Entropy extension*

A. E. Litvak V. D. Milman † A. Pajor
N. Tomczak-Jaegermann ‡

Dedication: *The paper is dedicated to the memory of an outstanding analyst B. Ya. Levin. The second named author would like to note that he was a student of B. Ya. Levin whose scientific integrity and honesty of his teacher have accompanied him all his life.*

Abstract

We prove "entropy extension-lifting theorem". It consists of two inequalities for covering numbers of two symmetric convex bodies. The first inequality, that can be called "entropy extension theorem", provides estimates in terms of entropy of sections and should be compared with the extension property of ℓ_∞. The second one, which can be called "entropy lifting theorem", provides estimates in terms of entropies of projections.

1 Introduction

One of important consequences of the Hahn-Banach theorem is so called "extension property of ℓ_∞". It states that given normed space X and a subspace $Y \subset X$ every linear operator $S : Y \to \ell_\infty$ can be extended to an operator $T : X \to \ell_\infty$ having the same norm as S. This theorem is used in proofs of many results of Banach space theory and related fields. In particular, it was one of ingredients of the following result on covering numbers, obtained recently in ([LPT]):

*Keywords: metric entropy, entropy extension, entropy lifting, entropy decomposition, covering numbers

†The research was partially supported by BSF grant.

‡This author holds the Canada Research Chair in Geometric Analysis.

Let $0 < a < r < A$ and $1 \leq k < n$. Let $K, L \subset \mathbb{R}^n$ be symmetric convex bodies, and let $K \subset AL$. Let $E \subset \mathbb{R}^n$ be k-codimensional subspace such that $K \cap E \subset aL$. Then

$$N(K, 2rL) \leq 2^k \left(\frac{A+r}{r-a}\right)^k .$$

Here, as usual, $N(K, L)$ denotes the covering number (see the definition below). In a sense, the latter result is a (weak) version of extension theorem for entropy: *if we control the norm of the identity operator (= the half of diameter of the unit ball) in a subspace then we control the entropy in the entire space.* Note that if $K \cap E \subset aL$ then trivially $N(K \cap E, aL) \leq 1$. However, why should the diameter play such a crucial role? Can one achieve a similar control of entropy in the whole space from the knowledge of the entropy (rather than the diameter) in a subspace? The intuition does not support such a hope. However, quite surprisingly, this is possible. In the present paper we prove a strong version of an extension theorem for entropy: *if we control the entropy in a subspace then we control the entropy in the entire space*, see Theorem 3.1 below for the precise statement.

We also provide a variant of the inverse statement in Theorem 4.1 below. In the last section we discuss the non-symmetric case.

2 Notation and preliminaries

By a convex body we always mean a closed convex set with non-empty interior. By a symmetric convex body we mean centrally symmetric (with respect to the origin) convex body.

Let $K \subset \mathbb{R}^m$ be a convex body with the origin in its interior. We denote by $|K|$ the volume of K, and by K^0 the polar of K, i.e.

$$K^0 = \{x \mid \langle x, y\rangle \leq 1 \text{ for every } y \in K\}.$$

Let K, L be subsets of $\mathbb{R}^m$. We recall that the covering number $N(K, L)$ of K by L is defined as the minimal number N such that there exist vectors x_1, ..., x_N in $\mathbb{R}^m$ satisfying

$$K \subset \bigcup_{i=1}^{N} (x_i + L).$$

We use notation $N_A(K,L)$, if additionally $x_i \in A$, for $1 \le i \le N$ and $A \subset \mathbb{R}^m$; and we let $\bar{N}(K,L) = N_K(K,L)$.

For a symmetric convex body $K \subset \mathbb{R}^m$ and $\varepsilon \in (0,1)$, we shall need an upper estimate for the covering number $N(K, \varepsilon K)$. The standard estimate is

$$N(K, \varepsilon K) \le \bar{N}(K, \varepsilon K) \le (1 + 2/\varepsilon)^m, \tag{2.1}$$

which follows by comparing volumes and which would be sufficient for our results. However, when positions of centers is not important, we prefer to use here a more sophisticated estimate which follows from a more general result by Rogers-Zong ([RZ]), namely

$$N(K, \varepsilon K) \le \theta_m (1 + 1/\varepsilon)^m, \tag{2.2}$$

where

$$\theta_m \le \min\left\{2^m,\ m(\ln m + \ln(\ln m) + 5)\right\}.$$

In fact, from Rogers-Zong Lemma one gets that θ_m is bounded from above by so-called covering density of K (see [R1], [R2] for precise definitions and upper bounds), while the bound 2^m follows immediately from (2.1).

