On the monotonicity of Minkowski sums towards convexity

Matthieu Fradelizi, Mokshay Madiman, Arnaud Marsiglietti, Artem Zvavitch

On the monotonicity of Minkowski sums towards convexity

Matthieu Fradelizi[*], Mokshay Madiman[†], Arnaud Marsiglietti[‡] and Artem Zvavitch[§]

Abstract

Let us define for a compact set $A \subset \mathbf{R}^n$ the sequence

$$A(k) = \left\{ \frac{a_1 + \cdots + a_k}{k} : a_1, \ldots, a_k \in A \right\} = \frac{1}{k}\Big(\underbrace{A + \cdots + A}_{k \text{ times}}\Big).$$

By a theorem of Shapley, Folkman and Starr (1969), $A(k)$ approaches the convex hull of A in Hausdorff distance as k goes to ∞. Bobkov, Madiman and Wang (2011) conjectured that $\mathrm{Vol}_n(A(k))$ is non-decreasing in k, where Vol_n denotes the n-dimensional Lebesgue measure, or in other words, that when one has convergence in the Shapley-Folkman-Starr theorem in terms of a volume deficit, then this convergence is actually monotone. We prove that this conjecture holds true in dimension 1 but fails in dimension $n \geq 12$. We also discuss some related inequalities for the volume of the Minkowski sum of compact sets, showing that this is fractionally superadditive but not supermodular in general, but is indeed supermodular when the sets are convex. Then we consider whether one can have monotonicity in the Shapley-Folkman-Starr theorem when measured using alternate measures of non-convexity, including the Hausdorff distance, effective standard deviation or inner radius, and a non-convexity index of Schneider. For these other measures, we present several positive results, including a strong monotonicity of Schneider's index in general dimension, and eventual monotonicity of the Hausdorff distance and effective standard deviation. Along the way, we clarify the interrelationships between these various notions of non-convexity, demonstrate applications of our results to combinatorial discrepancy theory, and suggest some questions worthy of further investigation.

2010 Mathematics Subject Classification. Primary 60E15 11B13; Secondary 94A17 60F15.

Keywords. Sumsets, Brunn-Minkowski, supermodular, Shapley-Folkman theorem, convex hull, inner radius, Hausdorff distance, discrepancy.

[*]supported in part by the Agence Nationale de la Recherche, project GeMeCoD (ANR 2011 BS01 007 01).

[†]supported in part by the U.S. National Science Foundation through grants DMS-1409504 (CAREER) and CCF-1346564.

[‡]supported in part by the Institute for Mathematics and its Applications with funds provided by the National Science Foundation.

[§]supported in part by the U.S. National Science Foundation Grant DMS-1101636, Simons Foundation and the Bézout Labex.

Contents

1 Introduction **3**

2 Measures of non-convexity **7**
2.1 Preliminaries and Definitions 7
2.2 Basic properties of non-convexity measures 10
2.3 Special properties of Schneider's index 13
2.4 Unconditional relationships 14
2.5 Conditional relationships 20

3 The behavior of volume deficit **24**
3.1 Monotonicity of volume deficit in dimension one and for Cartesian products . 24
3.2 A counterexample in dimension ≥ 12 26
3.3 Convergence rates for Δ 27

4 Volume inequalities for Minkowski sums **28**
4.1 A refined superadditivity of the volume for compact sets 28
4.2 Supermodularity of volume for convex sets 30

5 The behavior of Schneider's non-convexity index **35**
5.1 The refined monotonicity of Schneider's non-convexity index 35
5.2 Convergence rates for Schneider's non-convexity index 37

6 The behavior of the effective standard deviation v **38**
6.1 Subadditivity of v^2 38
6.2 Strong fractional subadditivity for large k 39
6.3 Convergence rates for v 40

7 The behavior of the Hausdorff distance **41**
7.1 Some basic properties of the Hausdorff distance 41
7.2 Strong fractional subadditivity for large k 42
7.3 Convergence rates for d 44

8 Connections to discrepancy theory **45**

9 Discussion **49**

1 Introduction

Minkowski summation is a basic and ubiquitous operation on sets. Indeed, the Minkowski sum $A+B = \{a+b : a \in A, b \in B\}$ of sets A and B makes sense as long as A and B are subsets of an ambient set in which the operation $+$ is defined. In particular, this notion makes sense in any group, and there are multiple fields of mathematics that are preoccupied with studying what exactly this operation does. For example, much of classical additive combinatorics studies the cardinality of Minkowski sums (called sumsets in this context) of finite subsets of a group and their interaction with additive structure of the concerned sets, while the study of the Lebesgue measure of Minkowski sums in $\mathbf{R}^n$ is central to much of convex geometry and geometric functional analysis. In this survey paper, which also contains a number of original results, our goal is to understand better the qualitative effect of Minkowski summation in $\mathbf{R}^n$– specifically, the "convexifying" effect that it has. Somewhat surprisingly, while the existence of such an effect has long been known, several rather basic questions about its nature do not seem to have been addressed, and we undertake to fill the gap.

The fact that Minkowski summation produces sets that look "more convex" is easy to visualize by drawing a non-convex set[1] in the plane and its self-averages $A(k)$ defined by

$$A(k) = \left\{ \frac{a_1 + \cdots + a_k}{k} : a_1, \ldots, a_k \in A \right\} = \frac{1}{k}\Big(\underbrace{A + \cdots + A}_{k \text{ times}}\Big). \tag{1}$$

This intuition was first made precise in the late 1960's independently[2] by Starr [67] (see also [68]), who credited Shapley and Folkman for the main result, and by Emerson and Greenleaf [31]. Denoting by $\mathrm{conv}(A)$ the convex hull of A, by B_2^n the n-dimensional Euclidean ball of radius 1, and by $d(A) = \inf\{r > 0 : \mathrm{conv}(A) \subset A + rB_2^n\}$ the Hausdorff distance between a set A and its convex hull, it follows from the Shapley-Folkman-Starr theorem that if $A_1, \ldots, A_k$ are compact sets in $\mathbf{R}^n$ contained inside some ball, then

$$d(A_1 + \cdots + A_k) \leq O\big(\sqrt{\min\{k, n\}}\big).$$

By considering $A_1 = \cdots = A_k = A$, one concludes that $d(A(k)) = O\big(\frac{\sqrt{n}}{k}\big)$. In other words, when A is a compact subset of $\mathbf{R}^n$ for fixed dimension n, $A(k)$ converges in Hausdorff distance to $\mathrm{conv}(A)$ as $k \to \infty$, at rate at least $O(1/k)$.

Our geometric intuition would suggest that in some sense, as k increases, the set $A(k)$ is getting progressively more convex, or in other words, that the convergence of $A(k)$ to $\mathrm{conv}(A)$ is, in some sense, monotone. The main goal of this paper is to examine this intuition, and explore whether it can be made rigorous.

One motivation for our goal of exploring monotonicity in the Shapley-Folkman-Starr theorem is that it was the key tool allowing Starr [67] to prove that in an economy with a

[1] The simplest nontrivial example is three non-collinear points in the plane, so that $A(k)$ is the original set A of vertices of a triangle together with those convex combinations of the vertices formed by rational coefficients with denominator k.

[2] Both the papers of Starr [67] and Emerson and Greenleaf [31] were submitted in 1967 and published in 1969, but in very different communities (economics and algebra); so it is not surprising that the authors of these papers were unaware of each other. Perhaps more surprising is that the relationship between these papers does not seem to have ever been noticed in the almost 5 decades since. The fact that $A(k)$ converges to the convex hull of A, at an $O(1/k)$ rate in the Hausdorff metric when dimension n is fixed, should perhaps properly be called the Emerson-Folkman-Greenleaf-Shapley-Starr theorem, but in keeping with the old mathematical tradition of not worrying too much about names of theorems (cf., Arnold's principle), we will simply use the nomenclature that has become standard.

sufficiently large number of traders, there are (under some natural conditions) configurations arbitrarily close to equilibrium even without making any convexity assumptions on preferences of the traders; thus investigations of monotonicity in this theorem speak to the question of whether these quasi-equilibrium configurations in fact get "closer" to a true equilibrium as the number of traders increases. A related result is the core convergence result of Anderson [3], which states under very general conditions that the discrepancy between a core allocation and the corresponding competitive equilibrium price vector in a pure exchange economy becomes arbitrarily small as the number of agents gets large. These results are central results in mathematical economics, and continue to attract attention (see, e.g., [60]).

Our original motivation, however, came from a conjecture made by Bobkov, Madiman and Wang [19]. To state it, let us introduce the volume deficit $\Delta(A)$ of a compact set A in $\mathbf{R}^n$: $\Delta(A) := \mathrm{Vol}_n(\mathrm{conv}(A) \setminus A) = \mathrm{Vol}_n(\mathrm{conv}(A)) - \mathrm{Vol}_n(A)$, where Vol_n denotes the Lebesgue measure in $\mathbf{R}^n$.

Conjecture 1.1 (Bobkov-Madiman-Wang [19]). *Let A be a compact set in $\mathbf{R}^n$ for some $n \in \mathbb{N}$, and let $A(k)$ be defined as in (1). Then the sequence $\{\Delta(A(k))\}_{k\geq 1}$ is non-increasing in k, or equivalently, $\{\mathrm{Vol}_n(A(k))\}_{k\geq 1}$ is non-decreasing.*

In fact, the authors of [19] proposed a number of related conjectures, of which Conjecture 1.1 is the weakest. Indeed, they conjectured a monotonicity property in a probabilistic limit theorem, namely the law of large numbers for random sets due to Z. Artstein and Vitale [6]; when this conjectured monotonicity property of [19] is restricted to deterministic (i.e., non-random) sets, one obtains Conjecture 1.1. They showed in turn that this conjectured monotonicity property in the law of large numbers for random sets is implied by the following volume inequality for Minkowski sums. For $k \geq 1$ being an integer, we set $[k] = \{1, \dots, k\}$.

Conjecture 1.2 (Bobkov-Madiman-Wang [19]). *Let $n \geq 1$, $k \geq 2$ be integers and let $A_1, \dots, A_k$ be k compact sets in $\mathbf{R}^n$. Then*

$$\mathrm{Vol}_n\left(\sum_{i=1}^{k} A_i\right)^{\frac{1}{n}} \geq \frac{1}{k-1}\sum_{i=1}^{k} \mathrm{Vol}_n\left(\sum_{j\in[k]\setminus\{i\}} A_j\right)^{\frac{1}{n}}. \tag{2}$$

Apart from the fact that Conjecture 1.2 implies Conjecture 1.1 (which can be seen simply by applying the former to $A_1 = \cdots = A_k = A$, where A is a fixed compact set), Conjecture 1.2 is particularly interesting because of its close connections to an important inequality in Geometry, namely the Brunn-Minkowski inequality, and a fundamental inequality in Information Theory, namely the entropy power inequality. Since the conjectures in [19] were largely motivated by these connections, we now briefly explain them.

The Brunn-Minkowski inequality (or strictly speaking, the Brunn-Minkowski-Lyusternik inequality) states that for all compact sets A, B in $\mathbf{R}^n$,

$$\mathrm{Vol}_n(A + B)^{1/n} \geq \mathrm{Vol}_n(A)^{1/n} + \mathrm{Vol}_n(B)^{1/n}. \tag{3}$$

It is, of course, a cornerstone of Convex Geometry, and has beautiful relations to many areas of Mathematics (see, e.g., [36, 62]). The case $k = 2$ of Conjecture 1.2 is exactly the Brunn-Minkowski inequality (3). Whereas Conjecture 1.2 yields the monotonicity described in Conjecture 1.1, the Brunn-Minkowski inequality only allows one to deduce that the subsequence $\{\mathrm{Vol}_n(A(2^k))\}_{k\in\mathbb{N}}$ is non-decreasing (one may also deduce this fact from the trivial inclusion $A \subset \frac{A+A}{2}$).

The entropy power inequality states that for all independent random vectors X, Y in $\mathbf{R}^n$,

$$N(X+Y) \geq N(X) + N(Y), \tag{4}$$

where

$$N(X) = \frac{1}{2\pi e} e^{\frac{2h(X)}{n}}$$

denotes the entropy power of X. Let us recall that the entropy of a random vector X with density function f_X (with respect to Lebesgue measure dx) is $h(X) = -\int f_X(x) \log f_X(x) dx$ if the integral exists and $-\infty$ otherwise (see, e.g., [27]). As a consequence, one may deduce that for independent and identically distributed random vectors X_i, $i \geq 0$, the sequence

$$\left\{ N\left(\frac{X_1 + \cdots + X_{2^k}}{\sqrt{2^k}} \right) \right\}_{k \in \mathbb{N}}$$

is non-decreasing. S. Artstein, Ball, Barthe and Naor [4] generalized the entropy power inequality (4) by proving that for any independent random vectors $X_1, \dots, X_k$,

$$N\left(\sum_{i=1}^{k} X_i \right) \geq \frac{1}{k-1} \sum_{i=1}^{k} N\left(\sum_{j \in [k] \setminus \{i\}} X_j \right). \tag{5}$$

In particular, if all X_i in the above inequality are identically distributed, then one may deduce that the sequence

$$\left\{ N\left(\frac{X_1 + \cdots + X_k}{\sqrt{k}} \right) \right\}_{k \geq 1}$$

is non-decreasing. This fact is usually referred to as "the monotonicity of entropy in the Central Limit Theorem", since the sequence of entropies of these normalized sums converges to that of a Gaussian distribution as shown earlier by Barron [11]. Later, simpler proofs of the inequality (5) were given by [46, 73]; more general inequalities were developed in [47, 64, 48].

There is a formal resemblance between inequalities (4) and (3) that was noticed in a pioneering work of Costa and Cover [26] and later explained by Dembo, Cover and Thomas [28] (see also [70, 74] for other aspects of this connection). In the last decade, several further developments have been made that link Information Theory to the Brunn-Minkowski theory, including entropy analogues of the Blaschke-Santaló inequality [45], the reverse Brunn-Minkowski inequality [17, 18], the Rogers-Shephard inequality [20, 49] and the Busemann inequality [9]; some of this literature is surveyed in [50]. On the other hand, natural analogues in the Brunn-Minkowski theory of Fisher information inequalities hold sometimes but not always [33, 7, 35]. In particular, it is now well understood that the functional $A \mapsto \mathrm{Vol}_n(A)^{1/n}$ in the geometry of compact subsets of $\mathbf{R}^n$, and the functional $f_X \mapsto N(X)$ in probability are analogous to each other in many (but not all) ways. Thus, for example, the monotonicity property desired in Conjecture 1.1 is in a sense analogous to the monotonicity property in the Central Limit Theorem implied by inequality (5), and Conjecture 1.2 from [19] generalizes the Brunn-Minkowski inequality (3) exactly as inequality (5) generalizes the entropy power inequality (4).

The starting point of this work was the observation that although Conjecture 1.2 holds for certain special classes of sets (namely, one dimensional compact sets, convex sets and their Cartesian product, as shown in subsection 3.1), both Conjecture 1.1 and Conjecture 1.2 fail to hold in general even for moderately high dimension (Theorem 3.3 constructs a counterexample in dimension 12). These results, which consider the question of the monotonicity of $\Delta(A(k))$ are stated and proved in Section 3. We also discuss there the question of when one has convergence of $\Delta(A(k))$ to 0, and at what rate, drawing on the work of the [31] (which seems not to be well known in the contemporary literature on convexity).

Section 4 is devoted to developing some new volume inequalities for Minkowski sums. In particular, we observe in Theorem 4.1 that if the exponents of $1/n$ in Conjecture 1.2 are removed, then the modified inequality is true for general compact sets (though unfortunately one can no longer directly relate this to a law of large numbers for sets). Furthermore, in the case of convex sets, Theorem 4.5 proves an even stronger fact, namely that the volume of the Minkowski sum of convex sets is supermodular. Various other facts surrounding these observations are also discussed in Section 4.

Even though the conjecture about $A(k)$ becoming progressively more convex in the sense of Δ is false thanks to Theorem 3.3, one can ask the same question when we measure the extent of non-convexity using functionals other than Δ. In Section 2, we survey the existing literature on measures of non-convexity of sets, also making some possibly new observations about these various measures and the relations between them. The functionals we consider include a non-convexity index $c(A)$ introduced by Schneider [61], the notion of inner radius $r(A)$ introduced by Starr [67] (and studied in an equivalent form as the effective standard deviation $v(A)$ by Cassels [23], though the equivalence was only understood later by Wegmann [75]), and the Hausdorff distance $d(A)$ to the convex hull, which we already introduced when describing the Shapley-Folkman-Starr theorem. We also consider the generalized Hausdorff distance $d^{(K)}(A)$ corresponding to using a non-Euclidean norm whose unit ball is the convex body K. The rest of the paper is devoted to the examination of whether $A(k)$ becomes progressively more convex as k increases, when measured through these other functionals.

In Section 5, we develop the main positive result of this paper, Theorem 5.3, which shows that $c(A(k))$ is monotonically (strictly) decreasing in k, unless $A(k)$ is already convex. Various other properties of Schneider's non-convexity index and its behavior for Minkowski sums are also established here, including the optimal $O(1/k)$ convergence rate for $c(A(k))$. We remark that even the question of convergence of $c(A(k))$ to 0 does not seem to have been explored in the literature.

Section 6 considers the behavior of $v(A(k))$ (or equivalently $r(A(k))$). For this sequence, we show that monotonicity holds in dimensions 1 and 2, and in general dimension, monotonicity holds eventually (in particular, once k exceeds n). The convergence rate of $r(A(k))$ to 0 was already established in Starr's original paper [67]; we review the classical proof of Cassels [23] of this result.

Section 7 considers the question of monotonicity of $d(A(k))$, as well as its generalizations $d^{(K)}(A(k))$ when we consider $\mathbf{R}^n$ equipped with norms other than the Euclidean norm (indeed, following [10], we even consider so-called "nonsymmetric norms"). Again here, we show that monotonicity holds in dimensions 1 and 2, and in general dimension, monotonicity holds eventually (in particular, once k exceeds n). In fact, more general inequalities are proved that hold for Minkowski sums of different sets. The convergence rate of $d(A(k))$ to 0 was already established in Starr's original paper [67]; we review both a classical proof, and also provide a new very simple proof of a rate result that is suboptimal in dimension for the Euclidean norm but sharp in both dimension and number k of summands given that it holds for arbitrary norms.

In Section 8, we show that a number of results from combinatorial discrepancy theory can be seen as consequences of the convexifying effect of Minkowski summation. In particular, we obtain a new bound on the discrepancy for finite-dimensional Banach spaces in terms of the

Banach-Mazur distance of the space from a Euclidean one.

Finally, in Section 9, we make various additional remarks, including on notions of non-convexity not considered in this paper.

Acknowledgments. Franck Barthe had independently observed that Conjecture 1.2 holds in dimension 1, using the same proof, by 2011. We are indebted to Fedor Nazarov for valuable discussions, in particular for the help in the construction of the counterexample in Theorem 3.3. We would like to thank Victor Grinberg for many enlightening discussions on the connections with discrepancy theory, which were an enormous help with putting Section 8 together. We also thank Franck Barthe, Dario Cordero-Erausquin, Uri Grupel, Bo'az Klartag, Joseph Lehec, Paul-Marie Samson, Sreekar Vadlamani, and Murali Vemuri for interesting discussions. Some of the original results developed in this work were announced in [34]; we are grateful to Gilles Pisier for curating that announcement.

2 Measures of non-convexity

2.1 Preliminaries and Definitions

Throughout this paper, we only deal with compact sets, since several of the measures of non-convexity we consider can have rather unpleasant behavior if we do not make this assumption.

The convex hull operation interacts nicely with Minkowski summation.

Lemma 2.1. *Let A, B be nonempty subsets of $\mathbf{R}^n$. Then,*

$$\mathrm{conv}(A+B) = \mathrm{conv}(A) + \mathrm{conv}(B).$$

Proof. Let $x \in \mathrm{conv}(A) + \mathrm{conv}(B)$. Then $x = \sum_{i=1}^k \lambda_i a_i + \sum_{j=1}^l \mu_j b_j$, where $a_i \in A$, $b_j \in B$, $\lambda_i \geq 0$, $\mu_j \geq 0$ and $\sum_{i=1}^k \lambda_i = 1$, $\sum_{j=1}^l \mu_j = 1$. Thus, $x = \sum_{i=1}^k \sum_{j=1}^l \lambda_i \mu_j (a_i + b_j)$. Hence $x \in \mathrm{conv}(A+B)$. The other inclusion is clear. □

Lemma 2.1 will be used throughout the paper without necessarily referring to it.

The Shapley-Folkman lemma, which is closely related to the classical Carathéodory theorem, is key to our development.

Lemma 2.2 (Shapley-Folkman)**.** *Let $A_1, \dots, A_k$ be nonempty subsets of $\mathbf{R}^n$, with $k \geq n+1$. Let $a \in \sum_{i\in[k]} \mathrm{conv}(A_i)$. Then there exists a set I of cardinality at most n such that*

$$a \in \sum_{i\in I} \mathrm{conv}(A_i) + \sum_{i\in[k]\setminus I} A_i.$$

Proof. We present below a proof taken from Proposition 5.7.1 of [16]. Let $a \in \sum_{i\in[k]} \mathrm{conv}(A_i)$. Then

$$a = \sum_{i\in[k]} a_i = \sum_{i\in[k]} \sum_{j=1}^{t_i} \lambda_{ij} a_{ij},$$

where $\lambda_{ij} \geq 0$, $\sum_{j=1}^{t_i} \lambda_{ij} = 1$, and $a_{ij} \in A_i$. Let us consider the following vectors of $\mathbf{R}^{n+k}$,

$$\begin{aligned} z &= (a, 1, \cdots, 1), \\ z_{1j} &= (a_{1j}, 1, 0, \cdots, 0), \quad j \in [t_1], \\ &\vdots \\ z_{kj} &= (a_{kj}, 0, \cdots, 0, 1), \quad j \in [t_k]. \end{aligned}$$

Notice that $z = \sum_{i=1}^{k} \sum_{j=1}^{t_i} \lambda_{ij} z_{ij}$. Using Carathéodory's theorem in the positive cone generated by z_{ij} in $\mathbf{R}^{n+k}$, one has

$$z = \sum_{i=1}^{k} \sum_{j=1}^{t_i} \mu_{ij} z_{ij},$$

for some nonnegative scalars μ_{ij} where at most $n+k$ of them are non zero. This implies that $a = \sum_{i=1}^{k} \sum_{j=1}^{t_i} \mu_{ij} a_{ij}$ and that $\sum_{j=1}^{t_i} \mu_{ij} = 1$, for all $i \in [k]$. Thus for each $i \in [k]$, there exists $j_i \in [t_i]$ such that $\mu_{ij_i} > 0$. But at most $n+k$ scalars μ_{ij} are positive. Hence there are at most n additional μ_{ij} that are positive. One deduces that there are at least $k-n$ indices i such that $\mu_{i\ell_i} = 1$ for some $\ell_i \in [t_i]$, and thus $\mu_{ij} = 0$ for $j \neq \ell_i$. For these indices, one has $a_i \in A_i$. The other inclusion is clear. □

The Shapley-Folkman lemma may alternatively be written as the statement that, for $k \geq n+1$,

$$\mathrm{conv}(\sum_{i\in[k]} A_i) = \bigcup_{I\subset[k]:|I|\leq n} \left[\sum_{i\in I} \mathrm{conv}(A_i) + \sum_{i\in[k]\setminus I} A_i \right], \tag{6}$$

where $|I|$ denotes the cardinality of I. When all the sets involved are identical, and $k > n$, this reduces to the identity

$$k\,\mathrm{conv}(A) = n\,\mathrm{conv}(A) + (k-n)\,A(k-n). \tag{7}$$

It should be noted that the Shapley-Folkman lemma is in the center of a rich vein of investigation in convex analysis and its applications. As explained by Z. Artstein [5], It may be seen as a discrete manifestation of a key lemma about extreme points that is related to a number of "bang-bang" type results. It also plays an important role in the theory of vector-valued measures; for example, it can be used as an ingredient in the proof of Lyapunov's theorem on the range of vector measures (see [43], [29] and references therein).

