ESI The Erwin Schrödinger International Institute for Mathematical Physics

Boltzmanngasse 9
A-1090 Wien, Austria

A Geometric Lemma and Duality of Entropy Numbers

V. D. Milman
S. J. Szarek

Vienna, Preprint ESI 687 (1999) April 2, 1999

Supported by Federal Ministry of Science and Transport, Austria
Available via http://www.esi.ac.at

A Geometric Lemma and Duality of Entropy Numbers

V. D. Milman (Tel Aviv)
S. J. Szarek (Cleveland & Paris)

1 Introduction

We shall study in this note the following conjecture, to which we shall refer as the "Geometric Lemma"; we state it first in a somewhat imprecise form.

Let n, N *be positive integers with* $k := \log N \ll n$. *If* $S \subset I\!R^n$ *is a finite set whose cardinality doesn't exceed* N *and such that its convex hull* $K := \operatorname{conv} S$ *admits an equally small Euclidean* 1*-net (i.e.,* K *can be covered by no more than* N *translates of the unit Euclidean ball* D*), then* $\frac{1}{2}D \not\subset K$.

At the first glance a statement of that nature may appear "trivial". Indeed, on some meta-mathematical level, we are asking whether a Euclidean ball may be described by less than exponential in n "bits of information" (by $\ll$ we mean above "much smaller than") . And our intuition says "NO", no matter what exactly those "bits" are supposed to mean. However, the more exact formulation brings together two very different assumptions: small number of vertices of the polytope K and small cardinality of the Euclidean "net" for K. As we shall see in this note, these data are not easy to "combine"; that creates difficulties and, at the same time, interest.

Our interest in the Lemma came as an outgrowth of the study of the following problem, which originally has been promoted by Carl and Pietsch and is well known in Geometric Functional Analysis (and in the Geometric Operator Theory), usually referred to as "the duality conjecture for entropy numbers of operators". Recall that if X, Y are Banach spaces, $u : X \to Y$ a compact operator and $\varepsilon > 0$, we denote by $N(u, \varepsilon)$ the minimal cardinality of an ε-net of the image $u(B_X)$ of the unit ball B_X of X (in the metric of Y; in other words, $N(u,\varepsilon)$ is the smallest number of balls in Y of radius ε that together cover $u(B_X)$). The conjecture referred to above asserts that there exist universal constants $a, b > 0$ such that

$$a^{-1} \log N(u, b\varepsilon) \leq \log N(u^*, \varepsilon) \leq a \log N(u, b^{-1}\varepsilon) \tag{1}$$

for any compact operator u and $\varepsilon > 0$. (Here and in what follows, all logarithms are to the base 2.) In terms of the so-called *entropy numbers* (for basic results concerning them and related concepts see [3] and [19]), defined for an operator u by

$$e_k(u) := \inf\{\varepsilon > 0 : N(u,\varepsilon) \leq 2^k\},$$

the assertion of the conjecture (roughly) becomes

$$b^{-1} e_{ak}(u) \leq e_k(u^*) \leq b e_{k/a}(u), \tag{2}$$

which is its more standard formulation. (For the most part we shall be working with the form (1).) Note that even though ak and k/a are not necessarily integers, (2) makes sense as the definition of $e_k(u)$ works also for noninteger k. Additionally, due to the asymptotic nature of the questions we investigate, the difference between the number and its integer part (or the nearest integer, or, usually, even its double) is immaterial, and so we shall pretend that all numbers are integers as necessary; see also Remark 8.3.

We point out here that, for an operator u, the sequence $(e_k(u))$ tends to 0 as $n \to \infty$ if and only if $(e_k(u^*))$ does (and if and only if u is compact). In such a weak sense the qualitative character of the two sequences is the same. The "Duality Conjecture", if true, would imply that the two sequences are equivalent "distributionally", i.e., in the strongest *quantitative* sense that we may reasonably expect. To the best of our knowledge, it is still not known whether (1) or (2) has a chance to hold with $a = 1$, as inquired originally by Pietsch (for sure not with $a = b = 1$); nevertheless, the above formulation seems to be the most natural one. See also Remark 4.2.

Up to now, the Duality Conjecture in the form stated here has been verified only under very strong assumptions on *both* spaces (see, e.g., [7], [18]), or under some regularity assumptions on the sequences $(e_k(u))$, $(e_k(u^*))$ ([29], [2]). Our Geometric Lemma is relevant to the case when one of the spaces, say Y, is a Hilbert space, and X is arbitrary; more specifically, it implies then the second inequality in (1). However, we believe that once a proper argument (assuming there is any) is found, a sort of a "dual" version for the other inequality will be figured out and, possibly, one will also be able to relax the assumptions on Y (paralleling the developments of ideas in [29] and [2], where "weaker" – but quantitative – equivalences were established).

There has been a substantial body of work on the Duality Conjecture. Rather than present the current state of the knowledge on the matter, we refer the reader to [2], [18], [21], their references and the survey [6] (in preparation).

As we indicated, the Geometric Lemma remains a conjecture. We did check it for various (classes of) convex bodies, including ℓ_p^n-balls of various radii (cf. one of the comments preceding Proposition 7.2) and some "random bodies" (specifically, "generic" projections of ℓ_1^N-balls – a "canonical" counterexample to many problems in high dimensional convexity, cf. [12]). In the general setting, we were able to prove it "up to a logarithmic factor". In particular, we did show in Theorem 9.3 :

The assertion of the Geometric Lemma holds if we replace the assumption $k \ll n$ *by* $k \leq c(1+\log n)^{-6} n$, *where* $c > 0$ *is some universal numerical constant.*

The corresponding "entropy duality" result is Theorem 9.4 :

There are numerical constants $a, C > 0$ *such that if* u *is a compact Hilbert space valued operator and* $k \in I\!N$, *then* $e_{ak}(u^*) \leq C(1+\log k)^3 e_k(u)$.

Thus, even though the principal appeal of the Geometric Lemma stems from its potential to resolve the Duality Conjecture in the relevant case, the partial solution indicated above already has nontrivial consequences. To the best of our knowledge, in absence of strong assumptions on *both* spaces or on "regularity" of the sequences, no result of the above type (i.e., with a factor, which is a function of k) appears in the literature (see also the comments following Theorem 9.4).

Finally, let us mention that Theorem 9.3 and, overall, an important part of the discussion deal with formally stronger statements: upper estimates on the mean width of the part of the body in question that is inside the Euclidean ball D. Theorem 9.3 asserts, in particular :

If K *is as in the Geometric Lemma, then the mean width of* $K \cap D$ *does not exceed* $C(1+\log k)^3 \sqrt{k/n}$, *where* C *is a universal numerical constant.*

2 Geometric Lemma – precise statements

In this section we shall state a geometric version of (the case of) the Duality Conjecture that is implied by the Geometric Lemma. We shall also give precise formulation(s) of the lemma and indicate the relations between various statements. Let us start with the Duality Conjecture, the relevant case of which is

If X is a Banach space, H – a Hilbert space, and $u : X \to H$ a compact operator, then, for any $\varepsilon > 0$, $\log N(u^*, b\varepsilon) \le a \log N(u, \varepsilon)$, where $a, b > 0$ are universal constants.

Note that one can restrict oneself to the case $\varepsilon = 1$ in the above (rescaling if necessary). Additionally, one may assume (by approximation) that the spaces in question are finite dimensional, and the operator u is one-to-one. In this context it is convenient to define the *covering number* $N(U, V)$ for U, V-subsets of $\mathbb{R}^n$ (say, with V a closed convex body) by

$$N(U, V) := \min\{N : U \textit{ may be covered by } N \textit{ translates of } V\}.$$

If we set $U = uB_X$, $V = \varepsilon B_Y$, we get $N(U, V) = N(u, \varepsilon)$. Similarly $N(u^*, \varepsilon) = N(V^\circ, U^\circ)$, where K° denotes the polar body of K (say, with respect to the canonical Euclidean structure). Accordingly, denoting again by $D = D_n$ the standard Euclidean ball (we shall use this convention throughout this note), we see that the above assertion is equivalent to

Conjecture 2.1 *If $U \subset \mathbb{R}^n$ is a symmetric compact convex body, then* $\log N(D, bU^\circ) \le a \log N(U, D)$ *where, $a, b > 0$ are universal constants.*

For future reference we point out that if V is a Euclidean ball (or, more generally, an ellipsoid), the translates of V in the definition of $N(U, V)$ may be further assumed to have centers contained in U (as required – in the general case – by some authors). This makes them insensitive to enlarging the ambient space, while the present definition ensures that they are *always* increasing in the first argument.

Conjecture 2.1 is the one we are *really* going to work with. In the next several sections we shall show how it is implied by the following version of the Geometric Lemma.

Conjecture 2.2 *For any $\gamma > 0$ there exists $c_1 = c_1(\gamma) > 0$ such that if $K \subset I\!R^n$ verifies*

(i) $K = \mathrm{conv}\, S$, $\log |S| \leq c_1 n$ and

(ii) $\log N(K, D) \leq c_1 n$,

then $\gamma D \not\subset K$.

Remark 2.3. In fact, Conjecture 2.1 (and Conjecture 2.4 below, see Proposition 5.1 and the comments preceding it) can be derived from the validity of Conjecture 2.2 for just one $\gamma < 1$, the constants involved (a, b, resp. C) depending on that γ and the corresponding c_1. □

It is possible to derive from Conjecture 2.2 a formally stronger statement, for which we need to introduce some notation. First, if $U \subset I\!R^n$ is a compact symmetric convex body containing the origin in the interior, one denotes by $\| \cdot \|_U$ its Minkowski functional, i.e. the norm, for which U is the unit ball. We shall use the same notation for gauges of nonsymmetric sets. We set also

$$M^*(U) := \int_{S^{n-1}} \sup_{y \in U} \langle x, y \rangle \, d\mu_n(x) = \int_{S^{n-1}} \|x\|_{U^\circ} \, d\mu_n(x)$$

where μ_n is the normalized (i.e., probability) Lebesgue measure on S^{n-1}. The second equality – without the integration signs – is in fact a definition of the polar U°. Both quantities on the right define $M^*(\cdot)$ also for nonsymmetric sets, and the first one even for nonconvex and not containing the origin (we will need that degree of generality later). We mention in passing that $M^*(U)$ equals the mean half-width of U, a well known geometric parameter. For future reference, we also set

$$M(U) := \int_{S^{n-1}} \|x\|_U d\mu_n(x).$$

In the sequel we will seldom formulate explicitly the hypotheses (convexity, symmetry or origin in the interior) on the arguments of $M^*(\cdot)$ and $M(\cdot)$; unless *some* hypotheses are stated, we shall implicitly assume the bare minimum for the formulae to make sense.

We are now ready to state the stronger version of Conjecture 2.2.

Conjecture 2.4 *Let $K = \operatorname{conv} S \subset I\!R^n$ and set $k = \max\{\log|S|, \log N(K, D)\}$. Then*

$$M^*(K \cap D) \le C\sqrt{\frac{k}{n}}, \tag{3}$$

where C is a universal constant.

Clearly Conjecture 2.4 $\Rightarrow$ Conjecture 2.2: if $\gamma D \subset K$, then $M^*(K \cap D) \ge M^*(\gamma D) = \gamma$, which is inconsistent with (3) if k/n is small.

In the next section we introduce more notation and recall some (more or less) known results about $M^*(\cdot)$, a few of them quite deep. In section 4 we prove the implications Conjecture 2.4 $\Rightarrow$ Conjecture 2.1 and Conjecture 2.2 $\Rightarrow$ Conjecture 2.4. (We do not know whether Conjecture 2.1 implies *formally* Conjecture 2.2, that is, whether the Geometric Lemma is in fact *equivalent* to "our" case of the Duality Conjecture. We devoted that matter some, but not too much, thought; the implication in question could be imaginably useful for constructing a possible counterexample to the Duality Conjecture.) Then, in section 5, we present some "almost isometric" refinements of arguments from section 4, in particular the one suggested in Remark 2.3. In sections 6 through 8 of we develop the ideas and tools needed for our study of the conjectures. That study culminates in section 9, where we formulate and prove the results sketched at the end of the introduction.

