D, \frac{R'_j}{2\psi\left(\frac{R'_{j+1}}{R'_j}\right)}(K \cap \frac{R'_{j+1}}{2} D)^\circ\right)^3 \\ &= N(D, \sqrt{R'_j}(K \cap \frac{R'_{j+1}}{2} D)^\circ)^3, \end{aligned}$$

where for the inequality we used (7) of Corollary 6. □

4 Telescoping the long product

In this last step we collapse the long products of covering numbers appearing in (10) and (11) to products consisting of just two terms. The largest R_s (respectively R'_s) will be chosen to exceed the diameter of the set, and so the terms $N(K, R'_s D)$ and $N(D, R_s K^\circ)$ will both equal 1. We need the following two super-multiplicativity inequalities for covering numbers which are valid for any symmetric convex body K.

Lemma 10 *Let $A > a > 3B > 3b$. Then*

$$\begin{aligned} N(K \cap AD, aD) N(K \cap BD, bD) &\leq N(K \cap AD, \frac{b}{4} D) \quad (12) \\ N(D, a(K \cap AD)^\circ) N(D, b(K \cap BD)^\circ) &\leq N(D, \frac{b}{4}(K \cap AD)^\circ). \quad (13) \end{aligned}$$

Proof Since K enters the inequlities only via its intersections with balls of radii $\leq A$, we may as well assume that $K = K \cap AD$ to begin with. For the first inequality, denote $N_1 = N(K, aD)$ and $N_2 = N(K \cap BD, bD)$.

Pick an a-separated set $x_1, \dots x_{N_1}$ in K and a b-separated set $y_1, \dots y_{N_2}$ in $K \cap BD$ (both separations with respect to the euclidean norm). Define a new set by $z_{i,j} = x_i/2 + y_j/2$. All these points are in K, and there are $N_1 N_2$ of them. We shall show that, in addition, $z_{i,j}$'s are $(b/2)$-separated; this will imply $N(K, \frac{b}{4}D) \geq N_1 N_2$, as required. To show the asserted separation, we consider two cases. First, if we look at $|z_{i,j} - z_{i,k}|$, this is simply $|y_j - y_k|/2$ and it exceeds $b/2$. On the other hand, if $k \neq i$, then $|z_{i,j} - z_{k,l}| \geq |x_i - x_k|/2 - |y_j - y_l|/2$, and using the fact that the y_i's are in BD we see that these quantities are greater than $a/2 - B$, which in turn exceeds $\frac{b}{2}$. This completes the proof of inequality (12).

For the second inequality in the Lemma, denote $N_1 = N(D, aK^\circ)$ and $N_2 = N(D, b(K \cap BD)^\circ)$. Pick sets $\{x_1, \dots, x_{N_1}\}$ and $\{y_1, \dots y_{N_2}\}$ in D which are respectively aK°-separated and $b(\alpha K^\circ + \frac{1-\alpha}{B} D)$-separated, where $\alpha = \frac{a}{2a-b} \in (0,1)$ (note that $\alpha K^\circ + \frac{1-\alpha}{B} D \subset \mathrm{conv}(K^\circ \cup \frac{1}{B} D) = (K \cap BD)^\circ$). Define $z_{i,j} = \frac{b}{2a} x_i + (1 - \frac{b}{2a}) y_j$. All these points are in D, and there are $N_1 N_2$ of them. As above, it will be enough to show that the $z_{i,j}$'s are $\frac{b}{2}K^\circ$-separated, i.e., whenever $(i,j) \neq (k,l)$, then

$$z_{k,l} \notin z_{i,j} + \frac{b}{2} K^\circ.$$

When looking at $j = l$, this is the same as asking that $\frac{b}{2a} x_k \notin \frac{b}{2a} x_i + \frac{b}{2} K^\circ$, which follows from the separation of x_i and x_k. When looking at $j \neq l$ and noticing that $x_i, x_k \in D$, we see that it suffices to show that

$$(1 - \frac{b}{2a}) y_j \notin (1 - \frac{b}{2a}) y_l + 2\frac{b}{2a} D + \frac{b}{2} K^\circ.$$