For a compact set A in $\mathbf{R}^n$, denote by

$$R(A) = \min_x \{r > 0 : A \subset x + rB_2^n\}$$

the radius of the circumscribed ball of A. By Jung's theorem [42], this parameter is close to the diameter, namely one has

$$\frac{\mathrm{diam}(A)}{2} \leq R(A) \leq \mathrm{diam}(A)\sqrt{\frac{n}{2(n+1)}} \leq \frac{\mathrm{diam}(A)}{\sqrt{2}},$$

where $\mathrm{diam}(A) = \sup_{x,y\in A} |x-y|$ is the Euclidean diameter of A. We also denote by

$$\mathrm{inr}(A) = \max_x \{r \geq 0 : x + rB_2^n \subset A\}$$

the inradius of A, i.e. the radius of a largest Euclidean ball included in A. There are several ways of measuring non-convexity of a set:

1. The Hausdorff distance from the convex hull is perhaps the most obvious measure to consider:

$$d(A) = d_H(A, \mathrm{conv}(A)) = \inf\{r > 0 : \mathrm{conv}(A) \subset A + rB_2^n\}.$$

A variant of this is to consider the Hausdorff distance when the ambient metric space is $\mathbf{R}^n$ equipped with a norm different from the Euclidean norm. If K is the closed unit ball of this norm (i.e., any symmetric[3], compact, convex set with nonempty interior), we define

$$d^{(K)}(A) = \inf\{r > 0 : \mathrm{conv}(A) \subset A + rK\}. \tag{8}$$

In fact, the quantity (8) makes sense for any compact convex set containing 0 in its interior– then it is sometimes called the Hausdorff distance with respect to a "nonsymmetric norm".

2. Another natural measure of non-convexity is the "volume deficit":

$$\Delta(A) = \mathrm{Vol}_n(\mathrm{conv}(A) \setminus A) = \mathrm{Vol}_n(\mathrm{conv}(A)) - \mathrm{Vol}_n(A).$$

Of course, this notion is interesting only when $\mathrm{Vol}_n(\mathrm{conv}(A)) \neq 0$. There are many variants of this that one could consider, such as $\log \mathrm{Vol}_n(\mathrm{conv}(A)) - \log \mathrm{Vol}_n(A)$, or relative versions such as $\Delta(A)/\mathrm{Vol}_n(\mathrm{conv}(A))$ that are automatically bounded.

3. The "inner radius" of a compact set was defined by Starr [67] as follows:

$$r(A) = \sup_{x\in\mathrm{conv}(A)} \inf\{R(T) : T \subset A, x \in \mathrm{conv}(T)\}.$$

4. The "effective standard deviation" was defined by Cassels [23]. For a random vector X in $\mathbf{R}^n$, let $V(X)$ be the trace of its covariance matrix. Then the effective standard deviation of a compact set A of $\mathbf{R}^n$ is

$$v^2(A) = \sup_{x\in\mathrm{conv}(A)} \inf\{V(X) : \mathrm{supp}\,(X) \subset A, |\mathrm{supp}\,(X)| < \infty, \mathbb{E}X = x\}.$$

Let us notice the equivalent geometric definition of v:

$$\begin{aligned} v^2(A) &= \sup_{x\in\mathrm{conv}(A)} \inf\{\sum p_i|a_i - x|^2 : x = \sum p_i a_i; p_i > 0; \sum p_i = 1, a_i \in A\} \\ &= \sup_{x\in\mathrm{conv}(A)} \inf\{\sum p_i|a_i|^2 - |x|^2 : x = \sum p_i a_i; p_i > 0; \sum p_i = 1, a_i \in A\}. \end{aligned}$$

5. Another non-convexity measure was defined by Cassels [23] as follows:

$$\rho(A) = \sup_{x\in\mathrm{conv}(A)} \inf_{a\in A_x} |x - a|,$$

where $A_x = \{a \in A : \exists b \in \mathrm{conv}(A), \exists\theta \in (0, 1) \text{ such that } x = (1 - \theta)a + \theta b\}$.

6. The "non-convexity index" was defined by Schneider [61] as follows:

$$c(A) = \inf\{\lambda \geq 0 : A + \lambda\,\mathrm{conv}(A) \text{ is convex}\}.$$

[3] We always use "symmetric" to mean centrally symmetric, i.e., $x \in K$ if and only if $-x \in K$.

2.2 Basic properties of non-convexity measures

All of these functionals are 0 when A is a convex set; this justifies calling them "measures of non-convexity". In fact, we have the following stronger statement since we restrict our attention to compact sets.

Lemma 2.3. *Let A be a compact set in $\mathbf{R}^n$. Then:*

1. *$c(A) = 0$ if and only if A is convex.*
2. *$d(A) = 0$ if and only if A is convex.*
3. *$r(A) = 0$ if and only if A is convex.*
4. *$\rho(A) = 0$ if and only if A is convex.*
5. *$v(A) = 0$ if and only if A is convex.*
6. *Under the additional assumption that* $\mathrm{conv}(A)$ *has nonempty interior, $\Delta(A) = 0$ if and only if A is convex.*

Proof. Directly from the definition of $c(A)$ we get that $c(A) = 0$ if A is convex (just select $\lambda = 0$). Now assume that $c(A) = 0$, then $\{A + \frac{1}{m}\mathrm{conv}(A)\}_{m=1}^{\infty}$ is a sequence of compact convex sets, converging in Hausdorff metric to A, thus A must be convex. Notice that this observation is due to Schneider [61].

The assertion about $d(A)$ follows immediately from the definition and the limiting argument similar to the above one.

If A is convex then, clearly $r(A) = 0$, indeed we can always take $T = (rB_2^n + x) \cap A \neq \emptyset$ with $r \to 0$. Next, if $r(A) = 0$, then using Theorem 2.14 below we have $d(A) \leq r(A) = 0$ thus $d(A) = 0$ and therefore A is convex.

The statements about $\rho(A)$ and $v(A)$ can be deduced from the definitions, but they will also follow immediately from the Theorem 2.14 below.

Assume that A is convex, then $\mathrm{conv}(A) = A$ and $\Delta(A) = 0$. Next, assume that $\Delta(A) = 0$. Assume, towards a contradiction, that $\mathrm{conv}(A) \neq A$. Then there exists $x \in \mathrm{conv}(A)$ and $r > 0$ such that $(x + rB_2^n) \cap A = \emptyset$. Since $\mathrm{conv}(A)$ is convex and has nonempty interior, there exists a ball $y + sB_2^n \subset \mathrm{conv}(A)$ and one has

$$\Delta(A) \geq \mathrm{Vol}_n(\mathrm{conv}(A) \cap (x + rB_2^n)) \geq \mathrm{Vol}_n(\mathrm{conv}(x, y + sB_2^n) \cap (x + rB_2^n)) > 0,$$

which contradicts $\Delta(A) = 0$. $\square$

The following lemmata capture some basic properties of all these measures of non-convexity (note that we need not separately discuss v and ρ henceforth owing to Theorem 2.14). The first lemma concerns the behavior of these functionals on scaling of the argument set.

Lemma 2.4. *Let A be a compact subset of $\mathbf{R}^n$, $x \in \mathbf{R}^n$, and $\lambda \in (0, \infty)$.*

1. *$c(\lambda A + x) = c(A)$. In fact, c is affine-invariant.*
2. *$d(\lambda A + x) = \lambda d(A)$.*
3. *$r(\lambda A + x) = \lambda r(A)$.*

4. $\Delta(\lambda A + x) = \lambda^n \Delta(A)$. *In fact, if* $T(x) = Mx + b$, *where* M *is an invertible linear transformation and* $b \in \mathbf{R}^n$, *then* $\Delta(T(A)) = |\det(M)|\Delta(A)$.

Proof. To see that c is affine-invariant, we first notice that $\mathrm{conv}(TA) = T\mathrm{conv}(A)$. Moreover writing $Tx = Mx + b$, where M is an invertible linear transformation and $b \in \mathbf{R}^n$, we get that

$$TA + \lambda \mathrm{conv}(TA) = M(A + \lambda \mathrm{conv}(A)) + (1 + \lambda)b,$$

which is convex if and only if $A + \lambda \mathrm{conv}(A)$ is convex.

It is easy to see from the definitions that d, r and Δ are translation-invariant, and that d and r are 1-homogeneous and Δ is n-homogeneous with respect to dilation. $\square$

The next lemma concerns the monotonicity of non-convexity measures with respect to the inclusion relation.

Lemma 2.5. *Let* A, B *be compact sets in* $\mathbf{R}^n$ *such that* $A \subset B$ *and* $\mathrm{conv}(A) = \mathrm{conv}(B)$. *Then:*

1. $c(A) \geq c(B)$.
2. $d(A) \geq d(B)$.
3. $r(A) \geq r(B)$.
4. $\Delta(A) \geq \Delta(B)$.

Proof. For the first part, observe that if $\lambda = c(A)$,

$$(1 + \lambda)\mathrm{conv}(B) \supset B + \lambda \mathrm{conv}(B) = B + \lambda \mathrm{conv}(A) \supset A + \lambda \mathrm{conv}(A) = (1 + \lambda)\mathrm{conv}(B),$$

where in the last equation we used that $A + \lambda \mathrm{conv}(A)$ is convex. Hence all relations in the above display must be equalities, and $B + \lambda \mathrm{conv}(B)$ must be convex, which means $c(A) = \lambda \geq c(B)$.

For the second part, observe that

$$d(A) = \sup_{x \in \mathrm{conv}(A)} d(x, A) = \sup_{x \in \mathrm{conv}(B)} d(x, A) \geq \sup_{x \in \mathrm{conv}(B)} d(x, B) = d(B).$$

For the third part, observe that

$$\inf\{R(T) : T \subset A, x \in \mathrm{conv}(T)\} \geq \inf\{R(T) : T \subset B, x \in \mathrm{conv}(T)\}.$$

Hence $r(A) \geq r(B)$.

For the fourth part, observe that

$$\Delta(A) = \mathrm{Vol}_n(\mathrm{conv}(B)) - \mathrm{Vol}_n(A) \geq \mathrm{Vol}_n(\mathrm{conv}(B)) - \mathrm{Vol}_n(B) = \Delta(B).$$

$\square$

As a consequence of Lemma 2.5, we deduce that $A(k)$ is monotone along the subsequence of powers of 2, when measured through all these measures of non-convexity.

Finally we discuss topological aspects of these non-convexity functionals, specifically, whether they have continuity properties with respect to the topology on the class of compact sets induced by Hausdorff distance.

Lemma 2.6. *Suppose $A_k \xrightarrow{d_H} A$, where all the sets involved are compact subsets of $\mathbf{R}^n$. Then:*

1. $\lim_{k\to\infty} d(A_k) = d(A)$, *i.e., d is continuous.*
2. $\liminf_{k\to\infty} \Delta(A_k) \geq \Delta(A)$, *i.e., Δ is lower semicontinuous.*
3. $\liminf_{k\to\infty} c(A_k) \geq c(A)$, *i.e., c is lower semicontinuous.*
4. $\liminf_{k\to\infty} r(A_k) \geq r(A)$, *i.e., r is lower semicontinuous.*

Proof. Let us first observe that $A_k \xrightarrow{d_H} A$ implies $\mathrm{conv}(A_k) \xrightarrow{d_H} \mathrm{conv}(A)$. Indeed, just by applying the convex hull operation to the inclusions $A_k \subset A + \varepsilon B_2^n$ and $A \subset A_k + \varepsilon B_2^n$, and invoking Lemma 2.1 yields this implication.

For the first part, fix an $\varepsilon > 0$ and observe that by the triangle inequality for the Hausdorff metric, we have the inequality

$$\begin{aligned} d(A_k) &= d_H(A_k, \mathrm{conv}(A_k)) \\ &\leq d_H(A_k, A) + d_H(A, \mathrm{conv}(A)) + d_H(\mathrm{conv}(A), \mathrm{conv}(A_k)) \end{aligned}$$

which rewrites as

$$d(A_k) - d(A) \leq d_H(A_k, A) + d_H(\mathrm{conv}(A), \mathrm{conv}(A_k)) \leq \varepsilon,$$

by choosing k large enough. On the other hand, we also have by the triangle inequality that

$$\begin{aligned} d(A) &= d_H(A, \mathrm{conv}(A)) \\ &\leq d_H(A, A_k) + d_H(A_k, \mathrm{conv}(A_k)) + d_H(\mathrm{conv}(A_k), \mathrm{conv}(A)) \end{aligned}$$

which rewrites as

$$d(A) - d(A_k) \leq d_H(A_k, A) + d_H(\mathrm{conv}(A), \mathrm{conv}(A_k)) \leq \varepsilon,$$

by choosing k large enough. Putting these conclusions together yields $d(A_k) - d(A) \to 0$ as $k \to \infty$, which proves the continuity of d.

For the second part, recall that, with respect to the Hausdorff distance, the volume is upper semicontinuous on the class of compact sets (see, e.g., [63, Theorem 12.3.6]) and continuous on the class of compact convex sets (see, e.g., [62, Theorem 1.8.20]). Thus

$$\limsup_{k\to\infty} \mathrm{Vol}_n(A_k) \leq \mathrm{Vol}_n(A)$$

and

$$\lim_{k\to\infty} \mathrm{Vol}_n(\mathrm{conv}(A_k)) = \mathrm{Vol}_n(\mathrm{conv}(A)),$$

so that subtracting the former from the latter yields the desired semicontinuity of Δ.

For the third part, observe that by definition,

$$A_k + \lambda_k \mathrm{conv}(A_k) = (1 + \lambda_k)\mathrm{conv}(A_k),$$

where $\lambda_k = c(A_k)$. Note that from Theorem 2.9 below due to Schneider [61] one has $\lambda_k \in [0, n]$, thus there exists a convergent subsequence $\lambda_{k_n} \to \lambda_*$ and

$$A + \lambda_* \mathrm{conv}(A) = (1 + \lambda_*)\mathrm{conv}(A),$$

Thus $\lambda_* \geq c(A)$, which is the desired semicontinuity of c.

Next we will study convergence of $r(A_k)$. Using $A_k \xrightarrow{d_H} A$ we get that $R(A_k)$ is bounded and thus $r(A_k)$ is bounded and there is a convergent subsequence $r(A_{k_m}) \to l$. Our goal is to show that $r(A) \leq l$. Let $x \in \mathrm{conv}(A)$. Then there exits $x_m \in A_{k_m}$ such that $x_m \to x$. From the definition of $r(A_{k_m})$ we get that there exists $T_m \subset A_{k_m}$ such that $x_m \in \mathrm{conv}(T_m)$ and $R(T_m) \leq r(A_{k_m})$. We can select a convergent subsequence $T_{m_i} \to T$, where T is compact (see [62, Theorem 1.8.4]), then $T \subset A$ and $x \in \mathrm{conv}(T)$ and $R(T_{m_i}) \to R(T)$ therefore $R(T) \leq l$. Thus $r(A) \leq l$. $\square$

We emphasize that the semicontinuity assertions in Lemma 2.6 are not continuity assertions for a reason and even adding the assumption of nestedness of the sets would not help.

Example 2.7. *Schneider [61] observed that c is not continuous with respect to the Hausdorff distance, even if restricted to the compact sets with nonempty interior. His example consists of taking a triangle in the plane, and replacing one of its edges by the two segments which join the endpoints of the edge to an interior point (see Figure 1). More precisely, let* $a_k = (\frac{1}{2} - \frac{1}{k}, \frac{1}{2} - \frac{1}{k})$, $A_k = \mathrm{conv}((0,0);(1,0);a_k) \cup \mathrm{conv}((0,0);(0,1);a_k)$, *and* $A = \mathrm{conv}((0,0);(0,1);(1,0))$. *Then* $d_H(A_k, A) \to 0$. *But one can notice that* $c(A_k) = 1$ *and* $r(A_k) = 1/\sqrt{2}$ *with* $r(A) = c(A) = 0$.

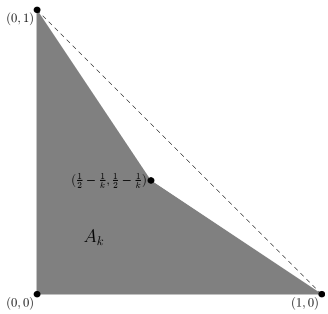

Figure 1: Discontinuity of c and r with respect to Hausdorff distance.

Example 2.8. *To see that there is no continuity for Δ, consider a sequence of discrete nested sets converging in d to $[0,1]$, more precisely:* $A_k = \{\frac{m}{2^k}; 0 \leq m \leq 2^k\}$.

2.3 Special properties of Schneider's index

All these functionals other than c can be unbounded. The boundedness of c follows for the following nice inequality due to Schneider [61].

Theorem 2.9. *[61] For any subset A of $\mathbf{R}^n$,*

$$c(A) \leq n.$$

Proof. Applying the Shapley-Folkman lemma (Lemma 2.2) to $A_1 = \cdots = A_{n+1} = A$, where $A \subset \mathbf{R}^n$ is a fixed compact set, one deduces that $(n+1)\mathrm{conv}(A) = A + n\mathrm{conv}(A)$. Thus $c(A) \leq n$. $\square$

Schneider [61] showed that $c(A) = n$ if and only if A consists of $n+1$ affinely independent points. Schneider also showed that if A is unbounded or connected, one has the sharp bound $c(A) \leq n-1$.

Let us note some alternative representations of Schneider's non-convexity index. First, we would like to remind the definition of the *Minkowski functional* of a compact convex set K containing zero:

$$\|x\|_K = \inf\{t > 0 : x \in tK\},$$

with the usual convention that $\|x\|_K = +\infty$ if $\{t > 0 : x \in tK\} = \emptyset$. Note that $K = \{x \in \mathbf{R}^n : \|x\|_K \leq 1\}$ and $\|x\|_K$ is a norm if K is symmetric with non empty interior.

For any compact set $A \subset \mathbf{R}^n$, define

$$A_\lambda = \frac{1}{1+\lambda}[A + \lambda \operatorname{conv}(A)],$$

and observe that

$$\operatorname{conv}(A_\lambda) = \frac{1}{1+\lambda}\operatorname{conv}(A + \lambda \operatorname{conv}(A)) = \frac{1}{1+\lambda}[\operatorname{conv}(A) + \lambda\operatorname{conv}(A)] = \operatorname{conv}(A).$$

Hence, we can express

$$c(A) = \inf\{\lambda \geq 0 : A_\lambda \text{ is convex}\} = \inf\{\lambda \geq 0 : A_\lambda = \operatorname{conv}(A)\}. \tag{9}$$

Rewriting this yet another way, we see that if $c(A) < t$, then for each $x \in \operatorname{conv}(A)$, there exists $a \in A$ and $b \in \operatorname{conv}(A)$ such that

$$x = \frac{a + tb}{1+t},$$

or equivalently, $x - a = t(b - x)$. In other words, $x - a \in tK_x$ where $K_x = \operatorname{conv}(A) - x$, which can be written as $\|x - a\|_{K_x} \leq t$ using the Minkowski functional. Thus

$$c(A) = \sup_{x \in \operatorname{conv}(A)} \inf_{a \in A} \|x - a\|_{K_x}$$

This representation is nice since it allows for comparison with the representation of $d(A)$ in the same form but with K_x replaced by the Euclidean unit ball.

Remark 2.10. *Schneider [61] observed that there are many closed unbounded sets $A \subset \mathbf{R}^n$ that satisfy $c(A) = 0$, but are not convex. Examples he gave include the set of integers in $\mathbf{R}$, or a parabola in the plane. This makes it very clear that if we are to use c as a measure of non-convexity, we should restrict attention to compact sets.*

2.4 Unconditional relationships

It is natural to ask how these various measures of non-convexity are related. First we note that d and $d^{(K)}$ are equivalent. To prove this we would like to present an elementary but useful observation:

Lemma 2.11. *Let $K \subset \mathbf{R}^n$ be an arbitrary convex body containing 0 in its interior. Consider a convex body $L \subset \mathbf{R}^n$ such that $K \subset L$ and $t > 0$. Then for any compact set $A \subset \mathbf{R}^n$,*

$$d^{(K)}(A) \geq d^{(L)}(A)$$

and

$$d^{(tK)}(A) = \frac{1}{t} d^{(K)}(A).$$

Proof. Notice that
$$A + d^{(K)}(A)L \supset A + d^{(K)}(A)K \supset \mathrm{conv}(A).$$
Hence, $d^{(K)}(A) \geq d^{(L)}(A)$. In addition, one has
$$A + d^{(K)}(A)K = A + \frac{1}{t}d^{(K)}(A)tK.$$
Hence, $d^{(tK)}(A) = \frac{1}{t}d^{(K)}(A)$. □

The next lemma follows immediately from Lemma 2.11:

Lemma 2.12. *Let K be an arbitrary convex body containing 0 in its interior. For any compact set $A \subset \mathbf{R}^n$, one has*
$$rd^{(K)}(A) \leq d(A) \leq Rd^{(K)}(A),$$
where $r, R > 0$ are such that $rB_2^n \subset K \subset RB_2^n$.