3 More notation, known results

As general references for notation and basic results of local theory of Banach spaces we suggest the books [16], [22] and [30] or the survey paper [5]; a handy source for "probability in Banach spaces" is the monograph [10].

For a subspace $E \subset I\!R^n$, we shall denote by P_E the orthogonal projection onto E. Given positive integer $k < n$, $G_{k,n}$ is the Grassmann manifold (of k-dimensional subspaces of R^n) endowed with the canonical rotationally invariant probability measure $\mathcal{P} = \mathcal{P}_{k,n}$. We shall say that a "generic k-dimensional subspace" E of $I\!R^n$ has certain property if the measure of E's with that property is close to 1 provided n is large (or k, n are large, depending on the context). Similarly, we shall talk about "generic rank k orthogonal projections".

The first auxiliary result is the following "relative" of the Dvoretzky Theorem (see, e.g., [14], Lemma 2.1 and its proof). Such results are usually stated for the symmetric case, but the arguments, based on the "concentration of measure phenomenon", do not *really* require that hypothesis (nor even, in our formulation, that the body in question contains 0, the assertion being insensitive to a "small" translation of the body, cf. Fact 3.2). Let us also remark here that even though in the context of the Duality Conjecture, as usually stated, only the case of symmetric sets is relevant, in the present note we drop the symmetry assumption whenever feasible. We believe that the setting of *general* convex bodies is the natural one, part of the motivation for their study coming from geometry and optimization, where the symmetry hypothesis is often artificial. Additionally, in our arguments nonsymmetric sets actually *have to* appear (e.g., as intersections of two symmetric sets with different centers). On the other hand, lack of symmetry is seldom a source of difficulties, and often – but not always – the passage to the more general framework is merely "formal". Such is, for example, the case of our Conjectures 2.2 and 2.4 from the preceding section: for a possibly nonsymmetric K, apply the statement to $(K-K)/2$ (noting that, e.g., $\log N((K-K)/2, D) \leq 2 \log N(K, D)$ while $M^*(K \cap D) \leq M^*((K-K)/2 \cap D)$; there is also an inequality sort of inverse to the latter for the sets we are interested in, see Remark 8.3). Similarly, if K is symmetric and S is its net, one can always assume that S is also symmetric and contains 0: replace S by $S \cup (-S) \cup \{0\}$, the resulting roughly two-fold increase in the cardinality of S is nonessential in our setting (see Remark 8.3).

Fact 3.1 *Let $V \subset D \subset \mathbb{R}^n$ be a convex set and let $m \leq n$. Let P be a generic orthogonal projection of rank m. Then*

(a) if $\sqrt{\frac{m}{n}} \geq \varepsilon_0 M^*(V)$, then $PV \subset C\sqrt{\frac{m}{n}} PD$

(b) if $\sqrt{\frac{m}{n}} \leq \varepsilon_0 M^*(V)$, then $cM^*(V)PD \subset PV \subset CM^*(V)PD$

(c) moreover, if for some $\varepsilon \leq \varepsilon_0$ we have $\sqrt{\frac{m}{n}} \leq \varepsilon M^*(V)$, then

$$(1 - C\varepsilon)M^*(V)PD \subset PV \subset (1 + C\varepsilon)M^*(V)PD.$$

Above, c, C and ε_0 are universal positive constants, independent of n, V and ε. In all cases "generic" means "except on a set (of projections) of measure $\leq \exp(-c'm)$", where $c' > 0$ is a universal numerical constant.

The next result is closely related to the fact that, for a fixed rank m projection P on $\mathbb{R}^n$, the Euclidean norm of Px is "strongly concentrated" around the value $\sqrt{\frac{m}{n}}$ as x varies over S^{n-1} (clearly, the *average* of $|Px|^2$ *equals* m/n; here and throughout the paper we use the notation $|\cdot|$ for the Euclidean norm $\|\cdot\|_D$). This well-known phenomenon has been often derived from the isoperimetric inequality for the sphere, but it can be approached also via a direct calculation. Here we choose an equivalent point of view : the point x stays fixed, while the projection P varies over $G_{m,n}$ endowed with the probability measure $\mathcal{P} = \mathcal{P}_{m,n}$.

Fact 3.2 *Let $x \in S^{n-1}$, let $m \leq n$ and let P be a generic orthogonal projection of rank m (i.e., considered as an element of $(G_{m,n}\,, \mathcal{P})$). Then $|Px|$ is strongly concentrated around the value $\sqrt{\frac{m}{n}}$. More precisely,*

(a) if $\varepsilon > 0$, then

$$\mathcal{P}\,(|\,|Px| - \,\sigma_{m,n}| > \varepsilon) \;<\; \exp\,(-\varepsilon^2 n/2),$$

where $\sigma_{m,n}$ is the median of $|Px|$ and $\left|\sigma_{m,n} - \sqrt{\frac{m}{n}}\right| \leq \frac{C}{\sqrt{n}}$
(b) consequently, if $\lambda > 1$, then

$$\mathcal{P}\left(|Px| > \lambda\sqrt{\frac{m}{n}}\right) \;<\; \exp\,(-c(\lambda-1)^2 m)$$

(c) and, additionally, for $\alpha > 0$,

$$\mathcal{P}\left(|Px| < \alpha\sqrt{\frac{m}{n}}\right) < (\sqrt{e}\alpha)^m\ .$$

Above, C and c are universal positive constants.

Part (a) of Fact 3.2 is just the isoperimetric inequality applied to the function $x \to |Px|$ (cf. [8], where it was employed in the spirit close to that of our paper). Part (c), better known in the case of the Gaussian measure (see, for example, [25]), can be recovered, e.g., from Lemma 6 in [15].

From Facts 3.1 and 3.2 we derive the following Milman-Pajor-Tomczak-Talagrand (cf. [13], [17], [28]) type result.

Proposition 3.3 *Let $k \leq n$, let $A > 0$ and let $K \subset \mathbb{R}^n$ be a symmetric convex body with* $\log N(K, D) \leq Ak$. *Set* $\omega := \max\{M^*(K \cap D), \sqrt{\frac{k}{n}}\}$. *Then, for a generic rank k orthogonal projection P, we have*

$$c_0\sqrt{\frac{k}{n}}\,|x| \leq \max\{\omega\|x\|_K, |Px|\} \quad \text{for all} \quad x \in \mathbb{R}^n, \tag{4}$$

where $c_0 > 0$ is a constant depending only on A. The assertion holds also for not-necessarily-symmetric bodies $K \ni 0$ after one replaces $M^*(K \cap D)$ by $\max_{x \in \mathbb{R}^n} M^*((K - x) \cap D)$ *in the definition of ω.*

Proof. For a smoother exposition we provide first a detailed proof in the (central) symmetric case and then sketch modifications needed to handle the general setting.

Noting that $\max\{\|\cdot\|_U, \|\cdot\|_V\} = \|\cdot\|_{U \cap V}$ and rescaling, we see that (4) is equivalent to

$$K \cap \omega\, P^{-1} D \;\subset\; c_0^{-1}\omega\,\sqrt{\frac{n}{k}}\,D$$

or

$$x \in K \setminus c_0^{-1}\omega\,\sqrt{\frac{n}{k}}\,D \;\Rightarrow\; |Px| > \omega. \tag{5}$$

Let S be a set with $|S| \leq 2^{Ak}$ such that $K \subset S + D$. A standard argument shows that then in fact $K \subset S + (D \cap 2K)$ (it is here that the symmetry is used; in general we would have $K - K$ in place of $2K$). Moreover, if $S_1 = S \setminus (c_0^{-1}\omega\sqrt{\frac{n}{k}} - 1)D$, then

$$K \setminus c_0^{-1}\omega\,\sqrt{\frac{n}{k}} D \;\subset\; S_1 + (D \cap 2K).$$

Accordingly, to prove (5), hence (4), it suffices to show that, for a generic P, $|P(s + y)| > \omega$ *simultaneously* for all $s \in S_1$ and all $y \in D \cap 2K$. To that end observe that, first, by Fact 3.1(a) or (b), $|Py| \leq 2C\omega$ for a generic P and *all* $y \in D \cap 2K$. On the other hand, by Fact 3.2(c), for any fixed $x \in \mathbb{R}^n$,

$$\mathcal{P}\left(|Px| \leq \delta\,\sqrt{\frac{k}{n}}\,|x|\right) \;<\; (\sqrt{e}\delta)^k$$

for any $\delta > 0$. Choosing δ small enough (say, $\delta = (\sqrt{e}2^{A+1})^{-1}$) we get that, for a generic projection P, *all* $s \in S_1$ (note $|S_1| \leq 2^{Ak}$) verify $|Ps| \geq \delta(c_0^{-1}\omega\sqrt{\frac{n}{k}} - 1)\sqrt{\frac{k}{n}} > \delta(c_0^{-1}\omega - 1)$, and so, for $x \in S_1 + (D \cap 2K)$,

$$|Px| \geq (\delta(c_0^{-1}\omega - 1) - 2C\omega)$$

which yields (5) if c_0 is chosen small enough.

If K is not symmetric, a more careful look shows that, in fact, one needs to control *simultaneously* (generic) projections of $K \cap (D + s)$ for all $s \in S_1$ or, equivalently, the projection of $W := \bigcup_{s \in S_1}(K - s) \cap D$. The argument used in the symmetric case carries over directly if A is small (specifically, if $A < c'$, where c' is the constant from Fact 3.1; cf. the proof of Proposition 5.1). For general A, it is more efficient to estimate $M^*(W)$ by $\max_{x \in I\!R^n} M^*((K-x) \cap D) + C_0\sqrt{\frac{k}{n}}$ via Lemma 8.1 and then "pipe in" $\mathrm{conv}\, W$ in place of $D \cap 2K$ in the argument above (the reader will readily verify that Lemma 8.1 depends only on Fact 3.8 and, morever, if all y_j's are 0 – the case which is relevant here – is independent from the rest of the paper). □

We will need an estimate of "covering numbers" known as "Sudakov's inequality".

Fact 3.4 ([10], Theorem 3.18) *If $U \subset I\!R^n$ and $\varepsilon > 0$, then*

$$\log N(U, \varepsilon D) \leq C(\frac{M^*(U)}{\varepsilon})^2 n$$

where C is a universal constant.

Recall also that the problem of duality of entropy numbers (say, in the form (2)) is solved for $k \geq \mathrm{rank}\, u$ ([9], see also [21]). We have

Fact 3.5 *Let $U, V \subset I\!R^n$ be convex bodies such that $U \ni 0$ and V is 0-symmetric. If $k \geq n$, then*

$$\log N(U, V) \leq k \Rightarrow \log N(V^\circ, \beta U^\circ) \leq \alpha k$$

for some universal constants $\alpha, \beta \geq 1$ (resp. $\alpha = \alpha(\tau)$ if we just assume that $k \geq \tau n$ for some $\tau \in (0, 1)$). Moreover, the above inequality holds – at least if U is also symmetric – with $\beta = \beta(\alpha)$, **for any** *$\alpha > 1$ (resp. $\beta = \beta(\alpha, \tau)$, for $\alpha > 1$ and $\tau \in (0, 1)$).*

The Fact was stated in [9] just in the case when *both* U and V are symmetric, but the present variant follows formally: just apply the symmetric version to $(U-U)/2$ (and $2k$ in place of k) and note that $((U-U)/2)^\circ \subset 2U^\circ$. We do not know whether the symmetry of U is needed in the last statement; in absence of that hypothesis the present argument yields $\alpha > 2$ in place of $\alpha > 1$. Let us also note that, at least for the first statement and with proper care, one may dispose of the symmetry assumptions altogether (see [15]).

We shall need a few more properties of the functional $M^*(\cdot)$.

Fact 3.6 *If $U, V \subset \mathbb{R}^n$ are convex sets, then the function defined on $\mathbb{R}^n$ by*

$$\phi(x) = M^*((x+U)\cap V)$$

is concave. In particular, if both U and V are 0-symmetric, then $\phi(x) \le \phi(0)$ for $x \in \mathbb{R}^n$.