3 Entropy extension-lifting theorem

The main result of this paper is the following "entropy extension-lifting theorem". It consists of two inequalities for entropies. The first inequality relates the entropy of K and L to the entropy of sections of small codimension and can be called "entropy extension theorem", while the second one assumes an information on entropies of projections of a small corank and can be called "entropy lifting theorem".

Theorem 3.1 *Let $0 < a < r < A$. Let K, L be symmetric convex bodies in $\mathbb{R}^n$ such that $K \subset AL$. Let E be a subspace of $\mathbb{R}^n$ and $P : \mathbb{R}^n \to \mathbb{R}^n$ be a projection with $\ker P = E$.*

(i) *If $\operatorname{codim} E = k$, then*

$$N(K, rL) \le \theta_k \left(1 + \frac{A}{r-a}\right)^k \ N(K \cap E, \frac{a}{3}L).$$

(ii) *If* $\dim E = k$*, then*

$$N(K, rL) \le \theta_k \left(\frac{2A + r}{r - a} \right)^k \; N(PK, \frac{a}{2} PL).$$

Let us notice one particular case of this theorem, namely the case when $N(K \cap E, (a/3)L \cap E) = 1$ (resp. $N(PK, (a/2)\ PL) = 1$). Taking $b = a/3$, $R = r/3$ in the first part and $b = a/2$, $R = r/2$ in the second part, we immediately obtain the following consequence of Theorem 3.1 (the first part of which has been already mentioned in the Introduction).

Corollary 3.2 *Let* $0 < b < R < A$*. Let* K*,* L *be symmetric convex bodies in* $\mathbb{R}^n$ *such that* $K \subset AL$*. Let* E *be a subspace of* $\mathbb{R}^n$ *and* $P : \mathbb{R}^n \to \mathbb{R}^n$ *be a projection with* $\ker P = E$*.*

(i) *If* $\operatorname{codim} E = k$ *and* $K \cap E \subset bL \cap E$ *then*

$$N(K, 3RL) \le \theta_k \left(1 + \frac{A}{3(R - b)} \right)^k$$

(ii) *If* $\dim E = k$ *and* $PK \subset bPL$ *then*

$$N(K, 2RL) \le \theta_k \left(\frac{A + R}{R - b} \right)^k .$$

This corollary was one of the main results on covering numbers in [LPT] (see Corollaries 1.6 and 1.7 there), which was essentially used in proofs of other results of [LPT] and of [LMPT]. Actually, our present work is inspired by this result.

Now we turn to the proof of Theorem 3.1. First we obtain a more general result estimating entropy of sets in terms of entropy of projections of these sets and entropy of sections of related (but a bit more complicated) sets, in a spirit of Rogers-Shephard lemma for volumes. We call it "entropy decomposition lemma". It will imply Theorem 3.1.

Theorem 3.3 *Let* K*,* L_1*, and* L_2 *be subsets of* $\mathbb{R}^n$*. Let* E *be a subspace of* $\mathbb{R}^n$ *and* $P : \mathbb{R}^n \to \mathbb{R}^n$ *be a projection with* $\ker P = E$*. Then*

$$N\left(K, L_1 + L_2\right) \le \bar{N}\left(PK, PL_1\right) \max_{z \in K} N\left((K - L_1 - z) \cap E, L_2\right)$$

$$\leq \bar{N}\left(PK, PL_1\right) \ N\left(\left(K-K-L_1\right)\cap E, L_2\right).$$

and

$$N\left(K, L_1+L_2\right) \leq N\left(PK, PL_1\right) \ \max_{z\in\mathbb{R}^n} N\left(\left(K-L_1-z\right)\cap E, L_2\right)$$

Proof: We prove the first estimate, the proof of the second one repeats the same lines with obvious modifications.