Under our hypotheses, the above follows from the separation of y_l and y_j. Indeed, $\frac{1}{2}(1 - \frac{b}{2a})^{-1} = \alpha$ just by the definition of α. On the other hand, it is readily verified that the assumption $a > 3B > 2B + b$ implies $\frac{1}{a}(1 - \frac{b}{2a})^{-1} < \frac{1-\alpha}{B}$. The proof is thus complete. $\square$

Remark 11 *(i) With more careful argument, any factor less than* $1/2$ *instead of* $1/4$ *can be obtained in the lemma (with then stronger conditions on* a, B*). Moreover, if we work with separated sets instead of covering numbers, then we may arrive at any factor less than 1: for a factor* $1 - \varepsilon$ *(which corresponds to* $\frac{1}{2} - \frac{\varepsilon}{2}$ *for covering numbers), we need the condition* $a > 3(\frac{1-\varepsilon}{\varepsilon})B$*. This may be used to improve slightly the constants in our main theorem.*

(ii) Notice that D *plays no special role in Lemma 10; the same inequalities hold for two general symmetric convex bodies* K *and* T *(i.e., with* D *replaced by an arbitrary* T*).*

Proof of the Main Theorem We will successively apply Lemma 10 to the long products in (10) and (11). However, an additional trick is required since for two neighboring factors in the products the condition of Lemma 10 does not hold, and so they cannot be "collapsed." For example, for two such factors in (10) one has $a = \sqrt{R_j}$ and $B = R_j$, and so one cannot hope for $a > 3B$. The trick is to split the product into two parts, by grouping separately the factors corresponding to the odd and the even j's. The growth of R_j is fast enough so that the conditions of Lemma 10 are satisfied for each two consecutive odd factors, and for each two consecutive even factors. We provide details for the product from (10); the analysis of (11) is fully analogous.

First choose s to be the smallest even number so that $R_s > \operatorname{diam}(K)$. Then the product in (10) which bounds $N(D, R_0K^\circ)$ can be written as (we omit the power 2 for the moment)

$$\prod_{j=1}^{s/2} N(K \cap R_{2j}D, \sqrt{R_{2j-1}}D) \prod_{j=1}^{s/2} N(K \cap R_{2j-1}D, \sqrt{R_{2j-2}}D).$$

For the first collapsing step in each of the two sub-products we need to check that $\sqrt{R_{s-1}} > 3R_{s-2}$, and that $\sqrt{R_{s-2}} > 3R_{s-3}$. When $R_{j+1} = \sqrt{R_j}\, \exp\big(((\sqrt{R_j} - 4 - 4C_2)/4C_2C_0)^{1/3}\big)$, these conditions clearly hold as long as R_{s-3} is larger than some numerical constant C. Since $R_j > R_0$ for $j > 0$, it is enough to start with R_0 which is big enough. (To be able to later compare the obtained expressions with $N(K, D)$, we also insist that $R_0 \geq 16$.) After this first step, using Lemma 10, the product becomes (bounded by)

$$N(K \cap R_sD, \frac{\sqrt{R_{s-3}}}{4}D) \prod_{j=1}^{s/2-2} N(K \cap R_{2j}D, \sqrt{R_{2j-1}}D)$$
$$N(K \cap R_{s-1}D, \frac{\sqrt{R_{s-4}}}{4}D) \prod_{j=1}^{s/2-2} N(K \cap R_{2j-1}D, \sqrt{R_{2j-2}}D).$$

From here onward all the steps are the same; we just need to make sure at each stage j that

$$(\sqrt{R_j}/4)/(R_{j-1}/2) \geq 3. \tag{14}$$

As before, this is indeed satisfied if $R_j > C$, which is assured since we insist that $R_0 > C$. We point out that the factors $1/4$ in $b/4$ do not accumulate, but enter into the quantity a of the next step. Continuing this way with all the factors of these products, we arrive at

$$N(D, R_0K^\circ) \leq N(K \cap R_sD, \frac{\sqrt{R_0}}{4}D)^2 N(K \cap R_{s-1}D, \frac{\sqrt{R_1}}{4}D)^2,$$

(we have inserted back the power 2 in (10)) which, having insisted that $R_0 \geq 16$, implies

$$N(D, R_0 K^\circ) \leq N(K, D)^4.$$

Similarly, in the other direction we use Corollary 9 and inequality (13) to obtain

$$N(K, R_0 D) \leq N(D, K^\circ)^6,$$

and the proof of the main theorem is complete. □

As mentioned earlier, a large part of this proof carries over to the case of two *general* convex bodies. We summarize this in the following conditional proposition. (Our decision to include this statement in the form below was influenced by discussions with Nicole Tomczak-Jaegermann.)