This follows from the facts that, under Minkowski addition, the set-valued map $x \to (x+U)\cap V$ is concave, while $M^*(\cdot)$ is additive and positively homogeneous.

Let γ_n be the standard Gaussian measure on $\mathbb{R}^n$ (i.e., the one with density $(2\pi)^{-n/2}\, e^{-|x|^2/2}$). The next result describes the very well known relationship between spherical averages and those with respect to γ_n, and is easily established by integrating the latter in spherical coordinates.

Fact 3.7 *For $U \subset \mathbb{R}^n$, the Gaussian average*

$$\ell_1(U) := \int_{\mathbb{R}^n} \|x\|_U \, d\gamma_n(x)$$

is "essentially the same as" $n^{1/2}M(U)$. More precisely, there are constants $\sigma_n < 1$ with $\sigma_n \to 1$ as $n \to \infty$ such that, for all U as above, $\ell_1(U) = \sigma_n n^{1/2} M(U)$. The same is true if we replace $\ell_1(U)$ by $\ell_1(U^\circ)$ and $M(U)$ by $M^*(U)$.

For the record, $\sigma_n = \sqrt{2}\Gamma(\frac{n}{2})/\Gamma(\frac{n}{2}+1) \in (\sqrt{1-\frac{1}{n}}, 1)$. We could have dispensed with σ_n's in Fact 3.7 if we had defined *both* the Gaussian average and $M(U)$ via second moments of $\|\cdot\|_U$, a rather insignificant modification by the Kahane-Khinchine inequality (see [1], Lemma 3.3, for the nonsymmetric case). However, that would not conform to the standard terminology.

Finally, we mention the following well known

Fact 3.8 *If $S \subset D \subset I\!\!R^n$ is a finite set, then*

$$M^*(\operatorname{conv} S) = M^*(S) \leq C\sqrt{\frac{\log|S|}{n}},$$

where C is a numerical constant.

Fact 3.8 is proved most easily by passing to the Gaussian average (Fact 3.7) and a direct computation using tail estimates for the Gaussian density; in the Gaussian setting it is a special case of a much more general phenomenon (see [10], (3.6) or our Lemma 8.1). Alternatively, it is implicit in our Fact 3.1. By Fact 3.4, the estimate is exact if the set S is uniformly separated. We recall that a set is called δ-separated if each two its different members are more than δ apart; this leads to the concept of *packing numbers* – equivalent, up to a factor of 2 (in the argument), to that of covering numbers.

In the sequel we shall occasionally write $\Phi \lesssim \Psi$ meaning that there exists a universal numerical constant C such that, for all values of the parameters involved in the definitions of (normally nonnegative) quantities Φ and Ψ, one has $\Phi \leq C\Psi$. E.g., the assertion of Fact 3.8 can be written as $M^*(S) \lesssim \sqrt{\frac{\log|S|}{n}}$. (We point out that this convention differs from the one employed often in, e.g., combinatorics, and using for that concept the symbol $\ll$, reserved in this note for "much smaller than" or "sufficiently smaller than".) Similarly, $\Phi \simeq \Psi$ will indicate two-sided estimates $C^{-1}\Psi \leq \Phi \leq C\Psi$. We will not use that convention when we want to make the dependence on other constants or parameters explicit. Unless stated otherwise, C, c, C_1, c' etc. will stand for numerical constants independent of the dimension or any other parameters, whose exact values *may* vary between occurrences.

4 The implications 2.4 $\Rightarrow$ 2.1 & 2.2 $\Rightarrow$ 2.4

The implication Conjecture 2.4 $\Rightarrow$ Conjecture 2.1. We shall assume the validity of Conjecture 2.4 and show how the results of the preceding section imply then Conjecture 2.1. More precisely, we prove

Proposition 4.1 *Let $k \in I\!\!N$. Let $w \geq 1$ be such that, for all n, S, K verifying the assumptions of Conjecture 2.4 for that particular k, one has*

$M^*(K \cap D) \le w\sqrt{\frac{k}{n}}$. *Then, for all* n *and for all convex sets* $U \subset I\!\!R^n$,

$$\log N(U, D) \le k \;\Rightarrow\; \log N(D, CwU^\circ) \le ak, \tag{6}$$

where $a, C > 0$ *are universal constants.*

Proof. Let $S \subset U$ with $|S| = N(U, D) \le 2^k$ be such that $S + D \supset U$. Denote $K = \mathrm{conv}\, S$. We first observe that, for $\rho > 0$,

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

In fact, any ρ-net of $D(= D^\circ)$ with respect to $\|\cdot\|_{K^\circ}$ is a $(\rho+2)$-net with respect to $\|\cdot\|_{U^\circ}$. To see that, observe that if $x, y \in D$ and $\|x - y\|_{K^\circ} \le \rho$, then

$$\begin{aligned} \|x - y\|_{U^\circ} &= \max_{u \in U} \langle x - y, u \rangle \\ &\le \max_{s \in S, z \in D} \langle x - y, s + z \rangle \\ &\le \|x - y\|_{K^\circ} + |x - y| \le \rho + 2. \end{aligned}$$

(We are being slightly careless here as, in principle, it is possible that $0 \notin K$ and so one can not really speak about $\|\cdot\|_{K^\circ}$. However, this is easily remedied by adding to S a single point, cf. Remark 8.3. Another potential difficulty, K being degenerate, is handled by passing to a lower dimension.)

To derive Conjecture 2.1, it is now enough to appropriately estimate $N(D, \rho K^\circ)$ by $N(K, D)$ for some $\rho \lesssim w$; notice that $N(K, D) \le N(U, D)$ (cf. the comment following Conjecture 2.1). To that end, apply Proposition 3.3 with the present choice of k, n, K (hence $A = 1$). Let $P = P_F$ be a (generic rank k) projection such that

$$c_0\sqrt{\frac{k}{n}}|\cdot| \le \max\{w\sqrt{\frac{k}{n}}\|\cdot\|_K, |P(\cdot)|\} \;.$$

After dividing out by $\sqrt{\frac{k}{n}}$, this dualizes to

$$c_0 D \subset \mathrm{conv}\,\{w\, K^\circ \cup \sqrt{\frac{n}{k}}\; (D \cap F)\} \;\; (\subset\; w\, K^\circ + \sqrt{\frac{n}{k}}(D \cap F)).$$

Hence

$$\begin{aligned} N(D,\rho K^\circ) &= N(c_0 D, c_0\rho K^\circ) \\ &\le N(wK^\circ + \sqrt{\frac{n}{k}}(D\cap F), c_0\rho K^\circ) \\ &\le N(\sqrt{\frac{n}{k}}(D\cap F), (c_0\rho - w)K^\circ) \end{aligned} \tag{7}$$

(where we tacitly assumed $c_0\rho - w > 0$). Observe that the polar of $K^\circ \cap F$ (inside the k-dimensional space F) is $P_F K$. Accordingly, if we knew that

$$\log N(P_F K, \beta^{-1}(c_0\rho - w)\sqrt{\frac{k}{n}} D) \le k, \tag{8}$$

we could conclude from Fact 3.5 that the last member of (7) is bounded by $N(K,D)^\alpha \le N(U,D)^\alpha \le 2^{\alpha k}$, as required (above α, β are the constants from Fact 3.5). We now argue as in the proof of Proposition 3.3. If K is symmetric, $K \subset S + D$ implies $K \subset S + (2K\cap D)$. Accordingly, $P_F K \subset P_F S + P_F(2K\cap D)$ (for any $P = P_F$), while, for a generic P_F, $P_F(2K\cap D) \subset 2Cw\sqrt{\frac{k}{n}}D$ by Fact 3.1(a) or (b) and so we get (8) as long as $2Cw \le \beta^{-1}(c_0\rho - w)$. In particular, $\rho = c_0^{-1}(2C\beta + 1)w \simeq w$ works, as required. In the general case (i.e., K not necessarily symmetric), $(2K\cap D)$ has to be replaced by $\bigcup_{s\in S}(K-s)\cap D$; cf. the end of the proof of Proposition 3.3. $\Box$

Remark 4.2. Assuming Conjecture 2.4 (or, by what follows, just Conjecture 2.2), the argument above yields (6), hence Conjecture 2.1, with $a = \alpha$, where α comes from Fact 3.5. In particular, we would obtain then the validity of the case of the Duality Conjecture stated at the beginning of section 2 for any $a > 1$, the price being paid in the magnitude of $b = b(a)$. We also emphasize the that the symmetry hypothesis in Conjecture 2.1 is not used (at least if one doesn't worry about the exact value of the constant a), we leave it there just "for historical reasons." (In any case, that hypothesis can be "disposed of" formally, see the comments following Fact 3.5.) $\Box$

The implication Conjecture 2.2 $\Rightarrow$ Conjecture 2.4. Let n, k, S, K be as in Conjecture 2.4. Assuming the validity of Conjecture 2.2, we must show that $\sqrt{\frac{n}{k}}M^*(K\cap D)$ can not be arbitrarily large. Accordingly, throughout the argument we may assume that that quantity is larger than an arbitrary preassigned numerical constant (as otherwise we would have been done). Let us denote $n_1 = (\varepsilon_0 M^*(K \cap D))^2 n$ (where ε_0 comes from Fact 3.1; as usual, we pretend that n_1 is an integer), then $\sqrt{\frac{n}{k}} M^*(K \cap D) = \varepsilon_0^{-1}\sqrt{\frac{n_1}{k}}$; clearly we may assume that n_1/k is "large". Apply Fact 3.1 with m replaced by n_1. This yields $K_0 = PK$, of which we may think to be contained in $\mathbb{R}^{n_1}$, such that (by the part (b) of the Fact) $K_0 \supset P(K \cap D_n) \supset c\sqrt{\frac{n_1}{n}} D_{n_1}$ while at the same time, by the part (a) of the Fact, $\log N(K_0, 2C\sqrt{\frac{n_1}{n}} D_{n_1}) \leq k$ (as in earlier arguments, we use here the equality $K = \bigcup_{s \in S} s + (K - s \cap D)$ and, in the symmetric case, the inclusion $(K - s) \cap D \subset 2K \cap D$, with appropriate modifications if K is not symmetric; see the end of the proof of Proposition 3.3). Now applying Conjecture 2.2 to $K_1 = (2C\sqrt{\frac{n_1}{n}})^{-1} K_0$ and $\gamma = c(2C)^{-1}$ (this can be done since the cardinality of the set of extreme points of K_1 doesn't exceed that of K) we see that we must have $k > c_1(\gamma) n_1$ or $\sqrt{\frac{n}{k}} M^*(K \cap D) = \varepsilon_0^{-1}\sqrt{\frac{n_1}{k}} < \varepsilon_0^{-1} c_1(\gamma)^{-1/2}$, as required. □

Remark 4.3. As was the case with the prior implication, the above argument is done "for fixed k", i.e., the validity od Conjecture 2.4 for given k, n is derived from the validity of Conjecture 2.2 for the same k and some other n. The same is (more explicitly) true for Proposition 5.1 from the next section.

5 The "almost isometric" variants

In this section we shall present some refinements of arguments from the preceding section allowing to prove stronger versions of the implication Conjecture 2.2 $\Rightarrow$ Conjecture 2.4, in particular the one announced in Remark 2.3, i.e. requiring the validity of the former for just *one* $\gamma < 1$.

We note first that in the preceding section we did not use the validity of Conjecture 2.2 for *all* $\gamma > 0$, but just for *some* specific (possibly rather small) $\gamma > 0$, depending on the absolute constants c, C from Fact 3.1(a), (b). Moreover, if we use Fact 3.1(c) instead of (b), an easy modification of the argument shows that we may derive Conjecture 2.4 from Conjecture 2.2 being valid for *some* fixed $\gamma < \frac{1}{2}$.

To get the "almost isometric" variant (any *fixed* $\gamma < 1$) we must work slightly harder; let us state it here for future reference.