It is also interesting to note a special property of $d^{(\mathrm{conv}(A))}(A)$:

Lemma 2.13. *Let A be a compact set in $\mathbf{R}^n$. If $0 \in \mathrm{conv}(A)$, then*
$$d^{(\mathrm{conv}(A))}(A) \leq c(A).$$
If $0 \in A$, then
$$d^{(\mathrm{conv}(A))}(A) \leq \min\{1, c(A)\}.$$

Proof. If $0 \in \mathrm{conv}(A)$, then $\mathrm{conv}(A) \subset (1 + c(A))\mathrm{conv}(A)$. But,
$$(1 + c(A))\mathrm{conv}(A) = A + c(A)\mathrm{conv}(A),$$
where we used the fact that by definition of $c(A)$, $A + c(A)\mathrm{conv}(A)$ is convex. Hence, $d^{(\mathrm{conv}(A))}(A) \leq c(A)$.

If $0 \in A$, in addition to the above argument, we also have
$$\mathrm{conv}(A) \subset A + \mathrm{conv}(A).$$
Hence, $d^{(\mathrm{conv}(A))}(A) \leq 1$. □

Note that the inequality in the above lemma cannot be reversed even with the cost of an additional multiplicative constant. Indeed, take the sets A_k from Example 2.7, then $c(A_k) = 1$ but $d^{(\mathrm{conv}(A_k))}(A_k)$ tends to 0.

Observe that d, r, ρ and v have some similarity in definition. Let us introduce the point-wise definitions of above notions: Consider $x \in \mathrm{conv}(A)$, define

- $d_A(x) = \inf_{a \in A} |x - a|$.
  More generally, if K is a compact convex set in $\mathbf{R}^n$ containing the origin,
- $d_A^{(K)}(x) = \inf_{a \in A} \|x - a\|_K$.
- $r_A(x) = \inf\{R(T) : T \subset A, x \in \mathrm{conv}(T)\}$.
- $v_A^2(x) = \inf\{\sum p_i|a_i|^2 - |x|^2 : x = \sum p_i a_i; p_i > 0; \sum p_i = 1, a_i \in A\}$.

- $\rho_A(x) = \inf_{a \in A_x} |x - a|$, where

$$A_x = \{a \in A : \exists b \in \mathrm{conv}(A), \exists \theta \in (0,1) \text{ such that } x = (1-\theta)a + \theta b\}.$$

Below we present a Theorem due to Wegmann [75] which proves that r, ρ and v are equal for compact sets and that they are equal also to d under an additional assumption. For the sake of completeness we will present the proof of Wegmann [75] which is simplified here for the case of compact sets.

Theorem 2.14 (Wegmann [75]). *Let A be a compact set in $\mathbf{R}^n$, then*

$$d(A) \leq \rho(A) = v(A) = r(A).$$

Moreover if $v_A(x_0) = v(A)$, for some x_0 in the relative interior of $\mathrm{conv}(A)$*, then $d(A) = v(A) = r(A) = \rho(A)$.*

Proof. 1) First observe that $d(A) \leq \rho(A) \leq v(A) \leq r(A)$ by easy arguments; in fact, this relation holds point-wise, i.e. $d_A(x) \leq \rho_A(x) \leq v_A(x) \leq r_A(x)$, indeed the first inequality follows directly from the definitions, because $A_x \subset A$. To prove the second inequality consider any convex decomposition of $x \in \mathrm{conv}(A)$, i.e. $x = \sum p_i a_i; p_i > 0; \sum p_i = 1, a_i \in A$, without loss of generality we may assume that $|x - a_1| \leq |x - a_i|$ for all $i \leq m$. Then

$$\sum p_i |x - a_i|^2 \geq |x - a_1|^2 \geq \rho_A^2(x),$$

because $a_1 \in A_x$ (indeed, $x = p_1 a_1 + (1 - p_1) \sum\limits_{i \geq 2} \frac{p_i}{1-p_1} a_i$). To prove the third inequality let $T = \{a_1, \ldots, a_m\} \subset A$ be such that $x \in \mathrm{conv}(T)$. Let $p_1, \ldots, p_m > 0$ be such that $\sum p_i = 1$ and $x = \sum p_i a_i$. Let c be the center of the smallest Euclidean ball containing T. Notice that the minimum of $\sum p_i |x - a_i|^2$ is reached for $x = \sum p_i a_i$, thus

$$v_A^2(x) \leq \sum p_i |x - a_i|^2 \leq \sum p_i |c - a_i|^2 \leq R^2(T),$$

and we take infimum over all T to finish the proof of the inequality.

2) Consider $x_0 \in \mathrm{conv}(A)$. To prove the theorem we will first show that $r_A(x_0) \leq v(A)$. After this we will show that $v_A(x_0) \leq \rho(A)$ and finally we will prove if x_0 is in the relative interior of $\mathrm{conv}(A)$ and maximizes $v_A(x)$, among $x \in \mathrm{conv}(A)$ then $d_A(x_0) \geq v(A)$.

2.1) Let us prove that $r_A(x_0) \leq v(A)$. Assume first that x_0 is an interior point of $\mathrm{conv}(A)$. Let us define the compact convex set $Q \subset \mathbf{R}^{n+1}$ by

$$Q = \mathrm{conv}\{(a, |a|^2); a \in A\}.$$

Next we define the function $f : \mathrm{conv}(A) \to \mathbf{R}^+$ by $f(x) = \min\{y : (x,y) \in Q\}$, note that

$$\begin{aligned} f(x) &= \min\{y : (x,y) = \sum \lambda_i (a_i, |a_i|^2); \lambda_1, \ldots, \lambda_m > 0 \text{ and } a_1, \ldots, a_m \in A\} \\ &= \min\{\sum \lambda_i |a_i|^2 : \lambda_1, \ldots, \lambda_m > 0 \text{ and } a_1, \ldots, a_m \in A, \sum \lambda_i = 1; x = \sum \lambda_i a_i\} \\ &= v_A^2(x) + |x|^2. \end{aligned}$$

Note that $(x_0, f(x_0))$ is a boundary point of Q hence there exists a support hyperplane H of Q at $(x_0, f(x_0))$. Since x_0 is an interior point of $\mathrm{conv}(A)$, the hyperplane H cannot be vertical because a vertical support plane would separate x_0 from boundary points of $\mathrm{conv}(A)$ and thus separate $(x_0, f(x_0))$ from boundary points of Q. Thus there exist $b \in \mathbf{R}^n$ and $\alpha \in \mathbf{R}$ such that $H = \{(x,t) \in \mathbf{R}^{n+1} : t = 2\langle b, x\rangle + \alpha\}$. Since $(x_0, f(x_0)) \in H$ one has

$$f(x_0) = 2\langle b, x_0\rangle + \alpha \tag{10}$$

and

$$f(x) \geq 2\langle b, x\rangle + \alpha, \text{ for all } x \in \mathrm{conv}(A).$$

By definition of f, there exists $a_1, \ldots, a_m \in A$ and $\lambda_1, \ldots, \lambda_m > 0$, $\sum \lambda_i = 1$ such that $x_0 = \sum \lambda_i a_i$ and

$$f(x_0) = \sum \lambda_i |a_i|^2 = \sum \lambda_i f(a_i).$$

From the convexity of Q we get that $(a_i, f(a_i)) \in H \cap Q$, for any i; indeed we note that

$$f(x_0) = 2\langle b, x_0\rangle + \alpha = \sum_i \lambda_i (2\langle b, a_i\rangle + \alpha) \leq \sum_i \lambda_i f(a_i) = f(x_0).$$

Thus $2\langle b, a_i\rangle + \alpha = f(a_i)$ for all i. Let $T = \{a_1, \ldots a_m\}$ and $W = \mathrm{conv}(T)$. Note that for any $x \in W \cap A$ we have

$$|x|^2 = f(x) = 2\langle b, x\rangle + \alpha$$

thus $\alpha + |b|^2 = |x - b|^2 \geq 0$. Define

$$R^2 = \alpha + |b|^2. \tag{11}$$

Notice that for any $x \in \mathrm{conv}(A)$

$$v_A^2(x) = f(x) - |x|^2 \geq 2\langle b, x\rangle + \alpha - |x|^2 = R^2 - |b - x|^2, \tag{12}$$

with equality if $x \in W$, in particular, $0 \leq v_A^2(x) = R^2 - |b - x|^2 \leq R^2$, for every $x \in W$. Consider the point $w \in W$ such that

$$v_A^2(w) = \max_{x \in W} v_A^2(x) = \max_{x \in W} (R^2 - |b - x|^2) = R^2 - \inf_{x \in W} |b - x|^2.$$

Then one has $|b - w| = \inf_{x \in W} |b - x|$, which means w is the projection of the point b on the convex set W. This implies that, for every $x \in W$, one has $\langle x - b, w - b\rangle \geq |w - b|^2$, thus

$$|x - w|^2 = |x - b|^2 - 2\langle x - b, w - b\rangle + |w - b|^2 \leq |x - b|^2 - |w - b|^2 \leq R^2 - |w - b|^2 = v_A^2(w).$$

We get $T \subset W \subset w + v_A(w) B_2^n$ and

$$R(T) \leq v_A(w) = \max_{x \in W} v_A(x).$$

Using that $x_0 \in W = \mathrm{conv}(T)$ and $T \subset A$, we conclude from the definition of r_A that

$$r_A(x_0) \leq R(T) \leq \max_{x \in W} v_A(x) \leq v(A).$$

If x_0 is a boundary point of $\mathrm{conv}(A)$, then using the boundary structure of the polytope $\mathrm{conv}(A)$ (see [62, Theorem 2.1.2, p. 75 and Remark 3, p. 78]) x_0 belongs to the relative interior of an exposed face F of $\mathrm{conv}(A)$. By the definition of the notion of exposed face (see

[62, p. 75]) we get that if $x = \sum \lambda_i a_i$ for $a_i \in A$ and $\lambda_i > 0$ with $\sum \lambda_i = 1$, then $a_i \in A \cap F$. Thus

$$v_A(x_0) = v_{A\cap F}(x_0), r_A(x_0) = r_{A\cap F}(x_0) \text{ and } \rho_A(x_0) = \rho_{A\cap F}(x_0). \tag{13}$$

If $\dim(F) = 0$ then $x_0 \in A$ and thus all proposed inequalities are trivial, otherwise we can reproduce the above argument for $A \cap F$ instead of A.

2.2) Now we will prove that $v_A(x_0) \leq \rho(A)$. Consider b, α and R defined in (10) and (11). Using that $v_A(a) = 0$, for every $a \in A$ and (12), we get $|b - a| \geq R$, for all $a \in A$. We will need to consider two cases

1. If $b \in \mathrm{conv}(A)$, then from the above $d_A(b) = \inf_{a\in A} |b - a| \geq R$ thus

$$v_A(x_0) \leq R \leq d_A(b) \leq \rho_A(b) \leq \rho(A). \tag{14}$$

2. If $b \notin \mathrm{conv}(A)$, then there exists $y \in \partial(\mathrm{conv}(A)) \cap [w, b]$, thus $|b - y| \leq |b - w|$. So, from (12) we have

$$v_A^2(y) \geq R^2 - |b - y|^2 \geq R^2 - |b - w|^2 = v_A^2(w) \geq v_A^2(x_0),$$

so it is enough to prove $v_A(y) \leq \rho(A)$, where $y \in \partial(\mathrm{conv}(A))$. Let F be the face of $\mathrm{conv}(A)$ containing y in its relative interior. Thus we can use the approach from (13) and reproduce the above argument for $A \cap F$ instead of A, in the end of which we will again get two cases (as above), in the first case we get $v_A(y) = v_{A\cap F}(y) \leq \rho(A \cap F) \leq \rho(A)$. In the second case, there exists $z \in \partial(\mathrm{conv}(A \cap F))$ such that $v_{A\cap F}(z) \geq v_{A\cap F}(y)$ and we again reduce the dimension of the set under consideration. Repeating this argument we will arrive to the dimension 1 in which the proof can be completed by verifying that $b \in \mathrm{conv}(A)$ (indeed, in this case $W = [a_1, a_2]$, $a_1, a_2 \in A$ and $|a_1 - b| = |a_2 - b|$, thus $b = (a_1 + a_2)/2 \in \mathrm{conv}(A)$) and thus $v_A(x_0) \leq \rho(A)$.

2.3) Finally, assume $v_A(x_0) = v(A)$, where x_0 is in the relative interior of $\mathrm{conv}(A)$. We may assume that $\mathrm{conv}(A)$ is n-dimensional (otherwise we would work in the affine subspace generated by A). Then using (12) we get that $v_A^2(x_0) = R^2 - |b - x_0|^2$ and $v_A^2(a) \geq R^2 - |b - a|^2$, for all $a \in \mathrm{conv}(A)$, thus

$$0 \leq v_A^2(x_0) - v_A^2(a) \leq |b - a|^2 - |b - x_0|^2,$$

for all $a \in \mathrm{conv}(A)$. So $|b - x_0| \leq |b - a|$ for all $a \in \mathrm{conv}(A)$, this means that the minimal distance between b and $a \in \mathrm{conv}(A)$ is reached at $a = x_0$. Notice that if $b \notin \mathrm{conv}(A)$ then x_0 must belong to $\partial(\mathrm{conv}(A))$, which contradicts our hypothesis. Thus $b \in \mathrm{conv}(A)$ and $x_0 = b$, and we can use (14) to conclude that $v(A) = v_A(x_0) \leq d_A(x_0) \leq d(A)$. □

Remark 2.15. *Let us note that the method used in the proof of Theorem 2.14 is reminiscent of the classical approach to Voronoi diagram and Delaunay triangulation, see for example [55, section 5.7]. Moreover the point b constructed above is exactly the center of the ball of the Delaunay triangulation to which the point x_0 belongs.*

The above relationships (summarized in Table 1) are the only unconditional relationships that exist between these notions in general dimension. To see this, we list below some examples that show why no other relationships can hold in general.

$\Rightarrow$	d	r	c	Δ
d	$=$	N (Ex. 2.7, 2.17)	N (Ex. 2.7, 2.16)	N (Ex. 2.8)
r	Y (Th. 2.14)	$=$	N (Ex. 2.16, 2.28)	N (Ex. 2.8)
c	N (Ex. 2.16, 2.19)	N (Ex. 2.16)	$=$	N (Ex. 2.8, 2.16)
Δ	N (Ex. 2.18, 2.19)	N (Ex. 2.7, 2.17)	N (Ex. 2.16, 2.28)	$=$

Table 1: When does convergence to 0 for one measure of non-convexity unconditionally imply the same for another?

Example 2.16. *By Lemma 2.4, we can scale a non convex set to get examples where* c *is fixed but* d, r *and* Δ *converge to 0, for example, take* $A_k = \{0; \frac{1}{k}\}$*; or to get examples where* c *goes to 0 but* d, r *are fixed and* Δ *diverges, for example take* $A_k = \{0, 1, \ldots, k\}$*.*

Example 2.17. *An example where* $\Delta(A_k) \to 0$*,* $d(A_k) \to 0$ *but* $r(A_k)$ *is bounded away from 0 is given by a right triangle from which a piece is shaved off leaving a protruding edge, see figure 1.*

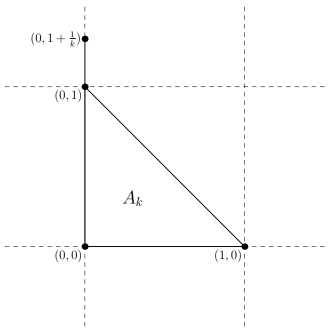

Figure 2: $\Delta(A_k) \to 0$ but $r(A_k) > \sqrt{2}/2$.

Example 2.18. *An example where* $\Delta(A_k) \to 0$ *but both* $c(A_k)$ *and* $d(A_k)$ *are bounded away from 0 is given by taking a 3-point set with 2 of the points getting arbitrarily closer but staying away from the third, see figure 3.*

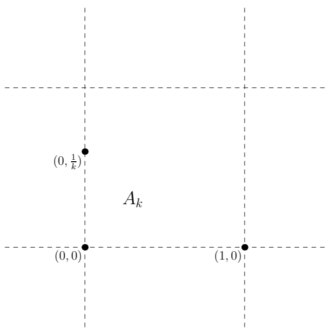

Figure 3: $\Delta(A_k) \to 0$ but $c(A_k) \geq 1$ and $d(A_k) \geq 1/2$.

Example 2.19. *An example where $\Delta(A_k) \to 0$ and $c(A_k) \to 0$ but $d(A_k) > 1/2$ can be found in figure 4.*

Figure 4: $\mathrm{Vol}_2(A_k) \geq 1$, $\Delta(A_k) \to 0$ and $c(A_k) \to 0$ but $d(A_k) > 1/2$.

2.5 Conditional relationships

There are some more relationships between different notions of non-convexity that emerge if we impose some natural conditions on the sequence of sets (such as ruling out escape to infinity, or vanishing to almost nothing).

A first observation of this type is that Hausdorff distance to convexity is dominated by Schneider's index of non-convexity if A is contained in a ball of known radius.

Lemma 2.20. *For any compact set $A \subset \mathbf{R}^n$,*

$$d(A) \leq R(A)c(A). \tag{15}$$

Proof. By translation invariance, we may assume that $A \subset R(A)B_2^n$. Then $0 \in \mathrm{conv}(A)$, and it follows that

$$\mathrm{conv}(A) \subset \mathrm{conv}(A) + c(A)\mathrm{conv}(A) = A + c(A)\mathrm{conv}(A) \subset A + c(A)R(A)B_2^n.$$

Hence $d(A) \leq R(A)c(A)$. □

This bound is useful only if $c(A)$ is smaller than 1, because we already know that $d(A) \leq r(A) \leq R(A)$.

In dimension 1, all of the non-convexity measures are tightly connected.

Lemma 2.21. *Let A be a compact set in $\mathbf{R}$. Then*

$$r(A) = d(A) = R(A)c(A) \leq \frac{\Delta(A)}{2}. \tag{16}$$

Proof. We already know that $d(A) \leq r(A)$. Let us prove that $r(A) \leq d(A)$. From the definition of $r(A)$ and $d(A)$, we have

$$r(A) = \sup_{x \in \mathrm{conv}(A)} \inf \left\{ \frac{\beta - \alpha}{2}; \alpha, \beta \in A, \alpha \leq x \leq \beta \right\}, \quad d(A) = \sup_{y \in \mathrm{conv}(A)} \inf_{\alpha \in A} |y - \alpha|.$$

Thus we only need to show that for every $x \in \mathrm{conv}(A)$, there exists $y \in \mathrm{conv}(A)$ such that

$$\inf \left\{ \frac{\beta - \alpha}{2}; \alpha, \beta \in A, \alpha \leq x \leq \beta \right\} \leq \inf_{\alpha \in A} |y - \alpha|.$$

By compactness there exists $\alpha, \beta \in A$, with $\alpha \leq x \leq \beta$ achieving the infimum in the left hand side. Then we only need to choose $y = \frac{\alpha+\beta}{2}$ in the right hand side to conclude that $r(A) \leq d(A)$. In addition, we get $(\alpha, \beta) \subset \mathrm{conv}(A) \setminus A$ thus $2r(A) = \beta - \alpha \leq \Delta(A)$.

Now we prove that $d(A) = R(A)c(A)$. From Lemma 2.20, we have $d(A) \leq R(A)c(A)$. Let us prove that $R(A)c(A) \leq d(A)$. By an affine transform, we may reduce to the case where $\mathrm{conv}(A) = [-1, 1]$, thus $-1 = \min(A) \in A$ and $1 = \max(A) \in A$. Notice that $R(A) = 1$ and denote $d := d(A)$. By the definition of $d(A)$, one has $[-1, 1] = \mathrm{conv}(A) \subset A + [-d, d]$. Thus using that $-1 \in A$ and $1 \in A$, we get

$$A + d(A)\mathrm{conv}(A) = A + [-d, d] \supset (-1 + [-d, d]) \cup [-1, 1] \cup (1 + [-d, d]) = [-1 - d, 1 + d],$$

we conclude that $A + d(A)\mathrm{conv}(A) \supset (1 + d(A))\mathrm{conv}(A)$ and thus $R(A)c(A) = c(A) \leq d(A)$. $\square$

Notice that the inequality on Δ of Lemma 2.21 cannot be reversed as shown by Example 2.8. The next lemma provides a connection between r and c in $\mathbf{R}^n$.

Lemma 2.22. *For any compact set $A \subset \mathbf{R}^n$,*

$$r(A) \leq 2\frac{c(A)}{1 + c(A)}R(A). \tag{17}$$

Proof. Consider x^* the point in $\mathrm{conv}(A)$ that realizes the maximum in the definition of $\rho(A)$ (it exists since $\mathrm{conv}(A)$ is closed). Then, for every $a \in A_{x^*}$, one has $\rho(A) \leq |x^* - a|$. By definition,

$$c(A) = \inf\{\lambda \geq 0 : \mathrm{conv}(A) = \frac{A + \lambda\mathrm{conv}(A)}{1 + \lambda}\}.$$

Hence,

$$x^* = \frac{1}{1 + c(A)}a + \frac{c(A)}{1 + c(A)}b,$$

for some $a \in A$ and $b \in \mathrm{conv}(A)$. Since $\frac{1}{1+c(A)} + \frac{c(A)}{1+c(A)} = 1$, one deduces that $a \in A_{x^*}$. Thus,

$$\rho(A) \leq |x^* - a|.$$

But,

$$x^* - a = \frac{1}{1 + c(A)}a + \frac{c(A)}{1 + c(A)}b - a = \frac{c(A)}{1 + c(A)}(b - a).$$

It follows that

$$\rho(A) \leq |x^* - a| = \frac{c(A)}{1 + c(A)}|b - a| \leq \frac{c(A)}{1 + c(A)}\mathrm{diam}(A) \leq 2\frac{c(A)}{1 + c(A)}R(A).$$

As shown by Wegmann (cf. Theorem 2.14), if A is closed then $\rho(A) = r(A)$. We conclude that

$$r(A) \leq 2\frac{c(A)}{1 + c(A)}R(A).$$

$\square$

Our next result says that the only reason for which we can find examples where the volume deficit goes to 0, but the Hausdorff distance from convexity does not, is because we allow the sets either to shrink to something of zero volume, or run off to infinity.