Proposition 12 *Let T be a convex symmetric body in a euclidean space such that, for some constants $c, C > 0$, the following holds: if K is a convex symmetric body with $K \subset 4T$, then*

$$N(K, T) \leq N(T^\circ, cK^\circ)^C.$$

Then, for some other constants $c', C' > 0$ (depending only on c, C) and **any** *convex symmetric body K*

$$N(K, T) \leq N(T^\circ, c'K^\circ)^{C'}.$$

Dually, if K is fixed and the hypothesis holds for all T's verifying $K \subset 4T$, then the assertion holds for **any** *T.*

Proof The argument consists of two parts. The first part is essentially a copy of the proof of Lemma 7 for the choice $R_j = 2^j$. The only difference is that an extra factor 2 appears in the analogue of (9) since at each step

$$\begin{aligned} N(K, 2^j T) &\leq N(K, 2^{j+2}T \cap 2K) N(2^{j+2}T \cap 2K, 2^j T) \\ &\leq N(K, 2^{j+1}T) N(2^{j+2}T \cap 2K, 2^j T) \end{aligned}$$

as we can no longer assume that the centers of the covering are inside K. In this way we show the inequality (similar to (9))

$$N(K, T) \leq N(K, 2^s T) \prod_{j=0}^{s-1} N(2^{1-j}K \cap 4T, T)$$

and, dually, another inequality similar to (8),

$$N(K, T) \leq N(K, 2^s T) \prod_{j=0}^{s-1} N(K, \mathrm{conv}(2^{j-1}T \cup \frac{1}{4}K)).$$

For each factor in these products the body that is being covered is included in 4 times the covering body, and so we may use the assumption (as before, we take s to be the smallest integer such that $N(K, 2^s T) = 1$) to pass to a product of dual covering numbers. Thus for example we get that

$$N(K,T) \le \prod_{j=0}^{s-1} N(T^\circ, c2^{j-1}(K \cap 2^{j+1}T)^\circ)^C$$

(and a similar estimate if we use the second inequality instead). We then collapse the remaining product in the same way as in the euclidean case, using Remark 11 (ii). Note that we may now have to split the product into more than 2, say l, subproducts to make sure that all neighboring factors in each subproduct satisfy the condition of Lemma 10. However, the resulting l depends only on c. We thus arrive (in both cases) at

$$N(K,T) \le N(T^\circ, \frac{c}{8}K^\circ)^{C\log_2(48/c)}.\Box$$

5 Improving the constant in the exponent

In this section we explain how to improve the constant $\beta = 6$ in (1) that we have obtained in sections 2-4 to a constant $\beta = 2(1+\varepsilon)$. The proof presented above is somewhat non-symmetric. As described, in one of the inequalities we get $\beta = 6$, and in the other $\beta = 4$. This $\beta = 4$ can be improved to $2(1+\varepsilon)$ if we continue working with a general ε in Lemma 3 and do not specify (as we did only to lighten the notation) $\varepsilon = 1$. However, as stated above, we cannot in a straightforward way obtain $\beta = 2(1+\varepsilon)$ in the other inequality. Below we explain how to arrive at $\beta = 2(1+\varepsilon)$ there as well. We take this opportunity to elaborate upon the conjecture called "the Geometric Lemma," which we mentioned in passing in Section 2.

Conjecture 13 (Geometric Lemma) *There exists an absolute constant C_0 such that for every dimension n and for every set $S = \{x_i\}_{i=1}^{2^k} \subset \mathbb{R}^n$ verifying $N(K,D) \le 2^k$, where $K = \mathrm{conv}(S)$, we have*

$$M^*(K \cap D) \le C_0\sqrt{\frac{k}{n}}.$$

Notice that this is precisely a version of Proposition 5 with no dependence on the radius of the set. We believe that its importance may transcend its relevance to entropy numbers.