Proposition 5.1 *Suppose that there exist constants $\gamma, \tau \in (0, 1)$ such that,* *for every $n \in \mathbb{N}$ and $K = \mathrm{conv}\, S \subset \mathbb{R}^n$ verifying* $\max\{\log |S|,\ \log N(K, D)\} \le \tau n$ *one has $\gamma D \not\subset K$. Then, for all $n \in \mathbb{N}$ and $K = \mathrm{conv}\, S \subset \mathbb{R}^n$ we have*

$$M^*(K \cap D) \le w\sqrt{\frac{k}{n}}, \tag{9}$$

where $k = \max\{\log |S|,\ \log N(K, D)\}$ *and w is a constant depending only on γ and τ.*

More precisely, if for some $\gamma, \tau \in (0,1)$, some $n_0 \in \mathbb{N}$, and all $K = \mathrm{conv}\, S \subset \mathbb{R}^{n_0}$ *(resp., for all* $K = \mathrm{conv}\, S \subset RD \subset \mathbb{R}^{n_0}$*; for some* $R > 0$*) the inequality* $k_0 := \max\{\log |S|,\ \log N(K, D)\} \le \tau n_0$ *implies $\gamma D \not\subset K$, then, for all $n \in \mathbb{N}$ and all* $K = \mathrm{conv}\, S \subset \mathbb{R}^n$ *(resp., for all* $K = \mathrm{conv}\, S \subset RD \subset \mathbb{R}^n$*; same R) such that* $\max\{\log |S|,\ \log N(K, D)\} \le k_0$*, one has* $M^*(K \cap D) \le w\sqrt{\frac{k_0}{n}}$*, with* $w \lesssim \tau^{-1/2}(1-\gamma)^{-1}$.

Proof. It is enough to prove the second statement. Observe first that if, for all $K = \mathrm{conv}\, S \subset \mathbb{R}^{n_0}$ with $\max\{\log |S|,\ \log N(K, D)\} \le k_0$, we have $\gamma D \not\subset K$, then the same is true with n_0 replaced by any $n \ge n_0$: any counterexample $K \subset \mathbb{R}^n$ can be projected back on an n_0-dimensional subspace.

Now choose $\varepsilon > 0$ so that $\gamma \le (1 - C\varepsilon)/(1 + C\varepsilon)$ and $\varepsilon \le \min\{\varepsilon_0, 1/2\}$, where C and ε_0 are as in Fact 3.1(c). (Note that the first restriction translates into $\varepsilon \le C^{-1}\frac{1-\gamma}{1+\gamma} \simeq 1 - \gamma$.) Let $n_1 = \varepsilon^2 M^*(K \cap D)^2 n$ and apply Fact 3.1(c) with m replaced by n_1 and K by $K \cap D$ to obtain, for a generic projection P of rank n_1,

$$PD \supset \frac{P(K \cap D)}{(1 + C\varepsilon)M^*(K)} \supset \gamma PD.$$

Without loss of generality we may assume that $\max_{x \in \mathbb{R}^n} M^*((x+D) \cap K)$ is attained at 0 (by Fact 3.6, this is automatically true if K is 0-symmetric) and so, again, generically $P((x + D) \cap K)$ is contained in a ball of radius $(1+C\varepsilon)M^*(D \cap K)$. This follows from Fact 3.1(c) if $\varepsilon M^*((x+D) \cap K) \ge \sqrt{\frac{n_1}{n}}$ (observe that, by the definition of ε, we have equality if $x = 0$) and holds *a fortiori* if the reverse inequality holds: just enlarge $(x + D) \cap K$ to a convex set – still contained in $x + D$ – for which one has the equality.

We now claim that we must have $k_0 > \min\{c'/2, \tau\} n_1$ (where c' is as in Fact 3.1), from which – in combination with the definition of n_1 – the inequality in the assertion immediately follows. Indeed, if that was not the case, i.e., if $k_0 \le c' n_1/2$ and $k_0 \le \tau n_1$, the first of these inequalities would imply that, for a generic projection P (of rank n_1), $P((x+D)\cap K)$ was contained in a ball of radius $(1+C\varepsilon)M^*(D\cap K)$ *simultaneously for all* x in a D-net of K. Setting $K_1 = \frac{PK}{(1+C\varepsilon)M^*(K)} \subset I\!R^{n_1}$ we see that then (generically) K_1 is contained in an union of less than 2^{k_0} balls of radius 1, hence $\log N(K,D) \le k_0$, while, on the other hand, K_1 contains (again generically) a ball of radius γ centered at the origin. (Alternatively, we could have applied P to $W = \bigcup_{s\in S}((K-s)\cap D)$, $M^*(W)$ having been estimated using Lemma 8.1; cf. the end of the proof of Proposition 3.3.) Since the number of extreme points of K_1 *never* exceeds that of K (and hence is $\le 2^{k_0}$), and since $k_0 \le \tau n_1$ and $\tau n_0 \equiv k_0$ imply $n_0 \le n_1$, the hypothesis of our statement applies to K_1 (cf. the remark at the beginning of the proof) yielding $K_1 \not\supset \gamma PD$, a contradiction.

To obtain the version of the statement involving R, we observe that, by Fact 3.2(b), $|Ps| \le 2\sqrt{n_1/n}\ |s|$ *simultaneously for all* $s \in S$ provided that n_1/k is larger than some numerical constant C_1 (to ensure the latter, we replace the condition $k > \min\{c'/2,\tau\}n_1$ above by $k > \min\{c'/2,\tau,C_1^{-1}\}n_1$). The radius R_1 of K_1 is then generically less than $\frac{2\sqrt{n_1/n}\,R}{(1+C\varepsilon)M^*(K)} = \frac{2\varepsilon R}{(1+C\varepsilon)} < R$ and so the hypothesis applies to K_1. (In fact we do have a "gain" in the radius as $R_1 \simeq \varepsilon R$, but since we are going to apply the Proposition for a "fixed" γ anyway, this is not going to be exploited.) $\square$

Remark 5.2. The Proposition above states, in essence, that in order to prove the inequality (3) for a fixed k, it is enough to verify whether it holds for the smallest n for which it is non-trivial, i.e., for which the right hand side is less than, say, $\frac{1}{2}$ (or even whether in that case $K \not\supset \frac{1}{2}D$, a weaker condition; same with $\frac{1}{2}$ replaced by any $\gamma < 1$). Going in the opposite direction, from smaller to larger n, is easy: any counterexample in $I\!R^n$ can be considered as a subset of $I\!R^m$ for $m > n$; the geometric parameters stay the same (see the comment following Conjecture 2.1), while the functionals $M^*(\cdot)$ in dimensions m and n differ (essentially) by a factor $\sqrt{n/m}$ (this can be seen most easily by replacing, via Fact 3.7, spherical averages with Gaussian means and noting that the latter do not change if we increase dimension). Thus, for a fixed k, the statements of type (3) for various n's are equivalent (in the range of n where the right hand side is *uniformly* non-trivial). $\square$

Remark 5.3. By applying a procedure similar to the proof of Proposition 5.1 for sufficiently small ε, one can show that to deduce Conjecture 2.2 or Conjecture 2.4 it is enough to have a "$\gamma = 1 - \delta$ version" of Conjecture 2.2, where δ is an "arbitrarily good" function of (say) $\frac{k}{n}$. A sample form:

Conjecture 2.2 *(or* Conjecture 2.4 *) is equivalent to the following:*

$$\text{If } k, n, S, K, \text{ are as in Conjecture 2.4, then } (1-(k/n)^2)D \not\subset K. \tag{10}$$

Indeed, suppose the above holds and we have a configuration which violates Conjecture 2.2 for some fixed $\gamma < 1$. Apply the previous argument with $\varepsilon = \alpha(\frac{k}{n})^{2/5}$, where α is a small constant. This leads to a $K_0 \subset I\!\!R^{n_1}, n_1 \simeq \varepsilon^2 n$ which admits a D-net of cardinality $\leq 2^k$, (and is spanned by $\leq 2^k$ points)with $(1 - C\varepsilon)D_{n_1} \subset K_0$. One routinely verifies that $C\varepsilon < (\frac{k}{n_1})^2$ if α is properly chosen (n_1 is now the "new" n).

Replacing $\alpha(\frac{k}{n})^{2/5}$ by an appropriate expression, one can obtain an analogue of (10) with $(\frac{k}{n})^2$ replaced by an arbitrary preassigned function of $(\frac{k}{n})$.

Let us remark here that, on the other hand, K cannot contain a ball of radius substantially larger than 1. Indeed, a simple volume comparison argument shows that if $\gamma D \subset K$, then $\gamma^n < 2^k$ and so $\gamma \leq 1 + \frac{k}{n}$. $\square$

6 Preliminary estimates for $M^*(K \cap D)$

Our setup is as in Conjecture 2.4, i.e. $K = \operatorname{conv} S \subset I\!\!R^n$, $k = \log N(K, D)$, $k_1 = \log |S|$; we shall normally assume that $k_1 \leq k$. We recall that, when needed, we may always assume that n/k is "large".

The first estimate is just a rewording of Fact 3.8.

Proposition 6.1 *If $K \subset RD$, then*

$$M^*(K) \leq CR\sqrt{\frac{k_1}{n}}$$

where C is a numerical constant.

The next estimate is much harder, even though the improvement seems rather minor.

Proposition 6.2 *If $K \subset RD$, then*

$$M^*(K \cap D) \leq C_0 \left(R\sqrt{\frac{k}{n}}\sqrt{\frac{k_1}{n}} \right)^{1/2},$$

where C_0 is a universal constant.

Proof. The conclusion of the Proposition can be rewritten as

$$R < c_0 \delta^2 \frac{n}{\sqrt{k_1 k}} \Rightarrow M^*(K \cap D) < \delta$$

for all $\delta \in (0,1)$, where $c_0 = C_0^{-2}$. We show first that, in fact, the Proposition is implied by a formally weaker statement

$$R \leq c_1 \delta^2 \frac{n}{\sqrt{k_1 k}} \;\Rightarrow\; \delta D \not\subset K. \tag{11}$$

(some $c_1 > 0$, with C_0 depending on c_1) and that, moreover, it suffices to obtain (11) just for some *fixed* $\delta \in (0,1)$, for example for $\delta = 1/6$. To that end, set $M^*(K \cap D) = \eta$. Let $\varepsilon = (2C)^{-1}$, where C comes from Fact 3.1(c), and set $n_1 = \varepsilon^2 \eta^2 n$. We may assume that $n_1 \geq k, k_1$ (as otherwise Proposition 6.2 clearly holds). Consider, as in prior arguments, a generic n_1-dimensional projection PK of K. We get, by Fact 3.1(c),

$$\frac{\eta}{2} D_{n_1} \subset P(K \cap D_n) \subset \frac{3\eta}{2} D_{n_1}.$$

Rescaling PK by a factor 3η we get an n_1-dimensional body $K_1 \supset D_{n_1}/6$, for which the respective parameters k, k_1 could only decrease. Now, if (11) held for $\delta = 1/6$, it would follow that the radius R_1 of K_1 would have to verify

$$R_1 \geq \frac{c_1}{36} \frac{n_1}{\sqrt{k_1 k}} = c_2 \frac{\eta^2 n}{\sqrt{k_1 k}}. \tag{12}$$

Now *a priori* we know only that $R_1 = (3\eta)^{-1} \cdot \text{radius}(PK) \leq (3\eta)^{-1} R$. However, for a generic rank m projection P and for any fixed set Σ with $\log |\Sigma| \leq m$, one has $|Px| \simeq \sqrt{\frac{m}{n}}|x|$ *simultaneously* for all $x \in \Sigma$ (by Fact 3.2; this can be made "almost isometric" if $\log |\Sigma|/m$ is "small"). Since the radius of PK is witnessed by $|Ps|$, $s \in S$, and since $\log |S| = k \leq n_1$, it would follow that in fact generically

$$\text{radius}\,(PK) \simeq \sqrt{\frac{n_1}{n}} R \simeq \eta R$$

and so $R_1 \simeq R$, which combined with (12) and the definition of η yields the assertion of Proposition 6.2.