Theorem 2.23. *Let A be a compact set in $\mathbf{R}^n$ with nonempty interior. Then*

$$d(A) \leq \left(\frac{n}{\mathrm{Vol}_{n-1}(B_2^{n-1})} \right)^{\frac{1}{n}} \left(\frac{2R(A)}{\mathrm{inr}(\mathrm{conv}(A))} \right)^{\frac{n-1}{n}} \Delta(A)^{\frac{1}{n}}. \tag{18}$$

Proof. From the definition of $d(A)$ there exists $x \in \mathrm{conv}(A)$ such that $\mathrm{Vol}_n((x + d(A)B_2^n) \cap A) = 0$. Thus $\Delta(A) \geq \mathrm{Vol}_n(\mathrm{conv}(A) \cap (x + d(A)B_2^n))$. Let us denote $r = \mathrm{inr}(\mathrm{conv}(A))$. From the definition of $\mathrm{inr}(\mathrm{conv}(A))$, there exists $y \in \mathrm{conv}(A)$ such that $y + rB_2^n \subset \mathrm{conv}(A)$. Hence

$$\Delta(A) \geq \mathrm{Vol}_n(\mathrm{conv}(x, y + rB_2^n) \cap (x + d(A)B_2^n)) \geq \frac{1}{n}\mathrm{Vol}_{n-1}(B_2^{n-1}) \left(\frac{rd(A)}{2R(A)} \right)^{n-1}.$$

Let $\{z\} = [x, y] \cap (x + d(A)S^{n-1})$ be the intersection point of the sphere centered at x and the segment $[x, y]$ and let h be the radius of the $(n-1)$-dimensional sphere $S_h = \partial(\mathrm{conv}(x, y + rB_2^n)) \cap (x + d(A)S^{n-1}))$. Then $h = \frac{d(A)r}{|x-y|}$ and $\mathrm{conv}(x, y + rB_2^n) \cap (x + d(A)B_2^n) \supset \mathrm{conv}(x, S_h, z)$. Thus

$$\Delta(A) \geq \mathrm{Vol}_n(\mathrm{conv}(x, S_h, z)) = \frac{d(A)}{n}\mathrm{Vol}_{n-1}(B_2^{n-1})h^{n-1} \geq \frac{d(A)^n}{n}\mathrm{Vol}_{n-1}(B_2^{n-1}) \left(\frac{r}{|x-y|} \right)^{n-1}.$$

$\square$

Observe that the first term on the right side in inequality (18) is just a dimension-dependent constant, while the second term depends only on the ratio of the radius of the smallest Euclidean ball containing A to that of the largest Euclidean ball inside it.

The next lemma enables to compare the inradius, the outer radius and the volume of convex sets. Such estimates were studied in [22], [59] where, in some cases, optimal inequalities were proved in dimension 2 and 3.

Lemma 2.24. *Let K be a convex body in $\mathbf{R}^n$. Then*

$$\mathrm{Vol}_n(K) \leq (n+1)\mathrm{Vol}_{n-1}(B_2^{n-1})\mathrm{inr}(K)(2R(K))^{n-1}.$$

Proof. From the definition of $\mathrm{inr}(K)$, there exists $y \in K$ such that $y + \mathrm{inr}(K)B_2^n \subset K$. Without loss of generality, we may assume that $y = 0$ and that $\mathrm{inr}(K) = 1$, which means that B_2^n is the Euclidean ball of maximal radius inside K. This implies that 0 must be in the convex hull of the contact points of S^{n-1} and $\partial(K)$, because if it is not, then there exists an hyperplane separating 0 from these contact points and one may construct a larger Euclidean ball inside K. Hence from Caratheodory, there exists $1 \leq k \leq n$ and $k+1$ contact points $a_1, \dots, a_{k+1}$ so that $0 \in \mathrm{conv}(a_1, \dots, a_{k+1})$ and $K \subset S = \{x : \langle x, a_i \rangle \leq 1, \forall i \in \{1, \dots, k+1\}\}$. Since $0 \in \mathrm{conv}(a_1, \dots, a_{k+1})$, there exists $\lambda_1, \dots, \lambda_{k+1} \geq 0$ such that $\sum_{i=1}^{k+1} \lambda_i a_i = 0$. Thus for every $x \in \mathbf{R}^n$, $\sum_{i=1}^{k+1} \lambda_i \langle x, a_i \rangle = 0$ hence there exists i such that $\langle x, a_i \rangle \geq 0$. Hence

$$S \subset \bigcup_{i=1}^{k+1} [0, a_i] \times \{x : \langle x, a_i \rangle = 0\}.$$

Moreover $K \subset \mathrm{diam}(K)B_2^n$ thus

$$K \subset S \cap \mathrm{diam}(K)B_2^n \subset \bigcup_{i=1}^{k+1} [0, a_i] \times \{x \in \mathrm{diam}(K)B_2^n : \langle x, a_i \rangle = 0\}.$$

Passing to volumes and using that $a_i \in S^{n-1}$, we get

$$\mathrm{Vol}_n(K) \leq (k+1)\mathrm{Vol}_{n-1}(B_2^{n-1})(\mathrm{diam}(K))^{n-1} \leq (n+1)\mathrm{Vol}_{n-1}(B_2^{n-1})(2R(K))^{n-1}.$$

□

An immediate corollary of the above theorem and lemma is the following.

Corollary 2.25. *Let A be a compact set in $\mathbf{R}^n$. Then*

$$d(A) \leq c_n \frac{R(A)^{n-1}}{\mathrm{Vol}_n(\mathrm{conv}(A))^{\frac{n-1}{n}}} \Delta(A)^{\frac{1}{n}},$$

where c_n is an absolute constant depending on n only. Thus for any sequence of compact sets (A_k) in $\mathbf{R}^n$ such that $\sup_k R(A_k) < \infty$ and $\inf_k \mathrm{Vol}_n(A_k) > 0$, the convergence $\Delta(A_k) \to 0$ implies that $d(A_k) \to 0$.

$\Rightarrow$	d	r	c	Δ
d	=	N (Ex. 2.7, 2.27)	N (Ex. 2.7, 2.27)	N (Ex. 2.26)
r	Y	=	N (Ex. 2.28)	N (Ex. 2.26)
c	Y (Lem. 2.20)	Y (Lem. 2.22)	=	N (Ex. 2.26)
Δ	Y (Cor. 2.25)	N (Ex. 2.7, 2.27)	N (Ex. 2.7, 2.27)	=

Table 2: When does convergence to 0 for one measure of non-convexity imply the same for another when we assume the sequence lives in a big ball and has positive limiting volume?

From the preceding discussion, it is clear that $d(A_k) \to 0$ is a much weaker statement than either $c(A_k) \to 0$ or $r(A_k) \to 0$.

Example 2.26. *Consider a unit square with a set of points in the neighboring unit square, where the set of points becomes more dense as $k \to \infty$ (see Figure 5). This example shows that the convergence in the Hausdorff sense is weaker than convergence in the volume deficit sense even when the volume of the sequence of sets is bounded away from 0.*

The following example shows that convergence in Δ does not imply convergence in r nor c:

Example 2.27. *Consider the set $A_k = \{(1 - \frac{1}{k}, 0)\} \cup ([1, 2] \times [-1, 1])$ in the plane.*

Note that the Example 2.27 also shows that convergence in d does not imply convergence in r nor c. The following example shows that convergence in r does not imply convergence in c:

Example 2.28. *Consider the set $A_k = B_2^2 \cup \{(1 + 1/k, 1/k); (1 + 1/k, -1/k)\}$ in the plane, the union of the Euclidean ball and two points close to it and close to each other (see Figure 6). Then we have $c(A_k) = 1$ because taking an homothety of scale 1/2 is the best you can do if you want to cover the point $(1 + 1/k, 0)$ with an homothety with center in a point of A_k. But for $r(A_k)$, we see that because of the roundness of the ball, one has $r(A_k) = \frac{\sqrt{k+1}}{\sqrt{2}k} \to 0$, when k grows.*

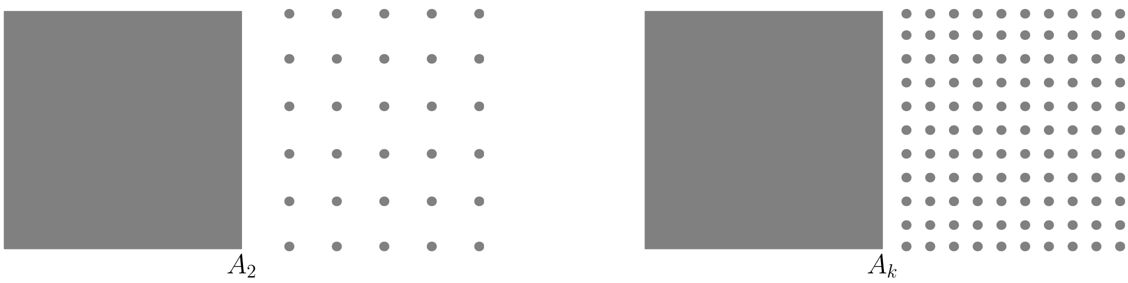

Figure 5: $d(A_k) \to 0$ and $\mathrm{Vol}_2(A_k) > c$ but $\Delta(A_k) > c$.

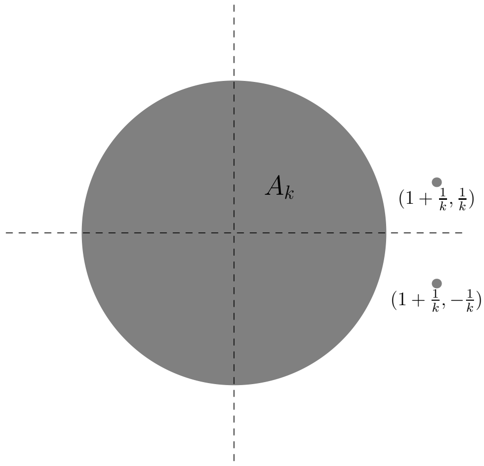

Figure 6: $c(A_k) = 1$ but $r(A_k) \to 0$, when k grows.

3 The behavior of volume deficit

3.1 Monotonicity of volume deficit in dimension one and for Cartesian products

In this section, we observe that Conjecture 1.2 holds in dimension one and also for products of one-dimensional compact sets. In fact, more generally, we prove that conjecture 1.2 passes to Cartesian product.

Theorem 3.1. *Conjecture 1.2 holds in dimension one,* i.e. *let $k \geq 2$ be an integer and let $A_1, \dots, A_k$ be k compact sets in $\mathbf{R}$. Then*

$$\mathrm{Vol}_1\left(\sum_{i=1}^k A_i\right) \geq \frac{1}{k-1}\sum_{i=1}^k \mathrm{Vol}_1\left(\sum_{j\in[k]\setminus\{i\}} A_j\right). \tag{19}$$

Proof. We adapt a proof of Gyarmati, Matolcsi and Ruzsa [40, Theorem 1.4] who established the same kind of inequality for finite subsets of the integers and cardinality instead of volume. The proof is based on set inclusions. Let $k \geq 1$. Set $S = A_1 + \cdots + A_k$ and for $i \in [k]$, let $a_i = \min A_i$, $b_i = \max A_i$,

$$S_i = \sum_{j\in[k]\setminus\{i\}} A_j,$$

$s_i = \sum_{j<i} a_j + \sum_{j>i} b_j$, $S_i^- = \{x \in S_i; x \le s_i\}$ and $S_i^+ = \{x \in S_i; x > s_i\}$. For all $i \in [k-1]$, one has

$$S \supset (a_i + S_i^-) \cup (b_{i+1} + S_{i+1}^+).$$

Since $a_i + s_i = \sum_{j \le i} a_j + \sum_{j>i} b_j = b_{i+1} + s_{i+1}$, the above union is a disjoint union. Thus for $i \in [k-1]$

$$\mathrm{Vol}_1(S) \ge \mathrm{Vol}_1(a_i + S_i^-) + \mathrm{Vol}_1(b_{i+1} + S_{i+1}^+) = \mathrm{Vol}_1(S_i^-) + \mathrm{Vol}_1(S_{i+1}^+).$$

Notice that $S_1^- = S_1$ and $S_k^+ = S_k \setminus \{s_k\}$, thus adding the above $k-1$ inequalities we obtain

$$\begin{aligned} (k-1)\mathrm{Vol}_1(S) &\ge \sum_{i=1}^{k-1} \big(\mathrm{Vol}_1(S_i^-) + \mathrm{Vol}_1(S_{i+1}^+)\big) \\ &= \mathrm{Vol}_1(S_1^-) + \mathrm{Vol}_1(S_k^+) + \sum_{i=2}^{k-1} \mathrm{Vol}_1(S_i) \\ &= \sum_{i=1}^{k} \mathrm{Vol}_1(S_i). \end{aligned}$$

We have thus established Conjecture 1.2 in dimension 1. □

Now we prove that Conjecture 1.2 passes to Cartesian product.

Theorem 3.2. *Let $k, m \ge 2$ and $n_1, \dots, n_m \ge 1$ be integers. Let $n = n_1 + \dots + n_m$. For $1 \le i \le k$ and $1 \le l \le m$, let A_i^l be some compact sets in $\mathbf{R}^{n_l}$. Assume that for any $1 \le l \le m$ the k compact sets $A_1^l, \dots, A_k^l \subset \mathbf{R}^{n_l}$ satisfy Conjecture 1.2. For $1 \le i \le k$, let $A_i = A_i^1 \times \dots \times A_i^m \subset \mathbf{R}^n = \mathbf{R}^{n_1} \times \dots \times \mathbf{R}^{n_m}$. Then Conjecture 1.2 holds for $A_1, \dots, A_k$.*

Proof. Let $S = \sum_{i=1}^k A_i$ and let $S_i = \sum_{j \ne i} A_j$ then let us prove that

$$(k-1)\mathrm{Vol}_n(S)^{\frac{1}{n}} \ge \sum_{i=1}^{k} \mathrm{Vol}_n(S_i)^{\frac{1}{n}}.$$

For all $1 \le i \le k$, one has

$$S_i = \sum_{j \ne i} A_j = \sum_{j \ne i} \prod_{l=1}^{m} A_j^l = \prod_{l=1}^{m} \left(\sum_{j \ne i} A_j^l \right).$$

For $1 \le i \le k$, denote $\sigma_i = (\mathrm{Vol}_{n_l}(\sum_{j \ne i} A_j^l)^{\frac{1}{n_l}})_{1 \le l \le m} \in \mathbf{R}^m$, and for $x = (x_l)_{1 \le l \le m} \in \mathbf{R}^m$, denote $\|x\|_0 = \prod_{l=1}^m |x_l|^{\frac{n_l}{n}}$. Then, using Minkowski's inequality for $\|\cdot\|_0$ (see, for example, Theorem 10 in [41]), we deduce that

$$\sum_{i=1}^{k} \mathrm{Vol}_n(S_i)^{\frac{1}{n}} = \sum_{i=1}^{k} \prod_{l=1}^{m} \mathrm{Vol}_{n_l} \left(\sum_{j \ne i} A_j^l \right)^{\frac{1}{n}} = \sum_{i=1}^{k} \|\sigma_i\|_0 \le \left\| \sum_{i=1}^{k} \sigma_i \right\|_0 = \prod_{l=1}^{m} \left(\sum_{i=1}^{k} \sigma_i^l \right)^{\frac{n_l}{n}}.$$

Using that for any $1 \le l \le m$ the k compact sets $A_1^l, \dots, A_k^l \subset \mathbf{R}^{n_l}$ satisfy Conjecture 1.2, we obtain

$$\sum_{i=1}^{k} \sigma_i^l = \sum_{i=1}^{k} \mathrm{Vol}_{n_l} \left(\sum_{j \ne i} A_j^l \right)^{\frac{1}{n_l}} \le (k-1)\mathrm{Vol}_{n_l} \left(\sum_{i=1}^{k} A_i^l \right)^{\frac{1}{n_l}}.$$

Thus

$$\sum_{i=1}^{k} \mathrm{Vol}_n(S_i)^{\frac{1}{n}} \leq \prod_{l=1}^{m} \left((k-1)\mathrm{Vol}_{n_l}\left(\sum_{i=1}^{k} A_i^l \right)^{\frac{1}{n_l}} \right)^{\frac{n_l}{n}} = (k-1)\mathrm{Vol}_n(S).$$

□

From Theorems 3.1 and 3.2, and the fact that Conjecture 1.2 holds for convex sets, we deduce that Conjecture 1.2 holds for Cartesian products of one-dimensional compact sets and convex sets.

3.2 A counterexample in dimension ≥ 12

In contrast to the positive results for compact product sets, both the conjectures of Bobkov, Madiman and Wang [19] fail in general for even moderately high dimension.

Theorem 3.3. *For every $k \geq 2$, there exists $n_k \in \mathbb{N}$ such that for every $n \geq n_k$ there is a compact set $A \subset \mathbf{R}^n$ such that $\mathrm{Vol}_n(A(k+1)) < \mathrm{Vol}_n(A(k))$. Moreover, one may take*

$$n_k = \min\left\{ n \in k\mathbb{Z} : n > \frac{\log(k)}{\log\left(1+\frac{1}{k}\right) - \frac{\log(2)}{k}} \right\}.$$

In particular, one has $n_2 = 12$, whence Conjectures 1.1 and Conjecture 1.2 are false in $\mathbf{R}^n$ for $n \geq 12$.

Proof. Let $k \geq 2$ be fixed and let n_k be defined as in the statement of Theorem 3.3 so that $n_k > \frac{\log(k)}{\log\left(1+\frac{1}{k}\right) - \frac{\log(2)}{k}}$ and $n_k = kd$, for a certain $d \in \mathbb{N}$. Let $F_1, \ldots, F_k$ be k linear subspaces of $\mathbf{R}^{n_k}$ of dimension d such that $\mathbf{R}^{n_k} = F_1 \oplus \cdots \oplus F_k$. Set $A = I_1 \cup \cdots \cup I_k$, where for every $i \in [k]$, I_i is a convex body in F_i. Notice that for every $l \geq 1$,

$$\underbrace{A + \cdots + A}_{l \text{ times}} = \bigcup_{m_i \in \{0,\cdots,l\}, \sum_{i=1}^{k} m_i = l} (m_1 I_1 + \cdots + m_k I_k),$$

where we used the convexity of each I_i to write the Minkowski sum of m_i copies of I_i as $m_i I_i$. Thus

$$k^{n_k}\mathrm{Vol}_{n_k}(A(k)) = \mathrm{Vol}_{n_k}(I_1 + \cdots + I_k) = \mathrm{Vol}_{n_k}(I_1 \times \cdots \times I_k),$$

and

$$\begin{aligned}
(k+1)^{n_k}\mathrm{Vol}_{n_k}(A(k+1)) &= \mathrm{Vol}_{n_k}((2I_1 + I_2 + \cdots + I_k) \cup \cdots \cup (I_1 + \cdots + I_{k-1} + 2I_k)) \\
&= \mathrm{Vol}_{n_k}((2I_1 \times I_2 \times \cdots \times I_k) \cup \cdots \cup (I_1 \times \cdots \times I_{k-1} \times 2I_k)) \\
&\leq \mathrm{Vol}_{n_k}(2I_1 \times I_2 \times \cdots \times I_k) + \cdots + \mathrm{Vol}_{n_k}(I_1 \times \cdots \times I_{k-1} \times 2I_k) \\
&= k2^d\mathrm{Vol}_{n_k}(I_1 \times \cdots \times I_k) \\
&= k^{n_k+1}2^d\mathrm{Vol}_{n_k}(A(k)).
\end{aligned}$$

The hypothesis on n_k enables us to conclude that $\mathrm{Vol}_{n_k}(A(k+1)) < \mathrm{Vol}_{n_k}(A(k))$. Now for $n \geq n_k$, we define $\tilde{A} = A \times [0,1]^{n-n_k}$. For every l, one has $\tilde{A}(l) = A(l) \times [0,1]^{n-n_k}$, thus $\mathrm{Vol}_n(\tilde{A}(l)) = \mathrm{Vol}_{n_k}(A(l))$. Therefore $\mathrm{Vol}_n(\tilde{A}(k+1)) < \mathrm{Vol}_n(\tilde{A}(k))$, which establishes that $\tilde{A}$ gives a counterexample in $\mathbf{R}^n$.

The sequence $\left\{ \frac{\log(k)}{\log\left(1+\frac{1}{k}\right) - \frac{\log(2)}{k}} \right\}_{k \geq 2}$ is increasing and $\frac{\log(2)}{\log\left(1+\frac{1}{2}\right) - \frac{\log(2)}{2}} \approx 11.77$. Hence, Conjecture 1.1 is false for $n \geq 12$. □

Remark 3.4. 1. *It is instructive to visualize the counterexample for $k = 2$, which is done in Figure 1 by representing each of the two orthogonal copies of $\mathbf{R}^6$ by a line.*

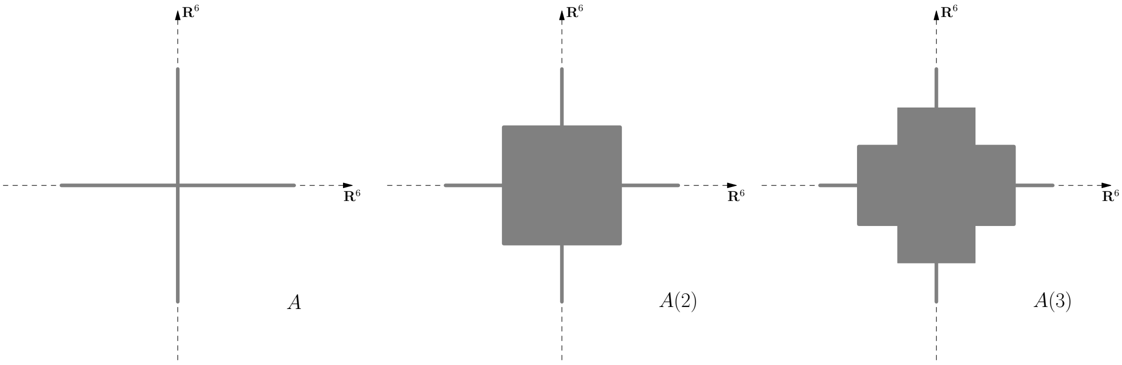

Figure 7: A counterexample in $\mathbf{R}^{12}$.