Considering now the dual situation, we can formulate a dual version of the Geometric Lemma, substituting the condition $N(K, D) \le 2^k$ by the condition $N(D, K^\circ) \le 2^k$. However, having proved the duality of entropy numbers, it is easy to see that the Geometric Lemma and its dual version are formally equivalent. Moreover, every estimate such as Proposition 5 can be applied now to the dual situation. For example the main theorem together with Proposition 5 gives the following

Proposition 14 (A dual $O((\log R)^3)$ estimate for $M^*(K \cap D)$) *There exists a universal constant C_0 such that if a set $S \subset RD \subset \mathbb{R}^n$ consists of 2^k points, and if $N(D, K^\circ) \le 2^k$ for $K = \mathrm{conv}S$, then*

$$M^*(K \cap D) \le C_0(\log R)^3\sqrt{\frac{k}{n}}.$$

We can thus define a new parameter, $\gamma'(K)$, to be

$$\gamma'(K) := \max\{1, M^*(K \cap D)\sqrt{\frac{n}{\log N(D, K^\circ)}}\},$$

and repeating the argument of Lemma 3 we obtain the following

Lemma 15 (Dual First step) *For every $\varepsilon > 0$ there is a constant $C_2' = C_2'(\varepsilon)$ such that for any dimension n and for any symmetric convex body $K \subset \mathbb{R}^n$, denoting $\gamma' = \gamma'(K)$ we have*

$$N(K, C_2'\gamma' D) \le N(D, K^\circ)^{1+\varepsilon}. \tag{15}$$

Employing the same line of argument as earlier, but using Proposition 14 as an estimate on γ', and inequality (15) at every step, we are now able to obtain $\beta = 2 + \varepsilon$, instead of $\beta = 6$, also in the other inequality involved in the duality of metric entropy.

To end this section, and the paper, we present another proposition which is an application of both our results and our methods, and gives an interesting link between Geometric Lemma type results and the behavior of covering numbers under projections. Its proof follows the same lines as that of Lemma 3, and other variants involving additional parameters are possible. Below we use the standard jargon of the asymptotic theory of normed spaces, saying that a property is satisfied for a "random" projection of given rank k if it holds for a set of (orthogonal) rank k projections whose measure tends to 1 as the relevant parameters (k, n below) tend to infinity (where "measure" = "the normalized Haar measure on the corresponding Grassmann manifold").

Proposition 16 *There exist universal constants c_1, C_1, and for every λ there exists a constant $C_2 = C_2(\lambda)$ depending only on λ such that, for any $K \subset \mathbb{R}^n$ with $N(K, D) = 2^k$ and any integer t_0 with $k \le t_0 \le n$ we have*

(i) If $M^*(K \cap D) \le \sqrt{\frac{t_0}{n}}$, then for every integer t with $t_0 \le t \le n$, the random rank t projection P_t satisfies

$$N(P_t K, C_1 \sqrt{\frac{t}{n}} D) \le 2^k \text{ and } N(P_t K, c_1 \sqrt{\frac{t}{n}} D) \ge 2^k.$$

(ii) In the other direction, if the random rank t_0 projection P_{t_0} verifies

$$N(P_{t_0} K, \lambda \sqrt{\frac{t_0}{n}} D) \le 2^k,$$

then necessarily $M^*(K \cap D) \le C_2 \sqrt{\frac{t_0}{n}}$ and, for any *integer t with $t_0 \le t \le n$, the random rank t projection P_t satisfies*

$$N(P_t K, C_2 \sqrt{\frac{t}{n}} D) \le 2^k \text{ and } N(P_t K, c_1 \sqrt{\frac{t}{n}} D) \ge 2^k.$$

Thus we observe - as is typical in the asymptotic geometric analysis - a unified form of behavior for *all* dimensions and *all* convex bodies.

We note that our Proposition 5 implies, in the case when K is a convex hull of 2^k points, that the critical t_0 in the above proposition is bounded from above by $C_0 k (\log k)^6$, for details see [MS1] (to pass from estimates on the diameter to estimates on $\log N(K, D)$). Notice that the validity of Conjecture 13 would imply that for this class of bodies in fact $t_0 \le C_0 k$ for a universal C_0. Also, our main theorem implies that Proposition 16 remains true if we replace the condition on $N(K, D)$ with a similar one on $N(D, K^\circ)$ (with additional universal constants). Similarly, we may replace the estimates on the behavior of covering numbers under projections with their dual analogues, describing the behavior of covering numbers under intersections with random subspaces.