It thus remains to show (11) (in fact just for $\delta = 1/6$, but since that doesn't really simplify the proof, we shall argue the general case). To that end, we need the following special case of "Maurey's Lemma" (see [20]).

Lemma 6.3 *If $S \subset RD$ and $K = \operatorname{conv} S$, then, for every $\varepsilon > 0$, setting $s = \lceil (R/\varepsilon)^2 \rceil$, we get that the set*

$$\{\frac{x_1 + x_2 + \ldots + x_s}{s} \,:\, x_j \in S,\ j = 1, \ldots, s\}$$

is an ε-net for K. In particular, if $k_1 = \log |S|$, then $\log N(K, \varepsilon D) \leq 4k_1(\frac{R}{\varepsilon} + 1)^2$.

Now, to prove (11), assume that $\delta D \subset K$. Set $k_2 = 2k$; we shall show that K contains a 2-separated set of cardinality $\geq 2^{k_2}$, which will contradict $\log N(K, D) \leq k$.

Consider a generic k_2-dimensional projection PK of K. Since we are assuming that $\delta D \subset K$, we also have $\delta D_{k_2} \subset PK$. Let Λ be a $\delta/4$-net of PK consisting (for appropriate s) of points of the form $s^{-1}(Px_1 + Px_2 + \ldots + Px_s)$, where $x_j \in S$, $j = 1, \ldots, s$. Since, by the same argument as in the paragraph following (12) and based on Fact 3.2, radius$(PK) \lesssim \sqrt{\frac{k_2}{n}} R$ in the generic case, Lemma 6.3 implies that it is enough to take $s \approx 4(\sqrt{\frac{k_2}{n}} R/(\delta/4))^2 \simeq \frac{k_2}{n}\frac{R^2}{\delta^2}$, hence $\log |\Lambda| \lesssim \frac{R^2}{\delta^2}\frac{k_2 k_1}{n}$. Now, let $\triangle$ be a maximal $\delta/4$-separated subset of Λ; noticing that $\triangle$ is a $\frac{\delta}{2}$-net for $PK \supset \delta D_{k_2}$ we infer that $|\triangle| \geq 2^{k_2}$. Set $\tilde{\Lambda} = \{s^{-1}(x_1 + \ldots + x_s) : x_j \in S, j = 1, \ldots, s\} \subset K$, in particular

$$\log |\tilde{\Lambda}| \lesssim \frac{R^2}{\delta^2}\frac{k_2 k_1}{n} \tag{13}$$

and let $\tilde{\triangle}$ be the subset of $\tilde{\Lambda}$ corresponding to elements of $\triangle$. We shall show that the elements of $\tilde{\triangle}$ are generically 2-separated; as $|\tilde{\triangle}| = |\triangle|$, this will yield the desired contradiction.

By Fact 3.2, a generic P shortens a given distance by the factor $\sqrt{\frac{k_2}{n}}$ and so, typically, the distance between two elements of $\tilde{\triangle}$ will be $\gtrsim \sqrt{\frac{n}{k_2}}\frac{\delta}{4} \gg 2$. Accordingly, we can afford to settle for a factor *smaller* than $\sqrt{\frac{n}{k_2}}$, but we need to control *all* distances between elements of $\tilde{\triangle}$. To this end, observe that, by Fact 3.2(b), for a fixed $x \in \mathbb{R}^n \setminus \{0\}$ and for $\lambda > 2$,

$$\mathcal{P}\left(|Px| > \lambda\sqrt{\frac{k_2}{n}}|x|\right) < \exp\left(-c'\lambda^2 k_2\right). \tag{14}$$

Choose $\lambda = \sqrt{\frac{n}{k_2}} \cdot \frac{\delta}{8}$ (we may assume $\lambda > 2$). If we knew that, for all $x \in (\tilde{\triangle} - \tilde{\triangle}) \setminus \{0\}$,

$$|Px| \leq \lambda\sqrt{\frac{k_2}{n}}|x| \tag{15}$$

we could infer that, for all such x, one has $|x| > 2$, as required (recall that $|Px| > \delta/4$, the elements of $\triangle$ being $\delta/4$-separated). However, we do not know *a priori* which elements of $\tilde{\Lambda}$ will end up in $\tilde{\triangle}$, and so we need to require (15) for a generic P and for *all* $x \in (\tilde{\Lambda} - \tilde{\Lambda})$. By (13) and (14), this can be assured provided that

$$|\tilde{\Lambda} - \tilde{\Lambda}| \cdot \exp\left(-c'\lambda^2 k_2\right) \leq \exp\left(C\frac{R^2}{\delta^2}\frac{k_2 k_1}{n}\right) \cdot \exp\left(-c'\lambda^2 k_2\right) \ll 1$$

or, say,

$$C\frac{R^2}{\delta^2}\frac{k_2 k_1}{n} \leq \frac{c'}{2}\lambda^2 k_2 = c'' n\delta^2 \,.$$

Considering that $k_2 = 2k$, the above is equivalent, for a properly chosen $c > 0$, to the estimate on R assumed in (11). This concludes the proof of Proposition 6.2. $\square$

7 "Boxing in" the set K.

In the preceding section we did obtain some estimates for $M^*(K \cap D)$ provided the set K was "nicely" bounded. Observe that, e.g.,the estimate from Proposition 6.2 is nontrivial (i.e., $\ll 1$) if $K \subset RD$ with $R \ll \frac{n}{k}$ ($k_1 \leq k$ is tacitly assumed). However, *a priori* no *reasonable* bound on the radius of

K is given (one only has, clearly, $R \le 2^{k+1}$). We shall show now that in fact it is enough to prove Conjecture 2.2 or Conjecture 2.4 in the case when K is "reasonably" bounded. The approach rests again on considering projections of K, this time *deterministic* ones. For simplicity, *in this section* we shall restrict our analysis to the 0-symmetric case.

Recall that, as explained in the last paragraph of section 5, no K verifying our assumptions can contain $(1+\frac{k}{n})D$, i.e., there exist $u_1, |u_1| = 1$ such that

$$K \subset \{|\langle \cdot, u_1\rangle| \le 1 + \frac{k}{n}.$$

Let $K' := P_{\{u_1\}^-}K$ (the projection onto orthogonal complement of u_1). Clearly K' verifies our standard assumptions in $\{u_1\}^-$ and so we can find $u_2 - u_1$, $|u_2| = 1$, such that

$$K_1 \subset \left\{|\langle \cdot, u_2\rangle| \le 1 + \frac{k}{n-1}\right\} \cap \{u_1\}^-$$

and hence

$$K \subset \left\{|\langle \cdot, u_2\rangle| \le 1 + \frac{k}{n-1}\right\}.$$

Continuing in this way we get an orthonormal sequence $u_1, \dots, u_{n_1}, n_1 = n/2$, such that if $E = [u_1, \dots, u_{n_1}]$ (where $[\,\cdot\,]$ denotes the linear span), then

$$K_1 = P_E K \subset \left\{|\langle \cdot, u_j\rangle| \le 1 + \frac{k}{n_1}, j = 1, 2, \dots, n_1\right\}.$$

We did thus show

Proposition 7.1 *If, for some $n, k_1, k \in I\!N$ and $\alpha > 0$ there exists a symmetric set $K = \operatorname{conv} S \subset I\!R^n$ such that*

$$\log N(K, D) \le k \text{ and } \log |S| \le k_1, \text{ while } \alpha D \subset K,$$

then, for $n_1 = n/2$, there exists (a symmetric set) $K_1 \subset I\!R^{n_1}$ satisfying

$$\log N(K_1, D) \le k, \; K_1 = \operatorname{conv} S_1, \; \log |S_1| \le k_1,$$

$$\alpha D_{n_1} \subset K_1 \quad \text{and} \quad K_1 \subset 2B_\infty^{n_1} \subset 2\sqrt{n_1} D_{n_1},$$

where $B_\infty^{n_1} = [-1,1]^{n_1}$ is the $\ell_\infty^{n_1}$ ball.

It follows that for our purposes (i.e. proving Conjecture 2.2 or Conjecture 2.4, or (10)) it is enough to consider sets $K \subset RD$, where $R \leq 2\sqrt{n}$, or even $K \subset 2B^n_\infty$.

A more precise analysis yields a slightly better bound on R; we do not *really* use it in the sequel but present here as the argument seems to be of some interest, in particular it can be adapted to show that Conjecture 2.4 holds for multiples of the unit ball of ℓ^n_1 (and similar sets). Again, only the estimate on $\log N(K, D)$ is used, and, again, it is enough to produce an $n/2$-dimensional projection of K which is contained in RD. The starting point is a well known formula for the asymptotic order of "covering numbers" of the ℓ^n_1-ball $B^n_1 \subset I\!R^n$. We have, for $R \in [1, \sqrt{n}]$ (see [24])

$$\log N(RB^n_1, D) \simeq R^2 \left(\log \frac{n}{R^2} + 1\right) \tag{16}$$

It follows from (16) by a direct calculation that if $\log N(RB^n_1, D) \leq k$, we must have $R \lesssim \sqrt{\frac{k}{\log \frac{n}{k}}}$. This example is representative for the general case, we have

Proposition 7.2 *In the notation and under the assumptions of Proposition 7.1, we have*

$$K_1 \subset C_1 \sqrt{\frac{k}{\log \frac{n}{k}}} D_{n_1}$$

where C_1 is a universal constant.

Proof. Let $R = C_1 \sqrt{\frac{k}{\log \frac{n}{k}}}$, the constant $C_1 > 0$ to be determined later. By a reasoning analogous to the one which led to Proposition 7.1, we see that either there is an $n_1 = n/2$-dimensional projection of K contained in RD (in which case we are done), or there exists a sequence $v_1, v_2, \ldots, v_{n_1}$ of elements of K such that

$$\mathrm{dist}(v_j, [v_i : i < j]) > R, \ j = 1, \ldots, n_1.$$

For simplicity, let us assume (as we may) that the Gramm-Schmidt orthonormalization applied to (v_j) yields the standard basis (e_j), and so

$$\langle v_j, e_i \rangle = 0 \ \text{ and } \ \langle v_i, e_i \rangle > R \ \text{ if } \ 1 \leq j < i \leq n_1. \tag{17}$$

Set $T := \operatorname{conv}\{\pm v_j\} \subset K$. We shall show that the covering numbers of T are "roughly" at least as large as those of $RB_1^{n_1}$: first for the ℓ_∞-norm and, as a consequence, for the Euclidean norm. We start by recalling an estimate "dual" to (16)

$$\log N(D, rB_\infty^n) \simeq \frac{\log(r^2 n + 1)}{r^2},$$

valid for $r \in [\frac{1}{\sqrt{n}}, 1]$, and a related one

$$\log N(tB_1^n, B_\infty^n) \simeq t\left(\log \frac{n}{t} + 1\right)$$

for $t \in [1, n]$, both obtained by, roughly speaking, counting the lattice points contained in respective bodies (in a quite general setting, an essentially equivalent problem to that of calculating the covering numbers $N(\,\cdot\,, B_\infty^n)$), cf. [24]. Let $t \in [1, n]$; it is elementary to show that (17) implies that the linear map u defined by $ue_i = t^{-1} v_i$ sends the integer lattice $\mathbb{Z}^{n_1}$ to a set which is R/t-separated in the ℓ_∞-norm. At the same time, $u(tB_1^{n_1}) \subset T$ and so

$$\log N\left(T, \frac{R}{2t} B_\infty^{n_1}\right) \geq \log N(tB_1^n, B_\infty^n) \geq ct \log\left(\frac{n}{t} + 1\right),$$

where $c > 0$ is a numerical constant. On the other hand, denoting $r = \frac{R}{2t}$, one has

$$\begin{aligned} \log N(T, rB_\infty^{n_1}) &\leq \log N(T, D) + \log N(D, rB_\infty^{n_1}) \\ &\leq \log N(T, D) + C_2 \frac{\log(r^2 n + 1)}{r^2}, \end{aligned}$$

whence

$$\log N(T, D) \geq ct \log\left(\frac{n}{t} + 1\right) - C_2 \frac{\log(r^2 n + 1)}{r^2}.$$