2. *It was shown by Bobkov, Madiman and Wang [19] that Conjecture 1.2 is true for convex sets. The constructed counterexample is a union of convex sets and is symmetric and star-shaped.*

3. *Notice that in the above example one has $\mathrm{Vol}_n(A(k-1)) = 0$. By adding to A a ball with sufficiently small radius, one obtains a counterexample satisfying $\mathrm{Vol}_n(A(k)) > \mathrm{Vol}_n(A(k-1)) > 0$ and $\mathrm{Vol}_n(A(k)) > \mathrm{Vol}_n(A(k+1))$.*

4. *The counterexample also implies that Conjecture 1.1 in [19], which suggests a fractional version of Young's inequality for convolution with sharp constant, is false. It is still possible that it may be true for a restricted class of functions (like the log-concave functions).*

5. *Conjectures 1.2 and 1.1 are still open in dimension $n \in \{2, \dots, 11\}$.*

3.3 Convergence rates for Δ

The asymptotic behavior of $\Delta(A(k))$ has been extensively studied by Emerson and Greenleaf [31]. In analyzing $\Delta(A(k))$, the following lemma about convergence of $A(k)$ to 0 in Hausdorff distance is useful.

Lemma 3.5. *If A is a compact set in $\mathbf{R}^n$,*

$$\mathrm{conv}(A) \subset A(k) + \frac{n\,\mathrm{diam}(A)}{k} B_2^n. \tag{20}$$

Proof. Using invariance of (20) under the shifts of A, we may assume that $0 \in \mathrm{conv}(A)$,

$$\mathrm{conv}(A) = \mathrm{conv}(A(k)) \subset (1 + c(A(k)))\mathrm{conv}(A) = A(k) + c(A(k))\mathrm{conv}(A).$$

Using $c(A(k)) \leq \frac{c(A)}{k}$ (see Section 5), as well as $c(A) \leq n$, we deduce that

$$\mathrm{conv}(A) \subset A(k) + \frac{n}{k}\mathrm{conv}(A).$$

To conclude, we note that since $0 \in \text{conv}(A)$, one has $|x| \leq \text{diam}(A)$ for every $x \in \text{conv}(A)$. Hence, $\text{conv}(A) \subset \text{diam}(A) B_2^n$. Finally, we obtain

$$\text{conv}(A) \subset A(k) + \frac{n\,\text{diam}(A)}{k} B_2^n.$$

$\square$

Note that Lemma 3.5 is essentially the same as the Shapley-Folkman-Starr theorem discussed in the introduction, and which we will prove in Section 7.3. Lemma 3.5 was contained in [31], but with an extra factor of 2.

One clearly needs assumption beyond compactness to have asymptotic vanishing of $\Delta(A(k))$. Indeed, a simple counterexample would be a finite set A of points, for which $\Delta(A(k))$ always remains at $\text{Vol}_n(\text{conv}(A))$ and fails to converge to 0. Once such an assumption is made, however, one has the following result.

Theorem 3.6. *[31] Let A be a compact set in $\mathbf{R}^n$ with nonempty interior. Then*

$$\Delta(A(k)) \leq \frac{C}{k} \text{Vol}_n(\text{conv}(A)),$$

for some constant C possibly depending on n.

Proof. By translation-invariance, we may assume that $\delta B_2^n \subset A$ for some $\delta > 0$. Then $\delta B_2^n \subset A(k_0)$, and by taking $k_0 \geq \frac{n\,\text{diam}(A)}{\delta}$, we have

$$\frac{n\,\text{diam}(A)}{k_0} B_2^n \subset A(k_0).$$

Hence using (20) we get

$$\text{conv}(A) \subset A(k) + \frac{k_0}{k} A(k_0) = \frac{k + k_0}{k} A(k + k_0),$$

so that by taking the volume we have

$$\text{Vol}_n(\text{conv}(A)) \leq \left(1 + \frac{k_0}{k}\right)^n \text{Vol}_n(A(k + k_0)),$$

and

$$\Delta(A(k + k_0)) \leq \left[\left(1 + \frac{k_0}{k}\right)^n - 1\right] \text{Vol}_n(A(k + k_0)) = O\left(\frac{1}{k}\right) \text{Vol}_n(\text{conv}(A)).$$

$\square$

4 Volume inequalities for Minkowski sums

4.1 A refined superadditivity of the volume for compact sets

In this section, we observe that if the exponents of $1/n$ in Conjecture 1.2 are removed, then the modified inequality is true (though unfortunately one can no longer directly relate this to a law of large numbers for sets).

Theorem 4.1. *Let $n \geq 1$, $k \geq 2$ be integers and let $A_1, \dots, A_k$ be k compact sets in $\mathbf{R}^n$. Then*

$$\mathrm{Vol}_n\left(\sum_{i=1}^{k} A_i\right) \geq \frac{1}{k-1}\sum_{i=1}^{k} \mathrm{Vol}_n\left(\sum_{j\in[k]\setminus\{i\}} A_j\right). \tag{21}$$

Proof. We use arguments similar to the proof of Theorem 3.1. Indeed, let us define the sets S and S_i in the same way as in the proof of Theorem 3.1. Let $\theta \in S^{n-1}$ be any fixed unit vector and let us define $a_i = \min\{\langle x, \theta\rangle; x \in A_i\}$, $b_i = \max\{\langle x, \theta\rangle; x \in A_i\}$, $S_i^- = \{x \in S_i; \langle x, \theta\rangle \leq s_i\}$ and $S_i^+ = \{x \in S_i; \langle x, \theta\rangle > s_i\}$. Then, the same inclusions hold true and thus we obtain

$$\begin{aligned}
(k-1)\mathrm{Vol}_n(S) &\geq \sum_{i=1}^{k-1}\left(\mathrm{Vol}_n(S_i^-) + \mathrm{Vol}_n(S_{i+1}^+)\right)\\
&= \mathrm{Vol}_n(S_1^-) + \mathrm{Vol}_n(S_k^+) + \sum_{i=2}^{k-1}\mathrm{Vol}_n(S_i)\\
&= \sum_{i=1}^{k}\mathrm{Vol}_n(S_i).
\end{aligned}$$

□

Applying Theorem 4.1 to $A_1 = \cdots = A_k = A$ yields the following positive result.

Corollary 4.2. *Let A be a compact set in $\mathbf{R}^n$ and $A(k)$ be defined as in (1). Then*

$$\mathrm{Vol}_n(A(k)) \geq \left(\frac{k-1}{k}\right)^{n-1}\mathrm{Vol}_n(A(k-1)). \tag{22}$$

In the following proposition, we improve Corollary 4.2 under additional assumptions on the set $A \subset \mathbf{R}^n$, for $n \geq 2$.

Proposition 4.3. *Let A be a compact subset of $\mathbf{R}^n$ and $A(k)$ be defined as in (1). If there exists a hyperplane H such that $\mathrm{Vol}_{n-1}(P_H(A)) = \mathrm{Vol}_{n-1}(P_H(\mathrm{conv}(A)))$, where $P_H(A)$ denotes the orthogonal projection of A onto H, then*

$$\mathrm{Vol}_n(A(k)) \geq \frac{k-1}{k}\mathrm{Vol}_n(A(k-1)).$$

Proof. By assumption, $\mathrm{Vol}_{n-1}(P_H(A)) = \mathrm{Vol}_{n-1}(P_H(\mathrm{conv}(A)))$. Thus, for every $k \geq 1$, $\mathrm{Vol}_{n-1}(P_H(A(k))) = \mathrm{Vol}_{n-1}(P_H(\mathrm{conv}(A)))$. Indeed, one has $A \subset A(k) \subset \mathrm{conv}(A)$. Thus, $P_H(A) \subset P_H(A(k)) \subset P_H(\mathrm{conv}(A))$. Hence,

$$\mathrm{Vol}_{n-1}(P_H(A)) \leq \mathrm{Vol}_{n-1}(P_H(A(k))) \leq \mathrm{Vol}_{n-1}(P_H(\mathrm{conv}(A))) = \mathrm{Vol}_{n-1}(P_H(A)).$$

It follows by the Bonnesen inequality (concave Brunn-Minkowski inequality, see [21]) that for every $k \geq 2$,

$$\begin{aligned}
\mathrm{Vol}_n(A(k)) &= \mathrm{Vol}_n\left(\frac{k-1}{k}A(k-1) + \frac{1}{k}A\right)\\
&\geq \frac{k-1}{k}\mathrm{Vol}_n(A(k-1)) + \frac{1}{k}\mathrm{Vol}_n(A) \geq \frac{k-1}{k}\mathrm{Vol}_n(A(k-1)).
\end{aligned}$$

□

Remark 4.4. *1. By considering the set $A = \{0, 1\}$ and $\delta_{\frac{1}{2}}$ the Dirac measure at $\frac{1}{2}$, one has*

$$\delta_{\frac{1}{2}}(A(2)) = 1 > 0 = \delta_{\frac{1}{2}}(A(3)).$$

Hence Conjecture 1.1 does not hold in general for log-concave measures in dimension 1.

2. *If A is countable, then for every $k \geq 1$, $\mathrm{Vol}_n(A(k)) = 0$, thus the sequence $\{\mathrm{Vol}_n(A(k))\}_{k\geq 1}$ is constant and equal to 0.*

3. *If there exists $k_0 \geq 1$ such that $A(k_0) = \mathrm{conv}(A)$, then for every $k \geq k_0$, $A(k) = \mathrm{conv}(A)$. Indeed,*

$$\begin{aligned}(k_0+1)A(k_0+1) &= k_0 A(k_0) + A = k_0 \mathrm{conv}(A) + A \\ &\supset \mathrm{conv}(A) + k_0 A(k_0) = (k_0+1)\mathrm{conv}(A).\end{aligned}$$

It follows that $A(k_0+1) = \mathrm{conv}(A)$. We conclude by induction. Thus, in this case, the sequence $\{\mathrm{Vol}_n(A(k))\}_{k\geq 1}$ is stationary to $\mathrm{Vol}_n(\mathrm{conv}(A))$, for $k \geq k_0$.

4. *It is natural to ask if the refined superadditivity of volume can be strengthened to fractional superadditivity as defined in Definition 4.11 below. However, this appears to be a difficult question even in dimension 1. In the case of compact subsets of $\mathbf{R}$, it was shown in unpublished work [12] of F. Barthe, M. Madiman and L. Wang that fractional superadditivity is true when dealing with up to $k = 4$ sets but only partial results were obtained for higher k.*

4.2 Supermodularity of volume for convex sets

If we restrict to convex sets, an even stronger inequality is true from which we can deduce Theorem 4.1 for convex sets.

Theorem 4.5. *Let $n \in \mathbb{N}$. For compact convex subsets B_1, B_2, B_3 of $\mathbf{R}^n$, one has*

$$\mathrm{Vol}_n(B_1 + B_2 + B_3) + \mathrm{Vol}_n(B_1) \geq \mathrm{Vol}_n(B_1 + B_2) + \mathrm{Vol}_n(B_1 + B_3). \tag{23}$$

We will provide two separate proofs of Theorem 4.5. We first observe that Theorem 4.5 is actually equivalent to a formal strengthening of it, namely Theorem 4.7 below. Let us first recall the notion of a supermodular set function.

Definition 4.6. *A set function $f : 2^{[k]} \to \mathbf{R}$ is supermodular if*

$$f(s \cup t) + f(s \cap t) \geq f(s) + f(t) \tag{24}$$

for all subsets s, t of $[k]$.

Theorem 4.7. *Let $B_1, \ldots, B_k$ be compact convex subsets of $\mathbf{R}^n$, and define*

$$v(s) = \mathrm{Vol}_n\left(\sum_{i\in s} B_i\right) \tag{25}$$

for each $s \subset [k]$. Then $v : 2^{[k]} \to [0, \infty)$ is a supermodular set function.

Theorem 4.7 implies Theorem 4.5, namely

$$\mathrm{Vol}_n(B_1 + B_2 + B_3) + \mathrm{Vol}_n(B_1) \geq \mathrm{Vol}_n(B_1 + B_2) + \mathrm{Vol}_n(B_1 + B_3) \tag{26}$$

for compact convex subsets B_1, B_2, B_3 of $\mathbf{R}^n$, since the latter is a special case of Theorem 4.7 when $k = 3$. To see the reverse, apply the inequality (26) to

$$B_1 = \sum_{i \in s \cap t} A_i, \quad B_2 = \sum_{i \in s \setminus t} A_i, \quad B_3 = \sum_{i \in t \setminus s} A_i.$$

The first proof of Theorem 4.5 combines a property of determinants that seems to have been first explicitly observed by Ghassemi and Madiman [48] with a use of optimal transport inspired by Alesker, Dar and Milman [1]. Let us prepare the ground by stating these results.

Lemma 4.8. *[48] Let K_1, K_2 and K_3 be $n \times n$ positive-semidefinite matrices. Then*

$$\det(K_1 + K_2 + K_3) + \det(K_1) \geq \det(K_1 + K_2) + \det(K_1 + K_3).$$

We state the deep result of [1] directly for k sets instead of for two sets as in [1] (the proof is essentially the same, with obvious modifications).

Theorem 4.9 (Alesker-Dar-Milman [1])**.** *Let $A_1, \ldots, A_k \subset \mathbf{R}^n$ be open, convex sets with $|A_i| = 1$ for each $i \in [k]$. Then there exist C^1-diffeomorphisms $\psi_i : A_1 \to A_i$ preserving Lebesgue measure, such that*

$$\sum_{i \in [k]} \lambda_i A_i = \Big\{ \sum_{i \in [k]} \lambda_i \psi_i(x) : x \in A_1 \Big\},$$

for any $\lambda_1, \ldots, \lambda_k > 0$.

First proof of Theorem 4.5. By adding a small multiple of the Euclidean ball B_2^n and then using the continuity of $\varepsilon \mapsto \mathrm{Vol}_n(B_i + \varepsilon B_2^n)$ as $\varepsilon \to 0$, we may assume that each of the B_i satisfy $\mathrm{Vol}_n(B_i) > 0$. Then choose λ_i such that $B_i = \lambda_i A_i$ with $|A_i| = 1$, so that

$$\begin{aligned} \mathrm{Vol}_n(B_1 + B_2 + B_3) &= \mathrm{Vol}_n\Big(\sum_i \lambda_i A_i\Big) = \int 1_{\sum_i \lambda_i A_i}(x)dx \\ &= \int 1_{\left\{ \sum_{i \in [M]} \lambda_i \psi_i(y) : y \in A_1 \right\}}(x)dx, \end{aligned}$$

using Theorem 4.9. Applying a change of coordinates using the diffeomorphism $x = \sum_{i \in [M]} \lambda_i \psi_i(y)$,

$$\begin{aligned} V &:= \mathrm{Vol}_n(B_1 + B_2 + B_3) = \int 1_{A_1}(y) \det\Big(\sum_i \lambda_i D\psi_i\Big)(y)dy \\ &\geq \int_{A_1} \det[(\lambda_1 D\psi_1 + \lambda_2 D\psi_2)(y)] + \det[(\lambda_1 D\psi_1 + \lambda_3 D\psi_3)(y)] - \det[\lambda_1 D\psi_1(y)]dy \\ &= \int 1_{A_1}(y) d[(\lambda_1\psi_1 + \lambda_2\psi_2)(y)] + \int 1_{A_1}(y) d[(\lambda_1\psi_1 + \lambda_3\psi_3)(y)] - \int 1_{A_1}(y) d[\lambda_1\psi_1(y)]dy \\ &= \int 1_{\{\lambda_1\psi_1(y) + \lambda_2\psi_2(y) : y \in A_1\}}(z)dz + \int 1_{\{\lambda_1\psi_1(y) + \lambda_3\psi_3(y) : y \in A_1\}}(z')dz' \\ &\qquad - \int 1_{\{\lambda_1\psi_1(y) : y \in A_1\}}(z'')dz'' \end{aligned}$$

where the inequality follows from Lemma 4.8, and the last equality is obtained by making multiple appropriate coordinate changes. Using Theorem 4.9 again,

$$\begin{aligned}\mathrm{Vol}_n(B_1+B_2+B_3) &\geq \int 1_{\lambda_1 A_1+\lambda_2 A_2}(z)dz + \int 1_{\lambda_1 A_1+\lambda_3 A_3}(z)dz - \int 1_{\lambda_1 A_1}(z)dz \\ &= \mathrm{Vol}_n(B_1+B_2)+\mathrm{Vol}_n(B_1+B_3)-\mathrm{Vol}_n(B_1).\end{aligned}$$

□

We now present a second proof of Theorem 4.5 using mixed volumes. Let us recollect some basic facts about these, which can be found, e.g., in chapter 5 of Schneider's book [62]. From Theorem 5.1.6. of [62], for any compact convex sets K and L in $\mathbf{R}^n$, the function $t \mapsto |K+tL|$ is a polynomial in t, for $t \geq 0$, and one has

$$\mathrm{Vol}_n(K+tL) = \sum_{k=0}^{n} \binom{n}{k} V(K[n-k], L[k])t^k, \tag{27}$$

for some nonnegative numbers $V(K[n-k], L[k])$, which are called the mixed volumes of K and L. One readily sees that the mixed volumes satisfy $V(K[n], L[0]) = |K|$ and

$$V((K+a)[n-k], L[k]) = V(K[n-k], L[k])$$

for any $a \in \mathbf{R}^n$, a property that we shall call translation-invariance. Moreover it is classical (see Schneider [62], equation (5.1.23)) that they satisfy the following monotonicity property: If $K \subset K'$, then $V(K[n-k], L[k]) \leq V(K'[n-k], L[k])$.

Second proof of Theorem 4.5. We apply the mixed volume formula (27) for $t = 1$, $K = B_1+B_2$ and $L = B_3$ to get

$$\mathrm{Vol}_n(B_1+B_2+B_3) - \mathrm{Vol}_n(B_1+B_2) = \sum_{k=0}^{n-1} \binom{n}{k} V((B_1+B_2)[n-k], B_3[k]).$$

By translation-invariance of mixed volumes, we may assume that $0 \in B_2$, and thus that $B_1 \subset B_1 + B_2$. Using the monotonicity property of mixed volumes, we have the inequality

$$\begin{aligned}\mathrm{Vol}_n(B_1+B_2+B_3) - \mathrm{Vol}_n(B_1+B_2) &\geq \sum_{k=0}^{n-1} \binom{n}{k} V(B_1[n-k], B_3[k]) \\ &= \mathrm{Vol}_n(B_1+B_3) - \mathrm{Vol}_n(B_3),\end{aligned}$$

using again the mixed volume formula at the end with $t = 1$, $K = B_1$ and $L = B_3$. □

For the purposes of discussion below, it is useful to collect some well known facts from the theory of supermodular set functions. Observe that if v is supermodular and $v(\emptyset) = 0$, then considering disjoint s and t in (24) implies that v is superadditive. In fact, a more general structural result is true. To describe it, we need some terminology.

Definition 4.10. *Given a collection $\mathcal{C}$ of subsets of $[k]$, a function $\alpha : \mathcal{C} \to \mathbf{R}^+$, is called a* fractional partition, *if for each $i \in [k]$, we have $\sum_{s\in\mathcal{C}: i\in s} \alpha_s = 1$.*

The reason for the terminology is that this notion extends the familiar notions of a partition of sets (whose indicator function can be defined precisely as in Definition 4.10 but with range restriction to $\{0, 1\}$) by allowing fractional values. An important example of a fractional partition of $[k]$ is the collection $\mathcal{C}_m = \binom{[k]}{m}$ of all subsets of size m, together with the coefficients $\alpha_s = \binom{k-1}{m-1}^{-1}$.

Definition 4.11. *A function $f : 2^{[k]} \to \mathbf{R}$ is fractionally superadditive if for any fractional partition $(\mathcal{C}, \beta)$,*

$$f([k]) \geq \sum_{s \in \mathcal{C}} \beta_s f(s).$$

The following theorem has a long history and is implicit in results from cooperative game theory in the 1960's but to our knowledge, it was first explicitly stated by Moulin Ollagnier and Pinchon [57].

Theorem 4.12. *[57] If $f : 2^{[k]} \to \mathbf{R}$ is supermodular and $f(\emptyset) = 0$, then f is fractionally superadditive.*

A survey of the history of Theorem 4.12, along with various strengthenings of it and their proofs, and discussion of several applications, can be found in [51]. If $\{A_i, i \in [k]\}$ are compact convex sets and $u(s) = \mathrm{Vol}_n(\sum_{i \in s} A_i)$ as defined in (25), then $u(\emptyset) = 0$ and Theorem 4.7 says that u is supermodular, whence Theorem 4.12 immediately implies that u is fractionally superadditive.