Set

$$N_1 := \bar{N}(PK, PL_1).$$

Then, by definition, there are z_1, ..., z_{N_1} with $z_i \in PK$ for $1 \leq i \leq N_1$, and such that

$$PK \subset \bigcup_{i=1}^{N_1}(z_i + PL_1).$$

For every $x \in K$ fix $i(x) \leq N_1$ and $y_x \in PL_1$ such that

$$Px = z_{i(x)} + y_x$$

(if more than one such $i(x)$ (or y_x) exists, choose any of them and fix in the further argument).

For $1 \leq i \leq N_1$ pick $\tilde{z}_i \in K$ such that $P\tilde{z}_i = z_i$, and for every $y \in PL_1$ pick $\tilde{y} \in L_1$ such that $P\tilde{y} = y$.

Now for every $x \in K$ define

$$v(x) = \tilde{z}_{i(x)} + \tilde{y}_x \in \tilde{z}_{i(x)} + L_1$$

and

$$w(x) = x - v(x) = x - \tilde{z}_{i(x)} - \tilde{y}_x.$$

Denote

$$T_i := K - L_1 - \tilde{z}_i, \qquad \text{for} \qquad i \leq N_1.$$

Then $w(x) \in T_{i(x)}$ for every $x \in K$. Note also that $w(x) \in E$ for every $x \in K$, since

$$Pw(x) = Px - Pv(x) = Px - z_{i(x)} - y_x = 0.$$

Thus $w(x) \in T_{i(x)} \cap E$ and

$$x = w(x) + v(x) \in T_{i(x)} \cap E + \tilde{z}_{i(x)} + L_1$$

for every $x \in K$. It implies

$$K \subset \bigcup_{i=1}^{N_1} \left(T_i \cap E + \tilde{z}_{i(x)} + L_1 \right).$$

Since for every $i \leq N_1$ we have

$$N\left(T_i \cap E, L_2\right) \leq \max_{z \in K} N\left((K - L_1 - z) \cap E, L_2\right),$$

the result follows. $\square$

Proof of Theorem 3.1: Let $\varepsilon := r - a$. To prove (i), first note that since $(1/A)K \subset L$, then $N(K, rL) \leq N(K, (\varepsilon/A)K + aL)$. Thus, using Theorem 3.3 with $L_1 = (\varepsilon/A)K$ and $L_2 = aL$ we get

$$N(K, rL) \leq \bar{N}\left(PK, \frac{\varepsilon}{A} PK\right) N\left(\left(2 + \frac{\varepsilon}{A}\right) K \cap E, aL\right).$$

Now, by estimate (2.2) the first factor is bounded by $\theta_k \left(1 + A/\varepsilon\right)^k$, while the second factor is less than or equal to $N(3K \cap E, aL) = N(K \cap E, (a/3)L)$. This concludes the proof of (i).

To prove (ii), we use Theorem 3.3 with $L_1 = aL$ and $L_2 = \varepsilon L$ to get

$$N(K, rL) \leq \bar{N}\left(PK, aPL\right)\, N\left((2K + aL) \cap E, \varepsilon L\right).$$

To estimate the first factor note a well-known general fact that for arbitrary sets K', L', with L' symmetric, we have $\bar{N}(K', L') \leq N(K', (1/2)L')$. For the second factor we use estimate (2.2) to get

$$N\left((2K + aL) \cap E, \varepsilon L\right) \leq N\left((2A + a)\, L \cap E, \varepsilon L\right) \leq \theta_k \left(\frac{2A + r}{r - a}\right)^k.$$

$\square$

4 Lower bounds for entropy

Here we prove a theorem which is in a sense inverse to Theorem 3.3.

Theorem 4.1 *Let $0 < t < 1$. Let K_1, K_2 be subsets of $\mathbb{R}^n$ and L_1, L_2 be symmetric convex bodies in $\mathbb{R}^n$. Let $P : \mathbb{R}^n \to \mathbb{R}^n$ be a projection and $E = \ker P$. Then*

$$N\Big(tK_1+(1-t)K_2, (tL_1)\cap((1-t)L_2)\Big) \geq \bar{N}\big(PK_1, 2PL_1\big)\,\bar{N}\big(K_2\cap E, 2L_2\cap E\big).$$

Let us note that taking $K_1 = K_2$ and, additionally, $L_1 = ((1-t)/t)L_2$, we have the following corollary.