Choosing t so that the first term on the right is twice bigger than the second (in particular $t \simeq R^2$), we get an estimate (16) with RB_1^n replaced by T. As before, this can be reconciled with $\log N(T, D) \leq \log N(K, D) \leq k$ only if $R \leq C_1 \sqrt{\frac{k}{\log \frac{n}{k}}}$, C_1 depending only on c and C_2. □

Remark 7.3. The argument above is based on the fact that

$$\log N(K, D) \leq k \Rightarrow \log N(K, rB_\infty^n) \lesssim k$$

if $r \simeq \sqrt{\frac{\log \frac{n}{k}}{k}}$. Accordingly, if we were able to obtain from K, say, by projections, a body K_1, for which $\log N(K_1, rB_\infty^{n_1}) \geq Ak$ for large A, this would yield a contradiction. $\square$

The last result of this section tells us that, for our purposes, we may additionally assume that $M^*(K)$ is "fairly small" (at least, temporarily, in the symmetric case). We have

Proposition 7.4 *Let $k_1, k \in \mathbb{N}$ and $\alpha > 0$. Suppose that, for some $n \in \mathbb{N}$, there exists a symmetric set $K = \operatorname{conv} S \subset \mathbb{R}^n$ (resp. additionally $K \subset RD$ for some $R > 0$) such that*

$$\log N(K, D) \leq k \text{ and } \log |S| \leq k_1, \text{ while } \alpha D \subset K,$$

then, for $n_1 = n/2$, there exists (a symmetric set) $K_1 \subset \mathbb{R}^{n_1}$ satisfying

$$\log N(K_1, D_{n_1}) \leq k,\ K = \operatorname{conv} S_1,\ \log |S_1| \leq k_1,\ \alpha D_{n_1} \subset K_1$$

and

$$M^*(K_1) \leq C(1 + \log n),$$

(resp. $M^*(K_1) \leq C(1 + \log \frac{R}{\alpha})$ and $K_1 \subset RD_{n_1}$), where C is a universal numerical constant.

If one replaces the hypothesis $\alpha D \subset K$ by a weaker one, $M^*(K \cap D) \geq \alpha$, one gets a similar conclusion, the only changes being that in the new setting $n_1 \simeq \alpha^2 n$, $\frac{1}{2} D_{n_1} \subset K_1$ and $M^*(K_1) \leq C(1 + \log(\alpha^2 n))$ (resp. $M^*(K_1) \leq C(1 + \log R)$).

Proof. By [4] and [22],Theorem 2.5, there exists $u \in GL(n)$ such that

$$M(uK) \cdot M^*(uK) \leq C(1 + \log n), \tag{18}$$

where C is a universal numerical constant or, more precisely, such that $M(uK) \cdot M^*(uK)$ does not exceed the so-called K-convexity constant of $(\mathbb{R}^n, \|x\|_K)$; uK is often referred to as the *ℓ-position* of K. It is well-known and easily seen that if $E \subset \mathbb{R}^n$ is an m-dimensional subspace, then $M(B \cap E)$ exceeds $M(B)$ by at most (asymptotically) $\sqrt{\frac{n}{m}}$. (Indeed, for Gaussian averages we have, identifying E with $\mathbb{R}^m$, $\int_E \|x\|_B \, d\gamma_m(x) \leq \int_{\mathbb{R}^n} \|x\|_B \, d\gamma_n(x)$ – essentially by the triangle inequality – and it remains to apply Fact 3.7.) A

fortiori, the same is true with $B \cap E$ replaced by $P_E(B)$ and, by duality, for $M^*(\cdot)$. Let us choose E, $\dim E = n_1 \geq n/2$, such that $P_E(uD)$ is a ball, say $P_E(uD) = \lambda D_{n_1}$ (we identify E with $I\!R^{n_1}$). We then have

$$\begin{aligned} \alpha\lambda D_{n_1} &\subset P_E(uK) \\ \log N(P_E(uK), \lambda D_{n_1)}\} &\leq k \end{aligned}$$

and so, if we set $K_1 = \lambda^{-1} P_E(uK)$ (again considered as a subset of $I\!R^{n_1}$, we also drop the subscript n_1 in D_{n_1} in what follows), then

$$\begin{aligned} \alpha D &\subset K_1 \\ \log N(K_1, D) &\leq k \\ M(K_1) \cdot M^*(K_1) &\leq C_0(1 + \log n). \end{aligned} \tag{19}$$

The entropy estimate in (19) implies that the volume of K_1 does not exceed 2^k times the volume of D. As a consequence,

$$M(K_1) \geq 2^{-\frac{k}{n_1}} \geq \frac{1}{2}$$

(this follows just from the Hölder inequality) and so

$$M^*(K_1) \leq C_1(1 + \log n).$$

as required. To settle the variant involving the condition $K \subset RD$ we observe that in that case we obtain (additionally) first $P_E(uK) \subset R\lambda D$ and then, after rescaling, $K_1 \subset RD$, as required.

To get the assertion when just $M^*(K \cap D) \geq \alpha$ is assumed, we argue as in the proof of the implication 2.2 $\Rightarrow$ 2.4 or, more precisely, the proof of Proposition 5.1, cf. Remark 5.2): we first apply to our configuration a generic projection of rank $n_0 \simeq \alpha^2 n$ to obtain K_0, $\frac{1}{2} D_{n_0} \subset K_0 \subset I\!R^{n_0}$, and then repeat the procedure described above. $\square$

Remark 7.5. Proposition 7.4 is the only point where symmetry intervenes in a significant way (the arguments of Propositions 7.1 and 7.2 can be routinely modified to yield nonsymmetric variants). Indeed, it is not known whether (18) can be achieved for a general convex body K (via an *affine* map u; see [1] and [23] for results to date). We could have approached the issue by using [11] to pass to an $n/2$-dimensional projection of K verifying (18). However, as mentioned already in the paragraph preceding Fact 3.1, our *final* estimates can be formally derived from the symmetric case, and so we decided to take the easy way here. See also Remark 8.3.

8 "Combining" the sets.

We start with the following lemma, which is a variant of Theorem 2 of [27] (cf. [10], (3.6)).

Lemma 8.1 *Let $(y_j), (A_j), j = 1, \ldots, N$, and $R > 0$, be such that $y_j \in RD$ and $A_j \subset RD$ for all $j \leq N$. Then*

$$M^* \left(\bigcup_{j \leq N} (y_j + A_j) \right) \leq \max_{j \leq N} M^*(A_j) + C_0 R \sqrt{\frac{\log N}{n}}. \tag{20}$$

Proof. Rescaling reduces the Lemma to the case when $R = 1$, which we shall assume from now on. We have

$$M^* \left(\bigcup_j (y_j + A_j) \right) \leq \max_{j \leq N} M^*(\{y_j\}) + M^* \left(\bigcup_j A_j \right).$$

Since the first term on the right does not exceed $C\sqrt{\frac{\log N}{n}}$ by Proposition 6.1, it is enough to prove (20) when all y_j's are 0. This in turn follows from the isoperimetric inequality (see [16]) : as $A_j \subset D$, the function $\|\cdot\|_{A_j^\circ}$ is 1-Lipschitz and so, for $t > 0$,

$$\mu_n(\|x\|_{A_j^\circ} - M^*(A_j) > t) \leq e^{-nt^2/2},$$

where μ_n is the normalized Lebesgue measure on S^{n-1} (note a slight abuse of notation: we write $\|x\|_{A^\circ} = \max\{\langle x, y \rangle : y \in A\}$ even though this is not necessarily a norm or even a seminorm; it would be more proper to employ the term "the support function of A" used in geometry). Hence

$$\begin{aligned} & \mu_n(\max_j \|x\|_{A_j^\circ} - \max_j M^*(A_j) > t) \\ \leq \quad & \mu_n(\max_j (\|x\|_{A_j^\circ} - M^*(A_j)) > t) \leq \min\{N \cdot e^{-nt^2/2}, 1\} \end{aligned} \tag{21}$$

and so

$$
\begin{aligned}
M^*(\bigcup_j A_j) &= \int \max_j \|x\|_{A_j^\circ} d\mu_n(x) \\
&\leq \max_j M^*(A_j) + \int_0^\infty \mu_n(\max_j \|x\|_{A_j^\circ} - \max_j M^*(A_j) > t)dt \\
&\leq \max_j M^*(A_j) + C_2\sqrt{\frac{\log N}{n}},
\end{aligned}
$$

where the last inequality follows easily from (21). □

¿From the Lemma we derive the following

Proposition 8.2 *Let $K_1, K_2 \subset I\!R^n$ be convex sets such that* $\log N(K_j, D) \leq k$ *for $j = 1, 2$. If K_1, K_2 are symmetric, then*

$$M^*((K_1 + K_2) \cap D) \leq M^*(K_1 \cap D) + M^*(K_2 \cap D) + C\sqrt{\frac{k}{n}}.$$

In the general case, the functional $M^*(\,\cdot\, \cap D)$ needs to be replaced **everywhere** *by* $\max_{x \in I\!R^n} M^*((\,\cdot - x) \cap D)$.

Proof. Let (x_i) and (y_j) be D-nets of K_1 and K_2 respectively. Then

$$
\begin{aligned}
K_1 + K_2 &= \bigcup_{i,j} ((x_i + D) \cap K_1) + ((y_j + D) \cap K_2) \\
&= \bigcup_{i,j} x_i + y_j + ((K_1 - x_i) \cap D + ((K_2 - y_j) \cap D).
\end{aligned}
$$

In particular, $(K_1 + K_2) \cap D$ is contained in the "subunion" restricted to $x_i + y_j \subset 3D$. Hence, if K_1, K_2 are symmetric, then, by Lemma 8.1 and Fact 3.6,

$$
\begin{aligned}
& M^*((K_1 + K_2) \cap D) \\
\leq\ & 3C_0\sqrt{\frac{2k}{n}} + \max_{i,j} M^*(((K_1 - x_i) \cap D) + ((K_2 - y_j) \cap D)) \\
=\ & 3C_0\sqrt{\frac{2k}{n}} + \max_i M^*((K_1 - x_i) \cap D) + \max_j M^*((K_2 - y_j) \cap D) \\
\leq\ & C\sqrt{\frac{k}{n}} + M^*(K_1 \cap D) + M^*(K_2 \cap D),
\end{aligned}
$$

as required. The not-necessarily-symmetric case is proved the same way. $\square$

Remark 8.3. The special case of the Proposition with $K_1 = K, K_2 = -K$ shows that if $\log N(K,D) \le k$, then

$$M^*((K-K)\cap D) \le 2 \max_{x \in \mathbb{R}^n} M^*((K-x)\cap D) + C\sqrt{\frac{k}{n}}$$

(a variant with $K-K$ replaced by $(K-K)/2$ and without factor 2 on the right hand side also holds, e.g., by the argument of Proposition 8.5 below). As already mentioned in the paragraph preceding Fact 3.1, inequalities going in the opposite direction are even easier. Consequently, when estimating $M^*(K \cap D)$, there is no major difference between a symmetric and a non-symmetric setting. More generally, Propositions 8.2 and 8.5 show that the functional in question is stable with respect to doubling the set S or the cardinality of the 1-net, and justify our occasional lack of rigor when adding a few points to S or not diferentiating between k and $k+1$. $\square$

Clearly, there is a lot of flexibility in applying Lemma 8.1, e.g. for "combining" more than two sets. For example, by iteration one gets (for the sake of brevity, we state this and the next result just in the symmetric case)

Corollary 8.4 *Let $K_1, K_2, \ldots, K_s \subset \mathbb{R}^n$ be symmetric convex sets verifying the assumptions from Proposition 8.2. Then*

$$M^*((K_1+K_2+\ldots+K_s)\cap D) \le M^*(K_1\cap D)+\ldots+M^*(K_s\cap D)+Cs\sqrt{\frac{k}{n}}.$$

For completeness, we also state a variant of Proposition 8.2 for convex hulls (rather than Minkowski sums), which we do not need for the direct purposes of this paper. Its appeal lies in the fact that the multiplicative constant on the right hand side is 1, a feature that is important in some contexts.