Corollary 4.13. *Let $B_1, \ldots, B_k$ be compact convex subsets of $\mathbf{R}^n$ and let β be any fractional partition using a collection $\mathcal{C}$ of subsets of $[k]$. Then*

$$\mathrm{Vol}_n\Big(\sum_{i \in [k]} B_i\Big) \geq \sum_{s \in \mathcal{C}} \beta_s \mathrm{Vol}_n\Big(\sum_{i \in s} B_i\Big).$$

Corollary 4.13 implies that for each $m < k$,

$$\mathrm{Vol}_n\Big(\sum_{i \in [k]} B_i\Big) \geq \binom{k-1}{m-1}^{-1} \sum_{|s|=m} \mathrm{Vol}_n\Big(\sum_{i \in s} B_i\Big). \tag{28}$$

Let us discuss whether these inequalities contain anything novel. On the one hand, if we consider the case $m = 1$ of inequality (28), the resulting inequality is not new and in fact implied by the Brunn-Minkowski inequality:

$$\mathrm{Vol}_n\Big(\sum_{i \in [k]} B_i\Big) \geq \Big[\sum_{i \in [k]} \mathrm{Vol}_n(B_i)^{\frac{1}{n}}\Big]^n \geq \sum_{i \in [k]} \mathrm{Vol}_n(B_i).$$

On the other hand, applying the inequality (28) to $m = k - 1$ yields precisely Theorem 4.1 for convex sets B_i, i.e.,

$$\mathrm{Vol}_n\Big(\sum_{i \in [k]} B_i\Big) \geq \frac{1}{k-1} \sum_{i \in [k]} \mathrm{Vol}_n\Big(\sum_{j \neq i} B_j\Big). \tag{29}$$

Let us compare this with what is obtainable from the refined Brunn-Minkowski inequality for convex sets proved in [19], which says that

$$\mathrm{Vol}_n\Big(\sum_{i \in [k]} B_i\Big) \geq \Big(\frac{1}{k-1}\Big)^n \Big[\sum_{i \in [k]} \mathrm{Vol}_n\big(\sum_{j \neq i} B_j\big)^{\frac{1}{n}}\Big]^n. \tag{30}$$

Denote the right hand sides of (29) and (30) by $R_{(29)}$ and $R_{(30)}$. Also set

$$c_i = \mathrm{Vol}_n\Big(\sum_{j\neq i} B_j\Big)^{\frac{1}{n}},$$

and write $c = (c_1, \dots, c_k) \in [0,\infty)^k$, so that $R_{(29)}^{\frac{1}{n}} = (k-1)^{-\frac{1}{n}}\|c\|_n$ and $R_{(30)}^{\frac{1}{n}} = (k-1)^{-1}\|c\|_1$. Here, for $m \geq 1$, $\|c\|_m = \Big(\sum_{i=1}^k c_i^m\Big)^{\frac{1}{m}}$. In other words,

$$\left[\frac{R_{(29)}}{R_{(30)}}\right]^{\frac{1}{n}} = (k-1)^{1-\frac{1}{n}}\frac{\|c\|_n}{\|c\|_1}.$$

Let us consider $n = 2$ for illustration. Then we have

$$\left[\frac{R_{(29)}}{R_{(30)}}\right]^{\frac{1}{2}} = \sqrt{k-1}\frac{\|c\|_2}{\|c\|_1},$$

which ranges between $\sqrt{1-\frac{1}{k}}$ and $\sqrt{k-1}$, since $\|c\|_2/\|c\|_1 \in [k^{-\frac{1}{2}}, 1]$. In particular, neither bound is uniformly better; so the inequality (28) and Corollary 4.13 do indeed have some potentially useful content.

Motivated by the results of this section, it is natural to ask if the volume of Minkowski sums is supermodular even without the convexity assumption on the sets involved, as this would strengthen Theorem 4.1. In fact, this is not the case.

Proposition 4.14. *There exist compact sets $A, B, C \subset \mathbf{R}$ such that*

$$\mathrm{Vol}_1(A+B+C) + \mathrm{Vol}_1(A) < \mathrm{Vol}_1(A+B) + \mathrm{Vol}_1(A+C).$$

Proof. Consider $A = \{0,1\}$ and $B = C = [0,1]$. Then,

$$\mathrm{Vol}_1(A+B+C) + \mathrm{Vol}_1(A) = 3 < 4 = \mathrm{Vol}_1(A+B) + \mathrm{Vol}_1(A+C).$$

$\square$

On the other hand, the desired inequality is true in dimension 1 if the set A is convex. More generally, in dimension 1, one has the following result.

Proposition 4.15. *If $A, B, C \subset \mathbf{R}$ are compact, then*

$$\mathrm{Vol}_1(A+B+C) + \mathrm{Vol}_1(\mathrm{conv}(A)) \geq \mathrm{Vol}_1(A+B) + \mathrm{Vol}_1(A+C).$$

Proof. Assume, as one typically does in the proof of the one-dimensional Brunn-Minkowski inequality, that $\max B = 0 = \min C$. (We can do this without loss of generality since translation does not affect volumes.) This implies that $B \cup C \subset B + C$, whence

$$(A+B) \cup (A+C) = A + (B \cup C) \subset A+B+C.$$

Hence

$$\begin{aligned}\mathrm{Vol}_1(A+B+C) &\geq \mathrm{Vol}_1((A+B)\cup(A+C)) \\ &= \mathrm{Vol}_1(A+B) + \mathrm{Vol}_1(A+C) - \mathrm{Vol}_1((A+B)\cap(A+C)).\end{aligned}$$

We will show that $(A+B)\cap(A+C)\subset \mathrm{conv}(A)$, which together with the preceding inequality yields the desired conclusion $\mathrm{Vol}_1(A+B+C)\geq \mathrm{Vol}_1(A+B)+\mathrm{Vol}_1(A+C)-\mathrm{Vol}_1(\mathrm{conv}(A))$.

To see that $(A+B)\cap(A+C)\subset \mathrm{conv}(A)$, consider $x\in(A+B)\cap(A+C)$. One may write $x=a_1+b=a_2+b$, with $a_1,a_2\in A$, $b\in B$ and $c\in C$. Since $\max B=0=\min C$ one has $b\leq 0\leq c$ and one deduces that $a_2\leq x\leq a_1$ and thus $x\in\mathrm{conv}(A)$. This completes the proof. □

Remark 4.16. 1. *One may wonder if Proposition 4.15 extends to higher dimension. More particularly, we do not know if the supermodularity inequality*

$$\mathrm{Vol}_n(A+B+C)+\mathrm{Vol}_n(A)\geq \mathrm{Vol}_n(A+B)+\mathrm{Vol}_n(A+C)$$

holds true in the case where A is convex and B and C are any compact sets.

2. *It is also natural to ask in view of the results of this section whether the fractional superadditivity (2) of $\mathrm{Vol}_n^{1/n}$ for convex sets proved in [19] follows from a more general supermodularity property, i.e., whether*

$$\mathrm{Vol}_n^{1/n}(A+B+C)+\mathrm{Vol}_n^{1/n}(A)\geq \mathrm{Vol}_n^{1/n}(A+B)+\mathrm{Vol}_n^{1/n}(A+C) \tag{31}$$

for convex sets $A,B,C\subset\mathbf{R}^n$. It follows from results of [48] that such a result does not hold (their counterexample to the determinant version of (31) corresponds in our context to choosing ellipsoids in $\mathbf{R}^2$). Another simple explicit counterexample is the following: Let $A=[0,2]\times[0,1/2]$, $B=[0,1/2]\times[0,2]$, and $C=\varepsilon B_2^2$, with $\varepsilon>0$. Then,

$$\mathrm{Vol}_2(A)^{1/2}=1,\quad \mathrm{Vol}_2(A+B+C)^{1/2}=\sqrt{25/4+10\varepsilon+\pi\varepsilon^2},$$

$$\mathrm{Vol}_2(A+B)^{1/2}=5/2,\quad \mathrm{Vol}_2(A+C)^{1/2}=\sqrt{1+5\varepsilon+\pi\varepsilon^2}.$$

Hence,

$$\begin{aligned}\mathrm{Vol}_2(A+B+C)^{1/2}+\mathrm{Vol}_2(A)^{1/2}&=1+5/2+2\varepsilon+o(\varepsilon)\\ \mathrm{Vol}_2(A+B)^{1/2}+\mathrm{Vol}_2(A+C)^{1/2}&=1+5/2+(5/2)\varepsilon+o(\varepsilon)\end{aligned}$$

For ε small enough, this yields a counterexample to (31).

3. *It is shown in [48] that the entropy analogue of Proposition 4.15 does not hold, i.e., there exist independent real-valued random variables X,Y,Z with log-concave distributions such that*

$$e^{2h(X+Y+Z)}+e^{2h(Z)}<e^{2h(X+Z)}+e^{2h(Y+Z)}.$$

5 The behavior of Schneider's non-convexity index

5.1 The refined monotonicity of Schneider's non-convexity index

In this section, our main result is that Schneider's non-convexity index c satisfies a strong kind of monotonicity in any dimension.

We state the main theorem of this section, and will subsequently deduce corollaries asserting monotonicity in the Shapley-Folkman-Starr theorem from it.

Theorem 5.1. *Let $n \geq 1$ and let A, B, C be subsets of $\mathbf{R}^n$. Then*

$$c(A+B+C) \leq \max\{c(A+B), c(B+C)\}.$$

Proof. Let us denote $\lambda = \max\{c(A+B), c(B+C)\}$. Then

$$\begin{aligned}
A+B+C+\lambda\mathrm{conv}(A+B+C) &= A+B+\lambda\mathrm{conv}(A+B)+C+\lambda\mathrm{conv}(C)\\
&= (1+\lambda)\mathrm{conv}(A+B)+C+\lambda\mathrm{conv}(C)\\
&\supset (1+\lambda)\mathrm{conv}(A)+B+\lambda\mathrm{conv}(B)+C+\lambda\mathrm{conv}(C)\\
&= (1+\lambda)\mathrm{conv}(A)+(1+\lambda)\mathrm{conv}(B+C)\\
&= (1+\lambda)\mathrm{conv}(A+B+C).
\end{aligned}$$

Since the opposite inclusion is clear, we deduce that $A+B+C+\lambda\mathrm{conv}(A+B+C)$ is convex, which means that $c(A+B+C) \leq \lambda = \max\{c(A+B), c(B+C)\}$. □

Notice that the same kind of proof also shows that if $A+B$ and $B+C$ are convex then $A+B+C$ is also convex. Moreover, Theorem 5.1 has an equivalent formulation for $k \geq 2$ subsets of $\mathbf{R}^n$, say $A_1, \ldots, A_k$: if $s, t \subset [k]$ with $s \cup t = [k]$, then

$$c\left(\sum_{i\in[k]} A_i\right) \leq \max\left\{c\left(\sum_{i\in s} A_i\right), c\left(\sum_{i\in t} A_i\right)\right\}. \tag{32}$$

To see this, apply Theorem 5.1 to

$$B = \sum_{i\in s\cap t} A_i, \quad A = \sum_{i\in s\setminus t} A_i, \quad C = \sum_{i\in t\setminus s} A_i.$$

From the inequality (32), the following corollary, expressed in a more symmetric fashion, immediately follows.

Corollary 5.2. *Let $n \geq 1$ and $k \geq 2$ be integers and let $A_1, \ldots, A_k$ be k sets in $\mathbf{R}^n$. Then*

$$c\left(\sum_{l\in[k]} A_l\right) \leq \max_{i\in[k]} c\left(\sum_{l\in[k]\setminus\{i\}} A_l\right).$$

The $k = 2$ case of Corollary 5.2 follows directly from the definition of c and was observed by Schneider in [61]. Applying Corollary 5.2 for $A_1 = \cdots = A_k = A$, where A is a fixed subset of $\mathbf{R}^n$, and using the scaling invariance of c, one deduces that the sequence $c(A(k))$ is non-increasing. In fact, for identical sets, we prove something even stronger in the following theorem.

Theorem 5.3. *Let A be a subset of $\mathbf{R}^n$ and $k \geq 2$ be an integer. Then*

$$c\left(A(k)\right) \leq \frac{k-1}{k} c\left(A(k-1)\right).$$

Proof. Denote $\lambda = c\left(A(k-1)\right)$. Since $\mathrm{conv}(A(k-1)) = \mathrm{conv}(A)$, from the definition of c, one knows that $A(k-1) + \lambda \mathrm{conv}(A) = \mathrm{conv}(A) + \lambda \mathrm{conv}(A) = (1+\lambda)\mathrm{conv}(A)$. Using that $A(k) = \frac{A}{k} + \frac{k-1}{k} A(k-1)$, one has

$$\begin{aligned} A(k) + \frac{k-1}{k}\lambda \mathrm{conv}(A) &= \frac{A}{k} + \frac{k-1}{k} A(k-1) + \frac{k-1}{k}\lambda \mathrm{conv}(A) \\ &= \frac{A}{k} + \frac{k-1}{k}\mathrm{conv}(A) + \frac{k-1}{k}\lambda \mathrm{conv}(A) \\ &\supset \frac{\mathrm{conv}(A)}{k} + \frac{k-1}{k} A(k-1) + \frac{k-1}{k}\lambda \mathrm{conv}(A) \\ &= \frac{\mathrm{conv}(A)}{k} + \frac{k-1}{k}(1+\lambda)\mathrm{conv}(A) \\ &= \left(1 + \frac{k-1}{k}\lambda\right)\mathrm{conv}(A). \end{aligned}$$

Since the other inclusion is trivial, we deduce that $A(k) + \frac{k-1}{k}\lambda \mathrm{conv}(A)$ is convex which proves that

$$c(A(k)) \leq \frac{k-1}{k}\lambda = \frac{k-1}{k} c\left(A(k-1)\right).$$

□

Remark 5.4. *1. We do not know if c is fractionally subadditive; for example, we do not know if $2\,c(A+B+C) \leq c(A+B) + c(A+C) + c(B+C)$. We know it with a better constant if $A = B = C$, as a consequence of Theorem 5.3. We also know it if we take a large enough number of sets; this is a consequence of the Shapley-Folkman lemma (Lemma 2.2).*

2. The Schneider index c (as well as any other measure of non-convexity) cannot be submodular. This is because, if we consider $A = \{0,1\}$, $B = C = [0,1]$, then $c(A+B) = c(A+C) = c(A+B+C) = 0$ but $c(A) > 0$, hence

$$c(A+B+C) + c(A) > c(A+B) + c(A+C).$$

5.2 Convergence rates for Schneider's non-convexity index

We were unable to find any examination in the literature of rates, or indeed, even of sufficient conditions for convergence as measured by c.

Let us discuss convergence in the Shapley-Folkman-Starr theorem using the Schneider non-convexity index. In dimension 1, we can get an $O(1/k)$ bound on $c(A(k))$ by using the close relation (16) between c and d in this case. In general dimension, the same bound also holds: by applying Theorem 5.3 inductively, we get the following theorem.

Theorem 5.5. *Let A be a compact set in $\mathbf{R}^n$. Then*

$$c(A(k)) \leq \frac{c(A)}{k}.$$

In particular, $c(A(k)) \to 0$ as $k \to \infty$.

Let us observe that the $O(1/k)$ rate of convergence cannot be improved, either for d or for c. To see this simply consider the case where $A = \{0,1\} \subset \mathbf{R}$. Then $A(k)$ consists of the $k+1$ equispaced points j/k, where $j \in \{0,1,\ldots,k\}$, and $c(A(k)) = 2d(A(k)) = 1/k$ for every $k \in \mathbb{N}$.

6 The behavior of the effective standard deviation v

6.1 Subadditivity of v^2

Cassels [23] showed that v^2 is subadditive.

Theorem 6.1 ([23]). *Let A, B be compact sets in $\mathbf{R}^n$. Then,*

$$v^2(A+B) \leq v^2(A) + v^2(B).$$

Proof. Recall that $v(A) = \sup_{x \in \mathrm{conv}(A)} v_A(x)$, where

$$v_A^2(x) = \inf\{\sum_{i \in I} \lambda_i |a_i - x|^2 : (\lambda_i, a_i)_{i \in I} \in \Theta_A(x)\},$$

and $\Theta_A(x) = \{(\lambda_i, a_i)_{i \in I} : I \text{ finite}, \ x = \sum \lambda_i a_i; \lambda_i > 0; \sum \lambda_i = 1, a_i \in A\}$. Thus

$$v(A+B) = \sup_{x \in \mathrm{conv}(A+B)} v_{A+B}(x) = \sup_{x_1 \in \mathrm{conv}(A)} \sup_{x_2 \in \mathrm{conv}(B)} v_{A+B}(x_1 + x_2).$$

And one has

$$v_{A+B}^2(x_1+x_2) = \inf\{\sum_{i \in I} \nu_i |c_i - x_1 - x_2|^2 : (\nu_i, c_i)_{i \in I} \in \Theta_{A+B}(x_1+x_2)\}.$$

For $(\lambda_i, a_i)_{i \in I} \in \Theta_A(x_1)$ and $(\mu_j, b_j)_{j \in J} \in \Theta_B(x_2)$ one has

$$(\lambda_i \mu_j, a_i + b_j)_{(i,j) \in I \times J} \in \Theta_{A+B}(x_1 + x_2),$$

and

$$\begin{aligned}
&\sum_{(i,j) \in I \times J} \lambda_i \mu_j |a_i + b_j - x_1 - x_2|^2 \\
&= \sum_{i \in I} \lambda_i |a_i - x_1|^2 + \sum_{j \in J} \mu_j |b_j - x_2|^2 + 2 \sum_{(i,j) \in I \times J} \lambda_i \mu_j \langle a_i - x_1, b_j - x_2 \rangle \\
&= \sum_{i \in I} \lambda_i |a_i - x_1|^2 + \sum_{j \in J} \mu_j |b_j - x_2|^2 + 2 \langle \sum_{i \in I} \lambda_i a_i - x_1, \sum_{j \in J} \mu_j b_j - x_2 \rangle \\
&= \sum_{i \in I} \lambda_i |a_i - x_1|^2 + \sum_{j \in J} \mu_j |b_j - x_2|^2.
\end{aligned}$$

Thus

$$\begin{aligned}
v_{A+B}^2(x_1+x_2) &\leq \inf_{(\lambda_i, a_i)_{i \in I} \in \Theta_A(x_1)} \inf_{(\mu_j, b_j)_{j \in J} \in \Theta_B(x_2)} \sum_{i \in I} \lambda_i |a_i - x_1|^2 + \sum_{j \in J} \mu_j |b_j - x_2|^2 \\
&= v_A^2(x_1) + v_B^2(x_2).
\end{aligned}$$

Taking the supremum in $x_1 \in \mathrm{conv}(A)$ and $x_2 \in \mathrm{conv}(B)$, we conclude. $\square$

Observe that we may interpret the proof probabilistically. Indeed, a key point in the proof is the identity (33), which is just the fact that the variance of a sum of independent random variables is the sum of the individual variances (written out explicitly for readability).

6.2 Strong fractional subadditivity for large k

In this section, we prove that the effective standard deviation v satisfies a strong fractional subadditivity when considering sufficient large numbers of sets.

Theorem 6.2. *Let $A_1, \dots, A_k$ be compact sets in $\mathbf{R}^n$, with $k \geq n+1$. Then,*

$$v\left(\sum_{i\in[k]} A_i\right) \leq \max_{I\subset[k]:|I|\leq n} \min_{i\in[k]\setminus I} v\left(\sum_{j\in[k]\setminus\{i\}} A_j\right).$$

Proof. Let $x \in \mathrm{conv}(\sum_{i\in[k]} A_i)$, where $k \geq n+1$. By using the Shapley-Folkman lemma (Lemma 2.2), there exists a set I of at most n indexes such that

$$x \in \sum_{i\in I} \mathrm{conv}(A_i) + \sum_{i\in[k]\setminus I} A_i.$$

Let $i_0 \in [k]\setminus I$. In particular, we have

$$x \in \mathrm{conv}\left(\sum_{i\in[k]\setminus\{i_0\}} A_i\right) + A_{i_0}.$$

Hence, by definition of the convex hull,

$$x = \sum_m p_m a_m + a_{i_0} = z + a_{i_0},$$

where $z = \sum_m p_m a_m$, $\sum_m p_m = 1$, $a_m \in \sum_{i\in[k]\setminus\{i_0\}} A_i$ and $a_{i_0} \in A_{i_0}$. Thus, by denoting $A_{\{i_0\}} = \sum_{i\in[k]\setminus\{i_0\}} A_i$, we have

$$\begin{aligned}
v^2_{A_{\{i_0\}}}(z) &= \inf\left\{\sum_m p_m|a_m - z|^2 : z = \sum_m p_m a_m; \sum_m p_m = 1; a_m \in A_{\{i_0\}}\right\} \\
&= \inf\left\{\sum_m p_m|a_m + a_{i_0} - (z+a_{i_0})|^2 : z = \sum_m p_m a_m; \sum_m p_m = 1; a_m \in A_{\{i_0\}}\right\} \\
&\geq \inf\left\{\sum_m p_m|a^*_m - (z+a_{i_0})|^2 : z + a_{i_0} = \sum_m p_m a^*_m; \sum_m p_m = 1; a^*_m \in \sum_{i\in[k]} A_i\right\} \\
&= v^2_{\sum_{i\in[k]} A_i}(x).
\end{aligned}$$

Taking supremum over all $z \in \mathrm{conv}(\sum_{i\in[k]\setminus\{i_0\}} A_i)$, we deduce that

$$v_{\sum_{i\in[k]} A_i}(x) \leq v\left(\sum_{i\in[k]\setminus\{i_0\}} A_i\right).$$

Since this is true for every $i_0 \in [k]\setminus I$, we deduce that

$$v_{\sum_{i\in[k]} A_i}(x) \leq \min_{i\in[k]\setminus I} v\left(\sum_{j\in[k]\setminus\{i\}} A_j\right).$$

Taking the supremum over all set $I \subset [k]$ of cardinality at most n yields

$$v_{\sum_{i\in[k]} A_i}(x) \leq \max_{I\subset[k]:|I|\leq n} \min_{i\in[k]\setminus I} v\left(\sum_{j\in[k]\setminus\{i\}} A_j\right).$$

We conclude by taking the supremum over all $x \in \mathrm{conv}(\sum_{i\in[k]} A_i)$.

$\square$

An immediate consequence of Theorem 6.2 is that if $k \geq n+1$, then

$$v\left(\sum_{i\in[k]} A_i\right) \leq \max_{i\in[k]} v\left(\sum_{j\in[k]\setminus\{i\}} A_j\right).$$

By iterating this fact as many times as possible (i.e., as long as the number of sets is at least $n+1$), we obtain the following corollary.

Corollary 6.3. *Let $A_1,\dots,A_k$ be compact sets in $\mathbf{R}^n$, with $k \geq n+1$. Then,*

$$v\left(\sum_{i\in[k]} A_i\right) \leq \max_{I\subset[k]:|I|=n} v\left(\sum_{j\in I} A_j\right).$$

In the case where $A_1 = \cdots = A_k = A$, we can repeat the above argument with $k \geq c(A)+1$ to prove that in this case,

$$v(A(k)) \leq \frac{k-1}{k} v(A(k-1)),$$

where $c(A)$ is the Schneider non-convexity index of A. Since $c(A) \leq n$, and $c(A) \leq n-1$ when A is connected, we deduce the following monotonicity property for the effective standard deviation.

Corollary 6.4.
1. *In dimension 1 and 2, the sequence $v(A(k))$ is non-increasing for every compact set A.*
2. *In dimension 3, the sequence $v(A(k))$ is non-increasing for every compact and connected set A.*

Remark 6.5. *It follows from the above study that if a compact set $A \subset \mathbf{R}^n$ satisfies $c(A) \leq 2$, then the sequence $v(A(k))$ is non-increasing. One can see that if a compact set $A \subset \mathbf{R}^n$ contains the boundary of its convex hull, then $c(A) \leq 1$; for such set $A \subset \mathbf{R}^n$, the sequence $v(A(k))$ is non-increasing.*

6.3 Convergence rates for v

It is classical that one has convergence in v at good rates.