Corollary 4.2 *Let $0 < t < 1$. Let K be a convex body in $\mathbb{R}^n$ and L, L_1, L_2 be symmetric convex bodies in $\mathbb{R}^n$. Let $P : \mathbb{R}^n \to \mathbb{R}^n$ be a projection and $E = \ker P$. Then*

$$N\Big(K, (tL_1)\cap((1-t)L_2)\Big) \geq \bar{N}\big(PK, 2PL_1\big)\,\bar{N}\big(K\cap E, 2L_2\cap E\big).$$

and

$$N(K, L) \geq \bar{N}(tPK, 2PL)\,\bar{N}((1-t)K\cap E, 2L\cap E).$$

In the proof we will use the notion of packing numbers. Recall that for K and L in $\mathbb{R}^n$ the packing number $P(K, L)$ of K by L is defined as the maximal number M such that there exist vectors x_1, ..., $x_M \in K$ satisfying

$$(x_i + L)\cap(x_j + L) = \emptyset \qquad \text{for every} \qquad i \neq j.$$

In other words, $x_i - x_j \notin L_0 := L - L$. Such set of points we also call L_0-separated set. It is well known (and easy to check) that if L is symmetric convex body (so $L - L = 2L$) then

$$\bar{N}(K, 2L) \leq P(K, L) \leq N(K, L).$$

Proof of Theorem 4.1: Let $N_1 = P(PK_1, PL_1) \geq \bar{N}(PK_1, 2PL_1)$. Then there exist $z_1, \ldots, z_{N_1} \in PK_1$ such that $z_i - z_j \notin 2PL_1$ whenever $i \neq j$. For $1 \leq i \leq N_1$ pick $\tilde{z}_i \in K_1$ such that $P\tilde{z}_i = z_i$.

Let $N_2 = P(K_2\cap E, L_2\cap E) \geq \bar{N}(K_2\cap E, 2L_2\cap E)$. Then there exist $w_1, \ldots, w_{N_2}$ in $K_2\cap E$ such that $w_k - w_\ell \notin 2L_2$ if $k \neq \ell$.

For every $i \leq N_1$ and $k \leq N_2$ denote $x_{i,k} := t\tilde{z}_i + (1-t)w_k$ and consider the set

$$\mathcal{A} = \{x_{i,k}\}_{i\leq N_1,\ k\leq N_2} \subset tK_1 + (1-t)K_2.$$

We claim that $x_{i,k} - x_{j,\ell} \notin (2tL_1) \cap (2(1-t)L_2)$ if the pair (i,k) is different from (j,ℓ). Indeed, if $i \neq j$ then $P(x_{i,k} - x_{j,\ell}) = t(z_i - z_j) \notin 2tPL_1$, and hence $x_{i,k} - x_{j,\ell} \notin 2tL_1$. If $i = j$ then $k \neq \ell$ and $x_{i,k} - x_{j,\ell} = (1-t)(w_k - w_\ell) \notin 2(1-t)L_2$. Thus $\mathcal{A}$ is $(2tL_1) \cap (2(1-t)L_2)$-separated, which implies

$$N\big(tK_1+(1-t)K_2, (tL_1)\cap((1-t)L_2)\big) \geq P\big(tK_1+(1-t)K_2, (tL_1)\cap((1-t)L_2)\big)$$
$$\geq N_1 N_2 \geq \bar{N}(PK_1, 2PL_1)\bar{N}(K_2 \cap E, 2L_2 \cap E).$$

It concludes the proof. $\square$

5 Additional observations

In this section we will extend to the case of non-symmetric bodies the theorem from [LPT], which was mentioned in the introduction and also as the first part of Corollary 3.2. To keep the present paper self-contained we will use formulation of Corollary 3.2.

First we extend it to the case when K is not symmetric body. We need the following simple lemma.

Lemma 5.1 *Let $a > 0$ and $1 \leq k \leq n$. Let K be a convex body in $\mathbb{R}^n$ and L be a symmetric convex body in $\mathbb{R}^n$. Let E be a k-codimensional subspace of $\mathbb{R}^n$. Assume that $2a$ is the maximal diameter of $K \cap (E - z)$ over all choices of $z \in \mathbb{R}^n$, that is*

$$\forall x \in \mathbb{R}^n \quad \exists y \in E \quad \text{such that} \quad (x + K) \cap E - y \subset aL.$$

Then

$$(K - K) \cap E \subset 2aL.$$

Proof: Let $z \in (K - K) \cap E$. Then $z = v - w$, where v, $w \in K$. Write $v = v_1 + v_2$ and $w = w_1 + w_2$, where v_1, $w_1 \in E^\perp$ and v_2, $w_2 \in E$. Since $z \in E$, we have $v_1 = w_1$.