Proposition 8.5 *Let $K_1, K_2 \subset \mathbb{R}^n$ be symmetric convex sets such that $\log N(K_j, D) \le k$ for $j = 1,2$. Then*

$$M^*(\mathrm{conv}\,(K_1 \cup K_2)\cap D) \le \max\{M^*(K_1\cap D), M^*(K_2\cap D)\} + C\sqrt{\frac{k}{n}} + C\sqrt{\frac{\log n}{n}}.$$

Proof. Set $M = \max\{M^*(K_1 \cap D), M^*(K_2 \cap D)\}$. Arguing as in the proof of Proposition 8.2, we get

$$\operatorname{conv}(K_1 \cup K_2) = \bigcup_{i,j,t\in[0,1]} (1-t)((x_i + D) \cap K_1) + t((y_j + D) \cap K_2)$$

$$= \bigcup_{i,j,t\in[0,1]} (1-t)x_i + ty_j + ((1-t)((K_1 - x_i) \cap D) + t((K_2 - y_j) \cap D))$$

Similarly as in Proposition 8.2, to analyze $\operatorname{conv}(K_1 \cup K_2) \cap D$ it is enough to consider only the subsegments of the segments $(1-t)x_i + ty_j$ that lie in $2D$. Given $\varepsilon > 0$, let S' be an ε-net for the union of such subsegments with $|S'| \leq (1 + 4/\varepsilon)2^{2k}$. Applying Lemma 8.1 gives

$$M^*(\operatorname{conv}(K_1 \cup K_2) \cap D) \leq M + \varepsilon + C_0\sqrt{\frac{\log(1 + 4/\varepsilon) + 2k}{n}}.$$

Optimizing over $\varepsilon > 0$ we get the assertion. □

9 Further estimates for $M^*(K \cap D)$

Similarly as in section 6, the setup is as in Conjecture 2.4, i.e.

$$K = \operatorname{conv} S, \quad k = \max\{\log|S|, \log N(K, D)\} \tag{22}$$

(we did supress above the dimension n of the ambient space as it may vary from point to point; cf. Remark 5.2). We recall that the objective is to show that $M^*(K \cap D)$ is "small" if n/k is "large". Thus far we did prove (in section 6, Proposition 6.2) that this holds provided $K \subset RD$ with $R \ll \frac{n}{k}$, while (in section 7, Proposition 7.2) it is shown that one may assume, without loss of generality, that $R \leq C\sqrt{\frac{k}{\log\frac{n}{k}}}$, where C is a numerical constant. Admittedly, the gap between the two estimates is significant. Still, they do allow to deduce our "objective" if $k \ll \frac{n^{2/3}}{(\log n)^{1/3}}$. In this section we shall narrow the gap substantially by proving

Proposition 9.1 *There exists a constant $c > 0$ such that whenever $S, K \subset \mathbb{R}^n$ and k verify* (22) *and $K \subset RD$, with $R \leq \exp(c(\frac{n}{k})^{1/6})$, then $\frac{1}{2}D \not\subset K$.*

Corollary 9.2 *If n, S, K and k are as in Proposition 9.1 and, for some $R \geq 1$, $K \subset RD$, then*

$$M^*(K \cap D) \leq C(1+\log R)^3 \sqrt{\frac{k}{n}},$$

where C is a universal constant.

Proof. More generally, any condition of the type $R \leq \psi(\sqrt{\frac{n}{k}})$ (for $\psi : I\!R^+ \to I\!R^+$, $\psi \nearrow +\infty$) in the theorem translates into an estimate $M^*(K \cap D) \leq C_1 \psi^{-1}(R)\sqrt{\frac{k}{n}}$ (here of course $\psi(x) = \exp(cx^{1/3})$). This follows from the second statement of Proposition 5.1 (with $\gamma = \frac{1}{2}$): the condition $R \leq \psi(\sqrt{\frac{n}{k}})$ (which assures $\frac{1}{2}D \not\subset K$) translates into $k \leq (\psi^{-1}(R))^{-2} n$ and so the hypothesis of that statement is satisfied with $\tau = (\psi^{-1}(R))^{-2}$, which yields $w \leq C_1\psi^{-1}(R)$ in (9), as required. $\square$

The next two corollaries summarize the progress obtained in this note towards the Geometric Lemma and the Duality Conjecture, and so we state them as theorems.

Theorem 9.3 *There exists a constant $C > 0$ such that if $S, K \subset I\!R^n$ and k are as in* (22)*, then $M^*(K \cap D) \leq C(1+\log k)^3 \sqrt{\frac{k}{n}}$. In particular, there exists a constant $c > 0$ such that if $k \leq c(1+\log n)^{-6} n$, then $M^*(K \cap D) < \frac{1}{2}$ (and, consequently, $\frac{1}{2}D \not\subset K$).*

Proof. Consider first the case when K is symmetric. Let $w = w(k)$ be the smallest constant such that the inequality $M^*(K \cap D) \leq w\sqrt{\frac{k}{n}}$ hold for all $n \geq k$ and for all S, K verifying the hypothesis (22) with K symmetric. Consider K, for which $M^*(K \cap D) = w\sqrt{\frac{k}{n}}$ (by compactness, the supremum of $M^*(K \cap D)/\sqrt{\frac{k}{n}}$ is achieved; one could of course devise an argument not using that fact). By the last part of Proposition 7.4, this yields a set $K_1 \subset I\!R^{n_1}$ with $n_1 \simeq w^2 k$, verifying (22) for the same value of k, and such that $\frac{1}{2}D_{n_1} \subset K_1$. After further halving the dimension (via Proposition 7.1) we may additionally attain $R \leq 2\sqrt{n_1} \simeq w\sqrt{k}$. Corollary 9.2 yields then

$$\frac{1}{2} \leq C(1+\log(w\sqrt{k}))^3 \sqrt{\frac{k}{n_1}} \simeq (1+\log(wk))^3 \frac{1}{w},$$

hence $w \lesssim (1 + \log(wk))^3$, which is only possible if $w \lesssim (1 + \log k)^3$, as required. The not-necessarily-symmetric case follows formally, see the comments preceding Fact 3.1. $\square$

Finally, let us restate Theorem 9.3 in terms of covering numbers and entropy numbers (the restating requires only the definitions and a direct application of Proposition 4.1).

Theorem 9.4 *There exist numerical constants $a, C > 0$ such that, for all n, all convex sets $K \subset \mathbb{R}^n$ and all k,*

$$\log N(K, D) \leq k \ \Rightarrow\ \log N(D, CwK^\circ) \leq ak.$$

where $w = (1 + \log k)^3$. Similarly, for a compact operator u, whose range is a Hilbert space,

$$e_{ak}(u^*) \leq Cwe_k(u),$$

with the same w. Moreover, the second statement (and the first in the symmetric case) holds for any given $a > 1$, with the price being paid then in the magnitude of $C = C(a)$.

This could be compared with the "best to date" duality results for general operators of rank $\leq n$ (see Corollary 2.4 of [21]), where an analogous estimate with $w = (1 + (\frac{n}{k})^2)(\log(2 + \frac{n}{k}))^2$ is obtained (our estimate is superior for $k \ll n(\log n)^{-3/2} \log\log n$).

Proof of Proposition 9.1. Let $n, k \in \mathbb{N}$, $R \in [1, \infty)$ and assume that $K = \operatorname{conv} S \subset \mathbb{R}^n$ is such that $\max\{\log|S|, \log N(K, D)\} \leq k$ and $\frac{1}{2}D \subset K \subset RD$. Since the symmetric set $(K - K)/2$ verifies the same hypotheses with k replaced by $2k$, we may and shall assume that K and S were symmetric to begin with, and that $S \ni 0$ (cf. Remark 8.3 and the comments in the paragraph preceding Fact 3.1). By Proposition 7.4, at the price of halving the dimension we may further assume that

$$M^*(K) \leq C'(1 + \log R). \qquad (23)$$

To take advantage of various estimates we obtained for $M^*(\cdot \cap D)$ we will, roughly speaking, decompose the set K into a Minkowski sum of "more easily manageable" sets. Let us first demonstrate a single step of such a decomposition. Let $t \in [1, R)$; by Fact 3.4 (Sudakov's inequality) combined with (23) we have

$$k_1 := \log N(K, tD) \lesssim (\frac{1 + \log R}{t})^2 n. \tag{24}$$

Consider the corresponding t-net of K, i.e. the set $\mathcal{N}_1$ verifying

$$S_1 + tD \supset K, \log |\mathcal{N}_1| \leq k_1.$$

Assign to each $s \in S$ an $s' \in \mathcal{N}_1$ such that $s \in s' + tD$ and let S_2 consist of all the differences $s - s'$, then $\log |S_2| \leq k$. (K and S being symmetric with $S \ni 0$, we may arrange that the same is true for $\mathcal{N}_1$ and S_2; these conditions are not indispensable for the argument, but they do clarify the picture.) Set $K_1 = \operatorname{conv} \mathcal{N}_1$ and $K_2 = \operatorname{conv} S_2$, then

$$K \subset K_1 + K_2 \tag{25}$$

and

$$K_2 \subset tD. \tag{26}$$

Now, by (25), Proposition 8.2 and the estimates for cardinalities of $\mathcal{N}_1$ and S_2,

$$\frac{1}{2} \leq M^*(K \cap D) \leq M^*(K_1 \cap D) + M^*(K_2 \cap D) + C_1 \sqrt{\frac{k}{n}}.$$

The first term on the right can be now efficiently handled via Proposition 6.2 (k_1 being rather small if t is "large"), while the second term is more susceptible to majorizing even via Proposition 6.1, the radius of K_2 being, by (26), significantly smaller than that of K if t is not "too large". To be absolutely precise, in the process we lost control of $N(K, D)$; we only know that $\log N(K, 2D) \leq k$ (which follows trivially from $K_2 \subset K - K = 2K$). This is readily remedied by considering instead the chain of inequalities

$$\frac{1}{4} \leq M^* \left(\frac{1}{2} K \cap D\right) \leq M^* \left(\frac{1}{2} K_1 \cap D\right) + M^* \left(\frac{1}{2} K_2 \cap D\right) + C_1 \sqrt{\frac{k}{n}}\ . \tag{27}$$

Now, by Proposition 6.2 and (24),

$$M^* \left(\frac{1}{2} K_1 \cap D\right) \lesssim \left(\frac{R}{2} \sqrt{\frac{k}{n}} \sqrt{\frac{k_1}{n}}\right)^{1/2} \lesssim \left(R \sqrt{\frac{k}{n}} \frac{1 + \log R}{t}\right)^{1/2}$$

and similarly (as $\log N(\frac{1}{2}K_2, D) \leq k$)

$$M^*(\frac{1}{2}K_2 \cap D) \lesssim \left(t\sqrt{\frac{k}{n}}\sqrt{\frac{k}{n}}\right)^{1/2}.$$

Combining these with (27) and optimizing over $t \in [1, R)$ we obtain

$$\frac{1}{4} \lesssim (R(1+\log R)(\frac{k}{n})^{3/2})^{1/4}.$$

This inequality (obtained assuming that $\frac{1}{2}D \subset K \subset RD$) is impossible if $k \leq c_1 n/(R(1+\log R))^{2/3}$ for sufficiently small $c_1 > 0$, leading (cf. the proof of Corollary 9.2) to the estimate

$$M^*(K \cap D) \lesssim (R(1+\log R))^{1/3}\sqrt{\frac{k}{n}},$$

which already is an improvement over Proposition 6.2 and (cf. the remarks at the beginning of this section) allows to deduce that $M^*(K \cap D)$ is "small" provided $R \ll (\frac{n}{k})^{3/2}/\log\frac{n}{k}$ or if $k \ll \frac{n^{3/4}}{(\log n)^{1/4}}$.