Theorem 6.6 ([23]). *Let $A_1,\dots,A_k$ be compact sets in $\mathbf{R}^n$. Then*

$$v(A_1+\cdots+A_k) \leq \sqrt{\min\{k,n\}} \max_{i\in[k]} v(A_i).$$

Proof. Firstly, by using subadditivity of v^2 (Theorem 6.1), one has

$$v^2(A_1+\cdots+A_k) \leq k \max_{i\in[k]} v^2(A_i).$$

Hence, $v(A_1+\cdots+A_k) \leq \sqrt{k}\max_{i\in[k]} v(A_i)$.

If $k \geq n+1$, we can improve this bound using Corollary 6.3, which gives us

$$\begin{aligned}
v^2\left(\sum_{i\in[k]} A_i\right) &\leq \max_{I\subset[k]:|I|=n} v^2\left(\sum_{j\in I} A_j\right)\\
&\leq \max_{I\subset[k]:|I|=n} \sum_{j\in I} v^2(A_j)\\
&\leq n \max_{i\in I} v^2(A_i) \leq n \max_{i\in[k]} v^2(A_i),
\end{aligned}$$

again using subadditivity of v^2 for the second inequality. ∎

By considering $A_1 = \cdots = A_k = A$, one obtains the following convergence rate.

Corollary 6.7. *Let A be a compact set in $\mathbf{R}^n$. Then,*

$$v(A(k)) \leq \min\left\{\frac{1}{\sqrt{k}}, \frac{\sqrt{n}}{k}\right\} v(A).$$

7 The behavior of the Hausdorff distance

7.1 Some basic properties of the Hausdorff distance

The Hausdorff distance is subadditive.

Theorem 7.1. *Let A, B be compact sets in $\mathbf{R}^n$, and K be an arbitrary convex body containing 0 in its interior. Then*

$$d^{(K)}(A+B) \leq d^{(K)}(A) + d^{(K)}(B).$$

Proof. The convexity of K implies that

$$A + B + (d^{(K)}(A) + d^{(K)}(B))K = A + d^{(K)}(A)K + B + d^{(K)}(B)K,$$

but since $A + d^{(K)}(A)K \supset \mathrm{conv}(A)$ and $B + d^{(K)}(B)K \supset \mathrm{conv}(B)$ by definition, we have

$$A + B + (d^{(K)}(A) + d^{(K)}(B))K \supset \mathrm{conv}(A) + \mathrm{conv}(B) = \mathrm{conv}(A+B).$$

∎

We can provide a slight further strengthening of Theorem 7.1 when dealing with Minkowski sums of more than 2 sets, by following an argument similar to that used for Schneider's non-convexity index.

Theorem 7.2. *Let A, B, C be compact sets in $\mathbf{R}^n$, and K be an arbitrary convex body containing 0 in its interior. Then*

$$d^{(K)}(A+B+C) \leq d^{(K)}(A+B) + d^{(K)}(B+C).$$

Proof. Notice that

$$\begin{aligned}
&A + B + C + \big(d^{(K)}(A+B) + d^{(K)}(B+C)\big)K \\
&= A + B + d^{(K)}(A+B)K + C + d^{(K)}(B+C)K \\
&\supset \mathrm{conv}(A+B) + C + d^{(K)}(B+C)K \\
&\supset \mathrm{conv}(A) + B + C + d^{(K)}(B+C)K \\
&\supset \mathrm{conv}(A) + \mathrm{conv}(B+C) \\
&= \mathrm{conv}(A+B+C).
\end{aligned}$$

∎

In particular, Theorem 7.2 implies that

$$d^{(K)}\left(\sum_{l\in[k]} A_l\right) \le 2\max_{i\in[k]} d^{(K)}\left(\sum_{l\in[k]\setminus\{i\}} A_l\right),$$

and, when the sets are the same,

$$d^{(K)}(A(k)) \le 2\frac{k-1}{k} d^{(K)}(A(k-1)). \tag{33}$$

While not proving monotonicity of $d^{(K)}(A(k))$, the inequality (33) does provide a bound on extent of non-monotonicity in the sequence in general dimension.

Remark 7.3. *1. Dyn and Farkhi [30] conjectured that*

$$d^2(A+B) \le d^2(A) + d^2(B). \tag{34}$$

This seems to be still open. The power of 2 must be optimal when A and B are distinct, as can be seen from the following example: consider $A = \{0,1\} \times \{0\} = \{(0,0),(1,0)\}$ and $B = \{0\} \times \{0,1\} = \{(0,0),(0,1)\}$ (in the plane). Thus, $A+B$ is the unit discrete cube $\{0,1\}^2$. Now $d(A) = d(B) = 1/2$, and $d(A+B) = 1/\sqrt{2}$, hence $d^2(A+B) = d^2(A) + d^2(B)$.

2. *As shown by Wegmann [75], if the set A is such that the supremum in the definition of $v(A)$ is achieved at a point in the relative interior of* $\mathrm{conv}(A)$, *then $d(A) = v(A)$. Thus Theorem 6.1 implies the following statement: If A, B are compact sets in $\mathbf{R}^n$ such that the supremum in the definition of $v(A)$ is achieved at a point in the relative interior of* $\mathrm{conv}(A)$, *and likewise for B, then*

$$d^2(A+B) \le d^2(A) + d^2(B).$$

3. *One cannot expect a better upper bound than $1/\sqrt{2}$ for the quotient $d(\frac{A+A}{2})/d(A)$ in general as can be seen from the example where $A = \{a_1, \cdots, a_{n+1}\}$ is a set of $n+1$ vertices of a regular simplex in $\mathbf{R}^n$, $n \ge 2$. In this case, it is not difficult to see that $d(A) = |g - a_1|$, where $g = (a_1 + \cdots + a_{n+1})/(n+1)$ is the center of mass of A and $d(\frac{A+A}{2}) = |g - \frac{a_1+a_2}{2}|$. Then, one easily concludes that*

$$\frac{d(\frac{A+A}{2})}{d(A)} = \frac{|g - \frac{a_1+a_2}{2}|}{|g - a_1|} = \frac{1}{\sqrt{2}}\sqrt{1 - \frac{1}{n}}.$$

7.2 Strong fractional subadditivity for large k

In this section, similarly as for the effective standard deviation v, we prove that the Hausdorff distance from the convex hull $d^{(K)}$ satisfies a strong fractional subadditivity when considering sufficient large numbers of sets.

Theorem 7.4. *Let K be an arbitrary convex body containing 0 in its interior. Let $A_1, \dots, A_k$ be compact sets in $\mathbf{R}^n$, with $k \ge n+1$. Then,*

$$d^{(K)}\left(\sum_{i\in[k]} A_i\right) \le \max_{I\subset[k]:|I|\le n}\ \min_{i\in[k]\setminus I} d^{(K)}\left(\sum_{j\in[k]\setminus\{i\}} A_j\right).$$

Proof. Let $x \in \mathrm{conv}(\sum_{i\in[k]} A_i)$. By using the Shapley-Folkman lemma (Lemma 2.2), there exists a set $I \subset [k]$ of cardinality at most n such that

$$x \in \sum_{i\in I} \mathrm{conv}(A_i) + \sum_{i\in[k]\setminus I} A_i.$$

Let $i_0 \in [k] \setminus I$. In particular, we have

$$x \in \sum_{i\in[k]\setminus\{i_0\}} \mathrm{conv}(A_i) + A_{i_0}.$$

Thus,

$$x = \sum_{i\in[k]\setminus\{i_0\}} x_i + x_{i_0} = z + x_{i_0},$$

for some $x_i \in \mathrm{conv}(A_i)$, $i \in [k] \setminus \{i_0\}$, and some $x_{i_0} \in A_{i_0}$, where $z = \sum_{i\in[k]\setminus\{i_0\}} x_i$. Hence,

$$\begin{aligned}
d^{(K)}_{\sum_{i\in[k]\setminus\{i_0\}} A_i}(z) &= \inf_{a\in\sum_{i\in[k]\setminus\{i_0\}} A_i} \|z-a\|_K \\
&= \inf_{a\in\sum_{i\in[k]\setminus\{i_0\}} A_i} \|z + x_{i_0} - (a + x_{i_0})\|_K \\
&\geq \inf_{a^*\in\sum_{i\in[k]} A_i} \|z + x_{i_0} - a^*\|_K \\
&= d^{(K)}_{\sum_{i\in[k]} A_i}(x).
\end{aligned}$$

Taking supremum over all $z \in \mathrm{conv}(\sum_{i\in[k]\setminus\{i_0\}} A_i)$, we deduce that

$$d^{(K)}_{\sum_{i\in[k]} A_i}(x) \leq d^{(K)}(\sum_{i\in[k]\setminus\{i_0\}} A_i).$$

Since this is true for every $i_0 \in [k] \setminus I$, we deduce that

$$d^{(K)}_{\sum_{i\in[k]} A_i}(x) \leq \min_{i\in[k]\setminus I} d^{(K)}(\sum_{j\in[k]\setminus\{i\}} A_j).$$

Taking the supremum over all set $I \subset [k]$ of cardinality at most n yields

$$d^{(K)}_{\sum_{i\in[k]} A_i}(x) \leq \max_{I\subset[k]:|I|\leq n} \min_{i\in[k]\setminus I} d^{(K)}(\sum_{j\in[k]\setminus\{i\}} A_j).$$

We conclude by taking the supremum over all $x \in \mathrm{conv}(\sum_{i\in[k]} A_i)$. $\square$

In the case where $A_1 = \cdots = A_k = A$, we can use the above argument to prove that for $k \geq c(A) + 1$,

$$d^{(K)}(A(k)) \leq \frac{k-1}{k} d^{(K)}(A(k-1)),$$

where $c(A)$ is the Schneider non-convexity index of A. Since $c(A) \leq n$, and $c(A) \leq n-1$ when A is connected, we deduce the following monotonicity property for the Hausdorff distance to the convex hull.

Corollary 7.5. *Let K be an arbitrary convex body containing 0 in its interior. Then,*

1. *In dimension 1 and 2, the sequence $d^{(K)}(A(k))$ is non-increasing for every compact set A.*
2. *In dimension 3, the sequence $d^{(K)}(A(k))$ is non-increasing for every compact and connected set A.*

Remark 7.6. *It follows from the above study that if a compact set $A \subset \mathbf{R}^n$ satisfies $c(A) \leq 2$, then the sequence $d^{(K)}(A(k))$ is non-increasing. One can see that if a compact set $A \subset \mathbf{R}^n$ contains the boundary of its convex hull, then $c(A) \leq 1$; for such set $A \subset \mathbf{R}^n$, the sequence $d^{(K)}(A(k))$ is non-increasing.*

It is useful to also record a simplified version of Theorem 7.4.

Corollary 7.7. *Let K be an arbitrary convex body containing 0 in its interior. Let $A_1, \dots, A_k$ be compact sets in $\mathbf{R}^n$, with $k \geq n+1$. Then,*

$$d^{(K)}\left(\sum_{i\in[k]} A_i\right) \leq \max_{I\subset[k]:|I|=n} d^{(K)}\left(\sum_{i\in I} A_i\right) \leq n \max_{i\in[k]} d^{(K)}(A_i).$$

Proof. By Theorem 7.4, provided $k > n$, we have in particular

$$d^{(K)}\left(\sum_{i\in[k]} A_i\right) \leq \max_{i\in[k]} d^{(K)}\left(\sum_{j\neq i} A_j\right).$$

Iterating the same argument as long as possible, we have that

$$d^{(K)}\left(\sum_{i\in[k]} A_i\right) \leq \max_{I\subset[k]:|I|=n} d^{(K)}\left(\sum_{j\in I} A_j\right),$$

which is the first desired inequality. Applying the subadditivity property of $d^{(K)}$ (namely, Theorem 7.1), we immediately have the second desired inequality. □

While Corollary 7.7 does not seem to have been explicitly written down before, it seems to have been first discovered by V. Grinberg (personal communication).

7.3 Convergence rates for d

Let us first note that having proved convergence rates for $v(A(k))$, we automatically inherit convergence rates for $d^{(K)}(A(k))$ as a consequence of Lemma 2.12 and Theorem 2.14.

Corollary 7.8. *Let K be an arbitrary convex body containing 0 in its interior. For any compact set $A \subset \mathbf{R}^n$,*

$$d^{(K)}(A(k)) \leq \frac{1}{r} \min\left\{\frac{1}{\sqrt{k}}, \frac{\sqrt{n}}{k}\right\} v(A),$$

where $r > 0$ is such that $rB_2^n \subset K$.

For Euclidean norm (i.e., $K = B_2^n$), this goes back to [67, 23].

Although we have a strong convergence result for $d^{(K)}(A(k))$ as a consequence of that for $v(A(k))$, we give below another estimate of $d^{(K)}(A(k))$ in terms of $d^{(K)}(A)$, instead of $v(A)$.

Theorem 7.9. *For any compact set $A \subset \mathbf{R}^n$,*

$$d^{(K)}(A(k)) \leq \min\left\{1, \frac{\lceil c(A) \rceil}{k}\right\} d^{(K)}(A).$$

Proof. As a consequence of Theorem 7.1, we always have $d^{(K)}(A(k)) \leq d^{(K)}(A)$. Now consider $k \geq c(A) + 1$, and notice that

$$kA(k) + \lceil c(A) \rceil d^{(K)}(A)K \supset (k - \lceil c(A) \rceil)A(k - \lceil c(A) \rceil) + \lceil c(A) \rceil \text{conv}(A) = \text{conv}(kA(k)).$$

Hence $d^{(K)}(kA(k)) \leq \lceil c(A) \rceil d^{(K)}(A)$, or equivalently, $d^{(K)}(A(k)) \leq \frac{\lceil c(A) \rceil d^{(K)}(A)}{k}$. $\square$

Using the fact that $c(A) \leq n$ for every compact set $A \subset \mathbf{R}^n$, we deduce that

$$d^{(K)}(A(k)) \leq \min\left\{1, \frac{n}{k}\right\} d^{(K)}(A).$$

8 Connections to discrepancy theory

The ideas in this section have close connections to the area known sometimes as "discrepancy theory", which has arisen independently in the theory of Banach spaces, combinatorics, and computer science. It should be emphasized that there are two distinct but related areas that go by the name of discrepancy theory. The first, discussed in this section and sometimes called "combinatorial discrepancy theory" for clarity, was likely originally motivated by questions related to absolute versus unconditional versus conditional convergence for series in Banach spaces. The second, sometimes called "geometric discrepancy theory" for clarity, is related to how well a finite set of points can approximate a uniform distribution on (say) a cube in $\mathbf{R}^n$. Our discussion here concerns the former; the interested reader may consult [72] for more on the latter. When looked at deeper, however, combinatorial discrepancy theory is also related to the ability to discretely approximate "continuous" objects. For example, a famous result of Spencer [65] says that given any collection $\{S_1, \dots, S_n\}$ of subsets of $[n]$, it is possible to color the elements of $[n]$ with two colors (say, red and blue) such that

$$\left| |S_i \cap R| - \frac{|S_i|}{2} \right| \leq 3\sqrt{n},$$

for each $i \in [n]$, where $R \subset [n]$ is the set of red elements. As explained for example by Srivastava [66]

> In other words, it is possible to partition $[n]$ into two subsets so that this partition is very close to balanced on *each one* of the test sets S_i. Note that a "continuous" partition which splits each element exactly in half will be exactly balanced on each S_i; the content of Spencer's theorem is that we can get very close to this ideal situation with an actual, discrete partition which respects the wholeness of each element.

Indeed, Srivastava also explains how the recent celebrated results of Marcus, Spielman and Srivastava [52, 53] that resulted in the solution of the Kadison-Singer conjecture may be seen from a discrepancy point of view.

For any n-dimensional Banach space E with norm $\|\cdot\|_E$, define the functional

$$V(k,E) = \max_{x_1,\ldots,x_k : \|x_i\|=1\, \forall i\in[k]} \ \min_{(\varepsilon_1,\ldots,\varepsilon_k)\in\{-1,1\}^k} \Big\| \sum_{i\in[k]} \varepsilon_i x_i \Big\|_E .$$

In other words, $V(k,E)$ answers the question: for any choice of k unit vectors in E, how small are we guaranteed to be able to make the signed sum of the unit vectors by appropriately choosing signs? The question of what can be said about the numbers $V(k,E)$ was first asked[4] by A. Dvoretzky in 1963. Let us note that the same definition also makes sense when $\|\cdot\|$ is a nonsymmetric norm (i.e., satisfies $\|ax\| = a\|x\|$ for $a > 0$, positive-definiteness and the triangle inequality), and we will discuss it in this more general setting.

It is a central result of discrepancy theory [39, 10] that when E has dimension n, it always holds that $V(k,E) \le n$. To make the connection to our results, we observe that this fact actually follows from Corollary 7.7.

Theorem 8.1. *Suppose $A_1,\ldots,A_k \subset K$, where K is a convex body in $\mathbf{R}^n$ containing 0 in its interior (i.e., the unit ball of a non-symmetric norm $\|\cdot\|_K$), and suppose $0 \in \mathrm{conv}(A_i)$ and $\dim(A_i) = 1$ for each $i \in [k]$. Then there exist vectors $a_i \in A_i$ ($i \in [k]$) such that*

$$\Big\| \sum_{i\in[k]} a_i \Big\|_K \le n.$$

In particular, by choosing $A_i = \{x_i, -x_i\}$, with $\|x_i\|_K = 1$, one immediately has $V(k,E_K) \le n$ for $E_K = (\mathbf{R}^n, \|\cdot\|_K)$.

Proof. We simply observe that since $0 \in \mathrm{conv}(\sum_{i\in[k]} A_i)$, there exists a point $a \in \sum_{i\in[k]} A_i$ such that

$$\|a\|_K \le \sup_{x\in\mathrm{conv}(\sum_{i\in[k]} A_i)} \ \inf_{a\in\sum_{i\in[k]} A_i} \|a-x\|_K = d^{(K)}\Big(\sum_{i\in[k]} A_i\Big) \le n \max_{i\in[k]} d^{(K)}(A_i),$$

where the last inequality follows from Corollary 7.7. Moreover, using that for each $i \in [k]$, $A_i \subseteq K$ and K is convex, we get $\mathrm{conv}(A_i) \subseteq K$. Thus by Lemmata 2.11 and 2.13, $d^{(K)}(A_i) \le d^{(\mathrm{conv}(A_i))}(A_i) \le c(A_i) \le 1$, where the last inequality uses Theorem 2.9 and the assumption that $\dim(A_i) = 1$. ☐

Remark 8.2. *Bárány and Grinberg [10] proved Theorem 8.1 without the condition $\dim(A_i) = 1$. They also proved it for symmetric bodies K under the weaker condition that $0 \in \mathrm{conv}(\sum_{i\in[k]} A_i)$; we will recover this fact as a consequence of Theorem 8.6 below.*

Remark 8.3. *As pointed out in [10], Theorem 8.1 is sharp. By taking $E = \ell_1^n$ and x_i to be the i-th standard basis vector e_i of $\mathbf{R}^n$, we see that for any choice of signs, $\|\sum_{i\in[n]} \varepsilon_i x_i\| = n$, which implies that $V(n,\ell_1^n) = n$.*

Remark 8.4. *It is natural to think that the sequence $V(k,E)$ may be monotone with respect to k. Unfortunately, this is not true. Swanepoel [69] showed that $V(k,E) \le 1$ for every odd k and every 2-dimensional Banach space E. Consequently, we have $V(1,\ell_1^2) = 1$ and $V(3,\ell_1^2) \le 1$, whereas we know from Remark 8.3 that $V(2,\ell_1^2) = 2$.*

[4] See [44, p. 496] where this question is stated as one in a collection of then-unsolved problems.

Not surprisingly, for special norms, better bounds can be obtained. In particular (see, e.g., [2, Theorem 2.4.1] or [13, Lemma 2.2]), $V(k,\ell_2^n)\le\sqrt{n}$. We will shortly present a proof of this, and more general, facts. But first let us discuss a quite useful observation about the quantity $V(k,E)$: it is an isometric invariant, i.e., invariant under nonsingular linear transformations of the unit ball. A way to measure the extent of isometry is using the Banach-Mazur distance d_{BM}: Let E, E' be two n-dimensional normed spaces. The Banach-Mazur distance between them is defined as

$$d_{BM}(E,E')=\inf\{\|T\|\cdot\|T^{-1}\|;T:E\to E' \text{ isomorphism}\}.$$

Thus $d_{BM}(E,E')\ge 1$ and $d_{BM}(E,E')=1$ if and only if E and E' are isometric. We also remind that the above notion have a geometrical interpretation. Indeed if we denote by $B(X)$ a unit ball of Banach space X, then $d_{BM}(E,E')$ is a minimal positive number such that there exists a linear transformation T with:

$$B(E)\subseteq T(B(E'))\subseteq d_{BM}(E,E')B(E).$$

Lemma 8.5. *If $d_{BM}(E,E')=1$, then*

$$V(k,E)=V(k,E').$$

Proof. Consider an invertible linear transformation T such that $T(B(E))=B(E')$ and thus $\|y\|_E=\|Ty\|_{E'}$, then

$$\begin{aligned}
V(k,E) &= \max_{x_1,\dots,x_k:\|x_i\|_E=1\,\forall i\in[k]}\ \min_{(\varepsilon_1,\dots,\varepsilon_k)\in\{-1,1\}^k}\left\|\sum_{i\in[k]}\varepsilon_i x_i\right\|_E\\
&= \max_{x_1,\dots,x_k:\|Tx_i\|_{E'}=1\,\forall i\in[k]}\ \min_{(\varepsilon_1,\dots,\varepsilon_k)\in\{-1,1\}^k}\left\|T\left(\sum_{i\in[k]}\varepsilon_i x_i\right)\right\|_{E'}\\
&= \max_{y_1,\dots,y_k:\|y_i\|_{E'}=1\,\forall i\in[k]}\ \min_{(\varepsilon_1,\dots,\varepsilon_k)\in\{-1,1\}^k}\left\|\sum_{i\in[k]}\varepsilon_i y_i\right\|_{E'}
\end{aligned}$$

□

Now we would like to use the ideas of the proof of Theorem 8.1 together with Lemma 8.5 to prove the following statement that will help us to provide sharper bounds for $V(k,E)$ for intermediate norms.