By the conditions of the lemma there exists $y \in E$ such that

$$(K - v_1) \cap E - y \subset aL.$$

Therefore $v_2 - y \subset aL$ and $w_2 - y \subset aL$, which implies

$$z = v - w = (v_2 - y) - (w_2 - y) \subset 2aL.$$

$\square$

Combining Lemma 5.1 and Corollary 3.2 (applied to $K-K$) we immediately obtain the following theorem.

Theorem 5.2 *Let $0 < a < A$ and $1 \le k \le n$. Let K be a convex body in $\mathbb{R}^n$ and L be a symmetric convex body in $\mathbb{R}^n$ such that $K \subset AL$. Let E be a k-codimensional subspace of $\mathbb{R}^n$. Assume that $2a$ is the maximal diameter of $K \cap (E-x)$ over all choices of $x \in \mathbb{R}^n$. Then for every $r > 2a$ one has*

$$N(K-K, 3rL) \le \theta_k \left(1 + \frac{2A}{3(r-2a)}\right)^k .$$

Now we consider the case when K is symmetric and L is not. First note that in this case the conclusion of Corollary 3.2 holds if we substitute L with $L \cap -L$. Indeed, if $K = -K$ is such that $K \subset RL$ and $K \cap E \subset aL \cap E$ then $-K \subset RL$ and $-K \cap E \subset aL \cap E$, which implies $K \subset R(L \cap -L)$ and $K \cap E \subset a(L \cap -L) \cap E$. Therefore, optimizing over all shifts of L, i.e. over all choices of center of L, we can extend Corollary 3.2 in the following way.

Theorem 5.3 *Let $0 < a < A$ and $1 \le k \le n$. Let K be a symmetric convex body in $\mathbb{R}^n$ and L be a convex body in $\mathbb{R}^n$. Let E be a k-codimensional subspace of $\mathbb{R}^n$. Assume that there exists $z \in \mathbb{R}^n$ satisfying*

$$K \subset A(L-z) \qquad \text{and} \qquad K \cap E \subset a(L-z).$$

Then for every $r > 0$ one has

$$N(K, 2r\bar{L}) \le \theta_k \left(1 + \frac{A}{3(r-a)}\right)^k ,$$

where $\bar{L} = (L-z) \cap (-L+z)$.

References

[LPT] A. E. Litvak, A. Pajor and N. Tomczak-Jaegermann, *Diameters of Sections and Coverings of Convex Bodies*, J. of Funct. Anal., 231 (2006), 438–457.

[LMPT] A. E. Litvak, V. D. Milman, A. Pajor and N. Tomczak-Jaegermann, *On the Euclidean metric entropy of convex bodies*, GAFA, Lecture Notes in Math., to appear.

[R1] C. A. Rogers, *A note on coverings*, Mathematica, 4 (1957), 1-6.

[R2] C. A. Rogers, Packing and Covering, Cambridge Tracts in Mathematics and Mathematical Physics., No. 54, Cambridge: University Press 1964.

[RZ] C. A. Rogers, C. Zong, *Covering convex bodies by translates of convex bodies*, Mathematica, 44 (1997), 215-218.

A. E. Litvak and N. Tomczak-Jaegermann,
Department of Mathematical and Statistical Sciences,
University of Alberta,
Edmonton, Alberta, Canada T6G 2G1,
alexandr@math.ualberta.ca
nicole.tomczak@ualberta.ca

V. D. Milman
Department of Mathematics,
Raymond and Beverly Sackler Faculty of Exact Sciences
Tel Aviv University, Tel Aviv, Israel
milman@math.tau.ac.il

A. Pajor,
Équipe d'Analyse et Mathématiques Appliquées,
Université de Marne-la-Vallée,
5, boulevard Descartes, Champs sur Marne,
77454 Marne-la-Vallée, Cedex 2, France
Alain.Pajor@univ-mlv.fr