To obtain a better estimate, we – roughly speaking – "decompose" K into a Minkowski sum of $\log R$ sets. Let us return to the setup described in (23) and the paragraph preceding it. To simplify the notation assume that $R = 2^m$ for some $m \in I\!N$. For $j = 1, 2, \ldots, m$ set $R_j = 2^{m-j}$ and let $\mathcal{N}_j$ be an R_j-net of K; by (23) and Fact 3.4 one may assume that

$$\log|\mathcal{N}_j| \leq C_2\left(\frac{1+\log R}{R_j}\right)^2 n.$$

This estimate is clearly not optimal for the last few j's, we improve it by setting $\mathcal{N}_j = S$ when the right hand side exceeds k. In particular we get $\mathcal{N}_m = S$ and

$$\log N(K, R_j D) \leq k_j := \log|\mathcal{N}_j| \leq \min\left\{C_2\left(\frac{1+\log R}{R_j}\right)^2 n\,,\ k\right\}. \qquad (28)$$

As in the "two term decomposition", we set $S_1 = \mathcal{N}_1$, while for $j > 1$ we assign to each $s \in \mathcal{N}_j$ an $s' \in \mathcal{N}_{j-1}$ such that $s \in s' + R_{j-1}D$ and let S_j consist of all the differences $s - s'$; then

$$\log |S_j| \leq k_j. \tag{29}$$

(For a more transparent argument, we may again arrange that all S_j's are symmetric and contain 0, and that $\mathcal{N}_j \subset \mathcal{N}_{j+1}$.) Set $K_j = \operatorname{conv} S_j$, then

$$\frac{K}{2} \subset \frac{K_1}{2} + \frac{K_2}{2} + \ldots + \frac{K_m}{2} \tag{30}$$

and

$$\frac{K_j}{2} \subset \frac{1}{2} R_{j-1} D = R_j D, \ \log N(\frac{K_j}{2}, D) \leq k. \tag{31}$$

Similarly as before, by (30), Proposition 8.4 and (28),

$$\frac{1}{4} \leq \sum_{j=1}^{m} M^* \left(\frac{K_j}{2} \cap D \right) + C_3 m \sqrt{\frac{k}{n}}.$$

On the other hand, by Proposition 6.2, (29), (31) and (28),

$$\begin{aligned} M^* \left(\frac{K_j}{2} \cap D \right) &\lesssim \left(R_j \sqrt{\frac{k}{n}} \sqrt{\frac{k_j}{n}} \right)^{1/2} \\ &\lesssim \left(R_j \sqrt{\frac{k}{n}} \frac{1 + \log R}{R_j} \right)^{1/2} = \left(\sqrt{\frac{k}{n}} (1 + \log R) \right)^{1/2} \end{aligned} \tag{32}$$

for $j = 1, 2, \ldots, m$. Combining the last two inequalities gives

$$\frac{1}{4} \lesssim m \sqrt{\frac{k}{n}} + m \left(\sqrt{\frac{k}{n}} (1 + \log R) \right)^{1/2} \lesssim (\log R)^{3/2} \left(\frac{k}{n} \right)^{1/4}, \tag{33}$$

which is impossible if $\frac{k}{n} < c_2 (\log R)^{-6}$ (for a properly chosen $c_2 > 0$) or, equivalently, $R < \exp\left(c(\frac{n}{k})^{1/6}\right)$ with $c = c_2 / \log e$, as required. This completes the proof of Proposition 9.1. $\square$

Remark 9.5. A significant step in the proof of Proposition 9.1 involved reducing the argument – via Proposition 7.4 – to the case when $M^*(K)$ is "controlled". We wish to point out that even if $M^*(K)$ is bounded by a universal constant, our argument doesn't give estimates substantially better than those contained in Proposition 9.1 (and Corollaries 9.2, 9.3) for the general case. The only improvement is that the exponents $1/6, 3$ and -6 in the respective statements are then replaced by $1/4, 2$ and -4. $\square$

Remark 9.6. Another reason for the logarithmic factor in, say, Corollary 9.2, is that we use a Sudakov type inequality (Fact 3.4) to estimate the cardinality of nets of K for different "degrees of resolution" and then put these estimates together to majorize $M^*(K \cap D)$. This has an inherent error as it doesn't capture the possible difference between the "Dudley majoration" and the "Sudakov minoration" (cf. [10], (12.2) and (12.3)) for the expectation of a supremum of a Gaussian process. The "obvious" way to (attempt to) remedy this problem would be to try to use the majorizing measures ([26]) as the basis for calculation. However, even if we were successful in implementing this program, it appears that we couldn't remove *all* logarithmic factors: the quantities k_j in (32) appear with the exponent $1/4$ as opposed to $1/2$ in the standard "entropy integral" and so the most improvement we could hope for would be replacing m by $m^{1/2}$ in the term $C_9 m(\sqrt{\frac{k}{n}}(1 + \log R))^{1/2}$ in (33), resulting in the same "gain" in the exponents as in the previous Remark. Moreover, even if we were able to simultaneously "force" the boundedness of $M^*(K)$, avoid the "Sudakov-Dudley discrepancy" and somehow handle better the term $C_3 m\sqrt{\frac{k}{n}}$ in (33) (coming from Proposition 8.4), we would still be left with a $m^{1/2} \simeq (\log R)^{1/2}$ factor at the right end of (33), leading to exponents $1/2, 1$ and -2 in Proposition 9.1 and Corollaries 9.2, 9.3. $\square$

Remark 9.7. The procedure of decomposing the set K into a Minkowski sum of "manageable" sets is actually somewhat noncanonical. Let us explain that point in the simpler case of "splitting" into a sum of just two terms (by demonstrating which we started our proof of Proposition 9.1). What happens is that the construction of the set K_2 is based on a kind of " retraction" of S to S_2 given by the correspondence $s \to s - s'$, which *a priori* can be a rather irregular map. The following approach is more natural. For a closed convex body $B \subset I\!\!R^n$ let $\mathcal{R}_B$ be the metric projection of $I\!\!R^n$ onto B (i.e., the "nearest point" map); then $\mathcal{R}_B$ and $\mathcal{Q}_B := I - \mathcal{R}_B$ are *contractions* (all operations being considered with respect to the Euclidean metric). Now redefine S_2 as $\mathcal{Q}_{K_1}(S)$. The prior argument carries over to this setting, in fact we do even have $N(S_2, D) \leq |S| \leq 2^k$, S_2 being a contraction of S. However, later in the process we use Proposition 6.2 to estimate $M^*(K_2 \cap D)$ and for that we need to control $N(K_2, D)$, which is not easily attainable : the maps $\mathcal{Q}_B$ being nonlinear, there is no reason why the set $\mathcal{Q}_{K_1}(K) = \mathcal{Q}_{K_1}(\mathrm{conv}\, S)$ would contain $K_2 := \mathrm{conv} S_2 = \mathrm{conv} \mathcal{Q}_{K_1}(S)$. Accordingly, this modification of the argument does not improve the estimates obtained in any substantial way. $\square$

References

[1] W. Banaszczyk, A. Litvak, A. Pajor and S. J. Szarek, *The flatness theorem for non-symmetric convex bodies via the local theory of Banach spaces. To appear*

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

[3] B. Carl and I. Stephani, *Entropy, compactness and the approximation of operators.* Cambridge Tracts in Mathematics, 98. Cambridge University Press, Cambridge, 1990.

[4] T. Figiel and N. Tomczak-Jaegermann (1979). *Projections onto Hilbertian Subspaces of Banach spaces.* Israel J. Math. **33**, 155-171.

[5] A. A. Giannopoulos and V. Milman, *Euclidean structure in Banach spaces.*, In *Handbook on the Geometry of Banach spaces*, W. B. Johnson, J. Lindenstrauss eds., Elsevier Science, *to appear*

[6] E. D. Gluskin, Y. Gordon and A. Pajor, *Entropy, approximation numbers and other parameters.* In *Handbook on the Geometry of Banach spaces*, W. B. Johnson, J. Lindenstrauss eds., Elsevier Science, *to appear*

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

[8] W. B. Johnson and J. Lindenstrauss, *Extensions of Lipschitz mappings into a Hilbert space.* In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, 189–206, Contemp. Math. **26**, Amer. Math. Soc., Providence, R.I., 1984.

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

[10] M. Ledoux and M. Talagrand, *Probability in Banach spaces. Isoperimetry and processes.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3), **23**. Springer-Verlag, Berlin, 1991.

[11] A. Litvak and N. Tomczak-Jaegermann, *Random aspects of high-dimensional convex bodies. Preprint.*

[12] P. Mankiewicz and N. Tomczak-Jaegermann, *Random Banach spaces.* In *Handbook on the Geometry of Banach spaces*, W. B. Johnson, J. Lindenstrauss eds., Elsevier Science, *to appear*

[13] V. Milman, *Random subspaces of proportional dimension of finite-dimensional normed spaces: approach through the isoperimetric inequality.* Banach spaces (Columbia, Mo., 1984), 106–115, Lecture Notes in Math., **1166**, Springer, Berlin-New York, 1985.

[14] V. Milman, *A note on a low M^*-estimate.* Geometry of Banach spaces (Strobl, 1989), 219–229, London Math. Soc. Lecture Note Ser., **158**, Cambridge Univ. Press, Cambridge, 1990.

[15] V. Milman and A. Pajor, *Entropy and asymptotic geometry of nonsymmetric convex bodies.* Preprint.

[16] V. Milman and G. Schechtman, *Asymptotic theory of finite-dimensional normed spaces. With an appendix by M. Gromov.* Lecture Notes in Mathematics, **1200**. Springer-Verlag, Berlin-New York, 1986.

[17] A. Pajor and N. Tomczak-Jaegermann, *Subspaces of small codimension of finite-dimensional Banach spaces.* Proc. Amer. Math. Soc. **97** (1986), no. 4, 637–642.

[18] A. Pajor and N. Tomczak-Jaegermann, *Volume ratio and other s-numbers of operators related to local properties of Banach spaces.* J. Funct. Anal. **87** (1989), no. 2, 273–293.

[19] A. Pietsch, *Eigenvalues and s-numbers.* Cambridge Studies in Advanced Mathematics, **13**. Cambridge University Press, Cambridge-New York, 1987.

[20] G. Pisier, *Remarques sur un résultat non publié de B. Maurey.* (French) Seminar on Functional Analysis, 1980–1981, Exp. No. V, 13 pp., Ecole Polytech., Palaiseau, 1981.

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

[22] G. Pisier, *The volume of convex bodies and Banach space geometry.* Cambridge Tracts in Mathematics, **94**. Cambridge University Press, Cambridge, 1989.

[23] M. Rudelson, *Distances between non–symmetric convex bodies and the MM^*-estimate.* Available at *http://xxx.lanl.gov/abs/math/9812010, preprint.*

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

[25] S. J. Szarek, *Spaces with large distance to ℓ_∞^n and random matrices.* Amer. J. Math. **112** (1990), 899–942.

[26] M. Talagrand, *Regularity of Gaussian processes.* Acta Math. **159** (1987), no. 1-2, 99–149.

[27] M. Talagrand, *A simple proof of the majorizing measure theorem.* Geom. Funct. Anal. **2** (1992), no. 1, 118–125.

[28] M. Talagrand, *Sudakov-type minoration for Gaussian chaos processes.* Israel J. Math. **79** (1992), no. 2-3, 207–224.

[29] N. Tomczak-Jaegermann, *Dualité des nombres d'entropie pour des opérateurs à valeurs dans un espace de Hilbert.* (French) C. R. Acad. Sci. Paris Sér. I Math. **305** (1987), no. 7, 299–301.

[30] N. Tomczak-Jaegermann, *Banach-Mazur distances and finite-dimensional operator ideals.* Pitman Monographs and Surveys in Pure and Applied Mathematics, **38**. Longman Scientific & Technical, Harlow, 1989.

Department of Mathematics, Tel Aviv University,
Ramat Aviv, Israel
E-mail: vitali@math.tau.ac.il

Department of Mathematics, Case Western Reserve University,
Cleveland, OH 44106-7058, U.S.A.
and
Université Pierre et Marie Curie, Equipe d'Analyse, Bte 186,
4, Place Jussieu, 75252 Paris, France
E-mail : szarek@ccr.jussieu.fr