References

[A] S. Artstein, *Proportional concentration phenomena on the sphere.* Israel J. Math. 132 (2002), 337–358.

[AMS1] S. Artstein, V.D. Milman, S.J. Szarek, *More on the duality conjecture for entropy numbers.* C. R. Math. Acad. Sci. Paris 336 (2003), no. 6, 479–482.

[AMS2] S. Artstein, V. D. Milman, S. J. Szarek, *Duality of metric entropy in Euclidean space.* C. R. Math. Acad. Sci. Paris 337 (2003), no 11, 711-714.

[BPST] J. Bourgain, A. Pajor, S. J. Szarek, N. Tomczak-Jaegermann, *On the duality problem for entropy numbers of operators.* Geometric aspects of functional analysis (1987–88), Lecture Notes in Math., 1376, Springer, Berlin-New York (1989) 50–63.

[D] R. M. Dudley, *The sizes of compact subsets of Hilbert space and continuity of Gaussian processes.* J. Funct. Anal. 1 (1967) 290–330.

[GKS] Y. Gordon, H. König, C. Schütt, *Geometric and probabilistic estimates for entropy and approximation numbers of operators.* J. Approx. Theory 49 (1987), no. 3, 219–239.

[KM] H. König, V. D. Milman, *On the covering numbers of convex bodies*, Geometric aspects of functional analysis (1985–86), Lecture Notes in Math., vol. 1267, Springer, Berlin-New York (1987) 82–95.

[KL] J. Kuelbs, W. V. Li, *Metric entropy and the small ball problem for Gaussian measures.* J. Funct. Anal. 116 (1993), no. 1, 133–157.

[L] M. Ledoux, *Isoperimetry and Gaussian analysis. Lectures on probability theory and statistics (Saint-Flour, 1994)*, 165–294, Lecture Notes in Math., 1648, Springer, Berlin, 1996.

[LL] W. V. Li, W. Linde, *Small deviations of stable processes via metric entropy.* J. Theor. Probab., to appear.

[MS1] V. D. Milman, S. J. Szarek, *A geometric lemma and duality of entropy numbers*, Geometric aspects of functional analysis (1996–2000), Lecture Notes in Math., vol. 1745, Springer, Berlin-New York (2000) 191–222.

[MS2] V. D. Milman, S. J. Szarek, *A geometric approach to duality of metric entropy*, C. R. Acad. Sci. Paris Ser. I Math. 332 (2001) no. 2, 157–162.

[PT] A. Pajor, N. Tomczak-Jaegermann, *Volume ratio and other s-numbers of operators related to local properties of Banach spaces*, J. Func. Anal. 87 (2) (1989), 273–293.

[P] A. Pietsch, Theorie der Operatorenideale (Zusammenfassung), Friedrich-Schiller-Universität Jena, 1972.

[Pi1] G. Pisier, *A new approach to several results of V. Milman*, J. Reine Angew. Math. 393 (1989), 115-131.

[Pi2] G. Pisier, *The Volume of Convex Bodies and Banach Space Geometry*, Cambridge Tracts in Mathematics, 94, Cambridge University Press, Cambridge (1989).

[S] C. Schütt, *Entropy numbers of diagonal operators between symmetric Banach spaces.* J. Approx. Theory 40 (1984), no. 2, 121–128.

[T] 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), no. 7, 299–301.

S. Artstein, V. D. Milman
School of Mathematical Sciences, Tel Aviv University,
Tel Aviv 69978, Israel
E-mail: artst@post.tau.ac.il
E-mail: milman@post.tau.ac.il

S. J. Szarek
Equipe d'Analyse Fonctionnelle, B.C. 186, Université Paris VI,
4 Place Jussieu, F-75252 Paris, France
and
Department of Mathematics, Case Western Reserve University,
Cleveland, OH 44106-7058, U.S.A.
E-mail: szarek@ccr.jussieu.fr