Theorem 8.6. *Suppose $A_1,\dots,A_k\subset K$, where K is a symmetric convex body in $\mathbf{R}^n$ (i.e., the unit ball of a norm $\|\cdot\|_K$), and suppose $0\in\operatorname{conv}(\sum_{i\in[k]}A_i)$. Then there exist vectors $a_i\in A_i$ ($i\in[k]$) such that*

$$\left\|\sum_{i\in[k]}a_i\right\|_K\le\sqrt{n}\,d_{BM}(E,\ell_2^n),$$

where $E=(\mathbf{R}^n,\|\cdot\|_K)$. In particular, by choosing $A_i=\{x_i,-x_i\}$, with $\|x_i\|_K=1$, one immediately has

$$V(k,E)\le\sqrt{n}\,d_{BM}(E,\ell_2^n).$$

Proof. Let $d = d_{BM}(E, \ell_2^n)$, then we may assume, using Lemma 8.5, that $B_2^n \subset K \subset dB_2^n$. Next, as in the proof of Theorem 8.1 we observe that since $0 \in \operatorname{conv}(\sum_{i\in[k]} A_i)$, there exists a point $a \in \sum_{i\in[k]} A_i$ such that

$$\|a\|_K \le d^{(K)}\left(\sum_{i\in[k]} A_i\right) \le \max_{I\subset[k]:|I|=n} d^{(K)}\left(\sum_{j\in I} A_j\right),$$

where the last inequality follows from Corollary 7.7. Next, we apply Lemma 2.11 together with $B_2^n \subset K$ to get

$$\max_{I\subset[k]:|I|=n} d^{(K)}\left(\sum_{j\in I} A_j\right) \le \max_{I\subset[k]:|I|=n} d\left(\sum_{j\in I} A_j\right).$$

Now we can apply Theorems 2.14 and 6.1 to get

$$\max_{I\subset[k]:|I|=n} d\left(\sum_{j\in I} A_j\right) \le \max_{I\subset[k]:|I|=n} v\left(\sum_{j\in I} A_j\right) \le \max_{I\subset[k]:|I|=n} \sqrt{\sum_{j\in I} v^2(A_j)} \le d\sqrt{n},$$

where the last inequality follows from the fact that $v(A_i) = r(A_i)$ is bounded by d since $A_i \subset K \subset dB_2^n$. $\square$

We note that it follows from F. John Theorem (see, e.g., [56, page 10]) that $d_{BM}(E, \ell_2^n) \le \sqrt{n}$ for any n-dimensional Banach space E. Thus we have the following corollary, which recovers a result of [10].

Corollary 8.7. *Suppose $A_1, \dots, A_k \subset K$, where K is a convex symmetric body in $\mathbf{R}^n$, and suppose $0 \in \operatorname{conv}(\sum_{i\in[k]} A_i)$. Then there exist vectors $a_i \in A_i$ ($i \in [k]$) such that*

$$\left\|\sum_{i\in[k]} a_i\right\|_K \le n.$$

In particular, by choosing $A_i = \{x_i, -x_i\}$, with $\|x_i\|_K = 1$, one immediately has

$$V(k, E) \le n,$$

where $E = (\mathbf{R}^n, \|\cdot\|_K)$.

The fact that $V(k, E) \le n$ appears to be folklore and the first explicit mention of it we could find is in [39].

It is well known that $d_{BM}(\ell_p^n, \ell_2^n) = n^{|\frac{1}{p}-\frac{1}{2}|}$ for $p \ge 1$ (see, e.g., [56, page 20]). Thus Theorem 8.6 gives:

Corollary 8.8. *For any $p \ge 1$ and any $n \in \mathbb{N}$,*

$$V(k, \ell_p^n) \le n^{\frac{1}{2}+|\frac{1}{p}-\frac{1}{2}|}.$$

In particular, we recover the classical fact that $V(k, \ell_2^n) \leq \sqrt{n}$, which can be found, e.g., in [2, Theorem 2.4.1]. V. Grinberg (personal communication) informed us of the following elegant and sharp bound generalizing this fact that he obtained in unpublished work: if A_i are subsets of $\mathbf{R}^n$ and $D = \max_i \operatorname{diam}(A_i)$, then

$$d\left(\sum_{i\in[k]} A_i\right) \leq \frac{D}{2}\sqrt{n}. \tag{35}$$

The special case of this when each A_i has cardinality 2 is due to Beck [13]. Let us note that the inequality (35) improves upon the bound of $\sqrt{n}\max_i v(A_i)$ that is obtained in the Shapley-Folkman theorem by combining Theorems 2.14 and 6.6.

Finally let us note that the fact that the quantities $V(k, E)$ are $O(n)$ for general norms and $O(\sqrt{n})$ for Euclidean norm is consistent with the observations in Section 7.3 that the rate of convergence of $d^{(K)}(A(k))$ for a compact set $A \subset \mathbf{R}^n$ is $O(n/k)$ for general norms and $O(\sqrt{n}/k)$ for Euclidean norm (i.e., $K = B_2^n$).

We do not comment further on the relationship of our study with discrepancy theory, which contains many interesting results and questions when one uses *different* norms to pick the original unit vectors, and to measure the length of the signed sum (see, e.g., [14, 37, 58]). The interested reader may consult the books [24, 54, 25] for more in this direction, including discussion of algorithmic issues and applications to theoretical computer science.

9 Discussion

Finally we mention some notions of non-convexity that we do not take up in this paper:

1. Inverse reach: The notion of reach was defined by Federer [32], and plays a role in geometric measure theory. For a set A in $\mathbf{R}^n$, the reach of A is defined as

$$\operatorname{reach}(A) = \sup\{r > 0 : \forall y \in A + rB_2^n, \text{ there exists a unique } x \in A \text{ nearest to } y\}.$$

A key property of reach is that $\operatorname{reach}(A) = \infty$ if and only if A is convex; consequently one may think of

$$\iota(A) = \operatorname{reach}(A)^{-1}$$

as a measure of non-convexity. Thäle [71] presents a comprehensive survey of the study of sets with positive reach (however, one should take into account the cautionary note in the review of this article on MathSciNet).

2. Beer's index of convexity: First defined and studied by Beer [15], this quantity is defined for a compact set A in $\mathbf{R}^n$ as the probability that 2 points drawn uniformly from A at random "see" each other (i.e., the probability that the line segment connecting them is in A). Clearly this probability is 1 for convex sets, and 0 for finite sets consisting of more than 1 point. Since our study has been framed in terms of measures of non-convexity, it is more natural to consider

$$b(A) = 1 - \mathbf{P}\{[X, Y] \subset A\},$$

where X, Y are i.i.d. from the uniform measure on A, and $[x, y]$ denotes the line segment connecting x and y.

3. Convexity ratio: The convexity ratio of a set A in $\mathbf{R}^n$ is defined as the ratio of the volume of a largest convex subset of A to the volume of A– it is clearly 1 for convex sets and can be arbitrarily close to 0 otherwise. For dimension 2, this has been studied, for example, by Goodman [38]. Balko et al. [8] discuss this notion in general dimension, and also give some inequalities relating the convexity ratio and Beer's index of convexity. Once again, to get a measure of non-convexity, it is more natural to consider

$$\kappa(A) = 1 - \frac{\mathrm{Vol}_n(L(A))}{\mathrm{Vol}_n(A)},$$

where $L(A)$ denotes a largest convex subset of A.

These notions of non-convexity are certainly very interesting, but they behave quite differently from the notions we have explored thus far. For example, if $b(A) = 0$ or $\kappa(A) = 0$, the compact set A may not be convex, but differ from a convex set by a set of measure zero. For example, if A is the union of a unit Euclidean ball and a point separated from it, then

$$b(A) = \kappa(A) = 0, \tag{36}$$

even though A is compact but non-convex. Even restricting to compact *connected* sets does not help– just connect the disc with a point by a segment, and we retain (36) though A remains non-convex.

It is possible that further restricting to connected open sets is the right thing to do here– this may yield a characterization of convex sets using b and κ, but it still is not enough to ensure stability of such a characterization. For example, $b(A)$ small would not imply that A is close to its convex hull even for this restricted class of sets, because we can take the previous example of a point connected to a disc by a segment and just slightly fatten the segment.

Generalizing this example leads to a curious phenomenon. Consider $A = B_2^n \cup \{x_1, ..., x_N\}$, where $x_1, \dots, x_N$ are points in $\mathbf{R}^n$ well separated from each other and the origin. Then $b(A) = \kappa(A) = 0$, but we can send $b(\frac{A+A}{2})$ and $\kappa(\frac{A+A}{2})$ arbitrarily close to 1 by making N go to infinity (since isolated points are never seen for A but become very important for the sumset). This is remarkably bad behavior indeed, since it indicates an extreme violation of the monotone decreasing property of $b(A(k))$ or $\kappa(A(k))$ that one might wish to explore, already in dimension 2.

Based on the above discussion, it is clear that the measures ι, b, κ of non-convexity are more sensitive to the topology of the set than the functionals we considered in most of this paper. Thus it is natural that the behavior of these additional measures for Minkowski sums should be studied with a different global assumption than in this paper (which has focused on what can be said for compact sets). We hope to investigate this question in future work.

References

[1] S. Alesker, S. Dar, and V. Milman. A remarkable measure preserving diffeomorphism between two convex bodies in $\mathbf{R}^n$. *Geom. Dedicata*, 74(2):201–212, 1999.

[2] N. Alon and J. H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul ErdHos.

[3] R. M. Anderson. An elementary core equivalence theorem. *Econometrica*, 46(6):1483–1487, 1978.

[4] S. Artstein, K. M. Ball, F. Barthe, and A. Naor. Solution of Shannon's problem on the monotonicity of entropy. *J. Amer. Math. Soc.*, 17(4):975–982 (electronic), 2004.

[5] Z. Artstein. Discrete and continuous bang-bang and facial spaces or: look for the extreme points. *SIAM Rev.*, 22(2):172–185, 1980.

[6] Z. Artstein and R. A. Vitale. A strong law of large numbers for random compact sets. *Ann. Probab.*, 3(5):879–882, 1975.

[7] S. Artstein-Avidan, D. Florentin, and Y. Ostrover. Remarks about mixed discriminants and volumes. *Commun. Contemp. Math.*, 16(2):1350031, 14, 2014.

[8] M. Balko, V. Jelínek, P. Valtr, and B. Walczak. On the Beer index of convexity and its variants. *Preprint,* `arXiv:1412.1769`, 2014.

[9] K. Ball, P. Nayar, and T. Tkocz. A reverse entropy power inequality for log-concave random vectors. *Studia Math.*, 235(1):17–30, 2016.

[10] I. Bárány and V. S. Grinberg. On some combinatorial questions in finite-dimensional spaces. *Linear Algebra Appl.*, 41:1–9, 1981.

[11] A.R. Barron. Entropy and the central limit theorem. *Ann. Probab.*, 14:336–342, 1986.

[12] F. Barthe, M. Madiman, and L. Wang. Notes on a fractional Brunn-Minkowski-type inequality. *Preprint*, 2016.

[13] J. Beck. On a geometric problem of ErdHos, Sárközy, and Szemerédi concerning vector sums. *European J. Combin.*, 4(1):1–10, 1983.

[14] J. Beck and T. Fiala. "Integer-making" theorems. *Discrete Appl. Math.*, 3(1):1–8, 1981.

[15] G. Beer. The index of convexity and the visibility function. *Pacific J. Math.*, 44:59–67, 1973.

[16] D. P. Bertsekas. *Convex optimization theory.* Athena Scientific, Nashua, NH, 2009.

[17] S. Bobkov and M. Madiman. Dimensional behaviour of entropy and information. *C. R. Acad. Sci. Paris Sér. I Math.*, 349:201–204, Février 2011.

[18] S. Bobkov and M. Madiman. Reverse Brunn-Minkowski and reverse entropy power inequalities for convex measures. *J. Funct. Anal.*, 262:3309–3339, 2012.

[19] S. G. Bobkov, M. Madiman, and L. Wang. Fractional generalizations of Young and Brunn-Minkowski inequalities. In C. Houdré, M. Ledoux, E. Milman, and M. Milman, editors, *Concentration, Functional Inequalities and Isoperimetry*, volume 545 of *Contemp. Math.*, pages 35–53. Amer. Math. Soc., 2011.

[20] S. G. Bobkov and M. M. Madiman. On the problem of reversibility of the entropy power inequality. In *Limit theorems in probability, statistics and number theory*, volume 42 of *Springer Proc. Math. Stat.*, pages 61–74. Springer, Heidelberg, 2013. Available online at `arXiv:1111.6807`.

[21] T. Bonnesen and W. Fenchel. *Theory of convex bodies.* BCS Associates, Moscow, ID, 1987. Translated from the German and edited by L. Boron, C. Christenson and B. Smith.

[22] K. Böröczky, Jr., M. A. Hernández Cifre, and G. Salinas. Optimizing area and perimeter of convex sets for fixed circumradius and inradius. *Monatsh. Math.*, 138(2):95–110, 2003.

[23] J. W. S. Cassels. Measures of the non-convexity of sets and the Shapley-Folkman-Starr theorem. *Math. Proc. Cambridge Philos. Soc.*, 78(3):433–436, 1975.

[24] B. Chazelle. *The discrepancy method.* Cambridge University Press, Cambridge, 2000. Randomness and complexity.

[25] W. Chen, A. Srivastav, and G. Travaglini, editors. *A panorama of discrepancy theory*, volume 2107 of *Lecture Notes in Mathematics.* Springer, Cham, 2014.

[26] M.H.M. Costa and T.M. Cover. On the similarity of the entropy power inequality and the Brunn-Minkowski inequality. *IEEE Trans. Inform. Theory*, 30(6):837–839, 1984.

[27] T.M. Cover and J.A. Thomas. *Elements of Information Theory.* J. Wiley, New York, 1991.

[28] A. Dembo, T.M. Cover, and J.A. Thomas. Information-theoretic inequalities. *IEEE Trans. Inform. Theory*, 37(6):1501–1518, 1991.

[29] J. Diestel and J. J. Uhl, Jr. *Vector measures.* American Mathematical Society, Providence, R.I., 1977. With a foreword by B. J. Pettis, Mathematical Surveys, No. 15.

[30] N. Dyn and E. Farkhi. Set-valued approximations with Minkowski averages—convergence and convexification rates. *Numer. Funct. Anal. Optim.*, 25(3-4):363–377, 2004.

[31] W. R. Emerson and F. P. Greenleaf. Asymptotic behavior of products $C^p = C + \cdots + C$ in locally compact abelian groups. *Trans. Amer. Math. Soc.*, 145:171–204, 1969.

[32] H. Federer. Curvature measures. *Trans. Amer. Math. Soc.*, 93:418–491, 1959.

[33] M. Fradelizi, A. Giannopoulos, and M. Meyer. Some inequalities about mixed volumes. *Israel J. Math.*, 135:157–179, 2003.

[34] M. Fradelizi, M. Madiman, A. Marsiglietti, and A. Zvavitch. Do Minkowski averages get progressively more convex? *C. R. Acad. Sci. Paris Sér. I Math.*, 354(2):185–189, February 2016.

[35] M. Fradelizi and A. Marsiglietti. On the analogue of the concavity of entropy power in the Brunn-Minkowski theory. *Adv. in Appl. Math.*, 57:1–20, 2014.

[36] R. J. Gardner. The Brunn-Minkowski inequality. *Bull. Amer. Math. Soc. (N.S.)*, 39(3):355–405 (electronic), 2002.

[37] A. A. Giannopoulos. On some vector balancing problems. *Studia Math.*, 122(3):225–234, 1997.

[38] J. E. Goodman. On the largest convex polygon contained in a nonconvex n-gon, or how to peel a potato. *Geom. Dedicata*, 11(1):99–106, 1981.

[39] V. S. Grinberg and S. V. Sevast'janov. Value of the Steinitz constant. *Funktsional. Anal. i Prilozhen.*, 14(2):56–57, 1980.

[40] K. Gyarmati, M. Matolcsi, and I. Z. Ruzsa. A superadditivity and submultiplicativity property for cardinalities of sumsets. *Combinatorica*, 30(2):163–174, 2010.

[41] G.H. Hardy, J.E. Littlewood and G. Pólya. Inequalities. 2d ed. *Cambridge, at the University Press,* 1952.

[42] H. Jung. Ueber die kleinste Kugel, die eine räumliche Figur einschliesst. *J. Reine Angew. Math.*, 123:241–257, 1901.

[43] M. A. Khan and K. P. Rath. The Shapley-Folkman theorem and the range of a bounded measure: an elementary and unified treatment. *Positivity*, 17(3):381–394, 2013.

[44] V. Klee, editor. *Proceedings of Symposia in Pure Mathematics. Vol. VII: Convexity.* Held at the University of Washington, Seattle, Wash., June 13-15, 1961. American Mathematical Society, Providence, R.I., 1963.

[45] E. Lutwak, D. Yang, and G. Zhang. Moment-entropy inequalities. *Ann. Probab.*, 32(1B):757–774, 2004.

[46] M. Madiman and A.R. Barron. The monotonicity of information in the central limit theorem and entropy power inequalities. In *Proc. IEEE Intl. Symp. Inform. Theory*, pages 1021–1025. Seattle, July 2006.

[47] M. Madiman and A.R. Barron. Generalized entropy power inequalities and monotonicity properties of information. *IEEE Trans. Inform. Theory*, 53(7):2317–2329, July 2007.

[48] M. Madiman and F. Ghassemi. Combinatorial entropy power inequalities: A preliminary study of the Stam region. *Preprint,* `arXiv:1704.01177`, 2017.

[49] M. Madiman and I. Kontoyiannis. Entropy bounds on abelian groups and the Ruzsa divergence. *Preprint,* `arXiv:1508.04089`, 2015.

[50] M. Madiman, J. Melbourne, and P. Xu. Forward and reverse entropy power inequalities in convex geometry. In E. A. Carlen, M. Madiman, and E. Werner, editors, *Convexity and Concentration*, IMA Volumes in Mathematics and its Applications. Springer, to appear. Available online at `arXiv:1604:04225`.

[51] M. Madiman and P. Tetali. Information inequalities for joint distributions, with interpretations and applications. *IEEE Trans. Inform. Theory*, 56(6):2699–2713, June 2010.

[52] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing families I: Bipartite Ramanujan graphs of all degrees. *Ann. of Math. (2)*, 182(1):307–325, 2015.

[53] A. W. Marcus, D. A. Spielman, and N. Srivastava. Interlacing families II: Mixed characteristic polynomials and the Kadison-Singer problem. *Ann. of Math. (2)*, 182(1):327–350, 2015.

[54] J. Matoušek. *Geometric discrepancy*, volume 18 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 2010. An illustrated guide, Revised paperback reprint of the 1999 original.

[55] J. Matoušek. *Lectures on discrete geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

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

[57] J. Moulin Ollagnier and D. Pinchon. Filtre moyennant et valeurs moyennes des capacités invariantes. *Bull. Soc. Math. France*, 110(3):259–277, 1982.

[58] A. Nikolov. The Komlós conjecture holds for vector colorings. *Preprint,* `arXiv:1301.4039`, 2013.

[59] J. R. Sangwine-Yager. A Bonnesen-style inradius inequality in 3-space. *Pacific J. Math.*, 134(1):173–178, 1988.

[60] L. D. W. Schmidt. On the dimensionality of bounds generated by the Shapley-Folkman theorem. *J. Math. Econom.*, 48(1):59–63, 2012.

[61] R. Schneider. A measure of convexity for compact sets. *Pacific J. Math.*, 58(2):617–625, 1975.

[62] R. Schneider. *Convex bodies: the Brunn-Minkowski theory*, volume 151 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, expanded edition, 2014.

[63] R. Schneider and W. Weil. *Stochastic and integral geometry*. Probability and its Applications (New York). Springer-Verlag, Berlin, 2008.

[64] D. Shlyakhtenko. A free analogue of Shannon's problem on monotonicity of entropy. *Adv. Math.*, 208(2):824–833, 2007.

[65] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289(2):679–706, 1985.

[66] N. Srivastava. Discrepancy, Graphs, and the Kadison-Singer Problem. https://windowsontheory.org/2013/07/11/discrepancy-graphs-and-the-kadison-singer-conjecture-2/

[67] R. M. Starr. Quasi-equilibria in markets with non-convex preferences. *Econometrica*, 37(1):25–38, January 1969.

[68] R. M. Starr. Approximation of points of convex hull of a sum of sets by points of the sum: an elementary approach. *J. Econom. Theory*, 25(2):314–317, 1981.

[69] K. J. Swanepoel. Balancing unit vectors. *J. Combin. Theory Ser. A*, 89(1):105–112, 2000.

[70] S. J. Szarek and D. Voiculescu. Shannon's entropy power inequality via restricted Minkowski sums. In *Geometric aspects of functional analysis*, volume 1745 of *Lecture Notes in Math.*, pages 257–262. Springer, Berlin, 2000.

[71] C. Thäle. 50 years sets with positive reach—a survey. *Surv. Math. Appl.*, 3:123–165, 2008.

[72] G. Travaglini. *Number theory, Fourier analysis and geometric discrepancy*, volume 81 of *London Mathematical Society Student Texts*. Cambridge University Press, Cambridge, 2014.

[73] A. M. Tulino and S. Verdú. Monotonic decrease of the non-gaussianness of the sum of independent random variables: A simple proof. *IEEE Trans. Inform. Theory*, 52(9):4295–7, September 2006.

[74] L. Wang and M. Madiman. Beyond the entropy power inequality, via rearrangements. *IEEE Trans. Inform. Theory*, 60(9):5116–5137, September 2014.

[75] R. Wegmann. Einige Maßzahlen für nichtkonvexe Mengen. *Arch. Math. (Basel)*, 34(1):69–74, 1980.

Matthieu Fradelizi
Laboratoire d'Analyse et de Mathématiques Appliquées UMR 8050
Université Paris-Est Marne-la-Vallée
5 Bd Descartes
Champs-sur-Marne
77454 Marne-la-Vallée Cedex 2, FRANCE
E-mail address: matthieu.fradelizi@u-pem.fr

Mokshay Madiman
University of Delaware
Department of Mathematical Sciences
501 Ewing Hall
Newark, DE 19716, USA
E-mail address: madiman@udel.edu

Arnaud Marsiglietti
Institute for Mathematics and its Applications
University of Minnesota
207 Church Street SE, 434 Lind Hall,
Minneapolis, MN 55455, USA
E-mail address: arnaud.marsiglietti@ima.umn.edu

Artem Zvavitch
Department of Mathematical Sciences
Kent State University
Kent, OH 44242, USA
E-mail address: zvavitch@math.kent.edu