Projections, Entropy and Sumsets

Paul Balister[*] Béla Bollobás[*†‡]

July 1, 2011

Abstract

In this paper we shall generalize Shearer's entropy inequality and its recent extensions by Madiman and Tetali, and shall apply projection inequalities to deduce extensions of some of the inequalities concerning sums of sets of integers proved recently by Gyarmati, Matolcsi and Ruzsa. We shall also discuss projection and entropy inequalities and their connections.

1 Introduction

In 1949, Loomis and Whitney [14] proved a fundamental inequality bounding the volume of a body in terms of its $(n-1)$-dimensional projections. Over forty years later, this inequality was extended considerably by Bollobás and Thomason [4]: they showed that a certain 'box' is a solution of much more general isoperimetric problems.

In 1978, Han [12] proved the exact analogue of the Loomis-Whitney inequality for the entropy of a family $\{X_1, \dots, X_n\}$ of random variables, and in the same year Shearer proved (implicitly) a considerable extension of this inequality, namely the entropy analogue of the projection inequality that was to be used some years later in [4] to deduce the Box Theorem. (This extension

[*] Department of Mathematical Sciences, University of Memphis, Memphis TN 38152, USA

[†] Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, UK

[‡] Research supported in part by NSF grants CCR-0225610, DMS-0505550 and W911NF-06-1-0076

was published only in 1986, in [6].) Recently, Madiman and Tetali [18, 19] strengthened Shearer's inequality to a two-sided inequality concerning the joint entropy $H(X_1, \dots, X_n)$.

In this paper we have two main aims. The first is to prove an entropy inequality that extends *both* sides of the Madiman-Tetali inequality. Surprisingly, this inequality is not only *much* more general than the earlier inequalities, but is also just about *trivial.* Our second aim is to point out that the projection inequalities imply extensions of some very recent inequalities of Ruzsa [20], and Gyarmati, Matolcsi and Ruzsa [10] (see also its follow-up, [9]) concerning sums of sets of integers.

Our paper is organized as follows. In the next two sections we shall review some of the projection and entropy inequalities. In Section 4 we shall prove our extremely simple but very general entropy inequality extending those of Shearer, and Madiman and Tetali. In Section 5 we shall turn to sumsets, and continue the work of Gyarmati, Matolcsi and Ruzsa. Finally, in Section 6, we shall state some related unsolved problems.

2 Projection inequalities

As in [4], we call a compact subset of $\mathbb{R}^n$ which is the closure of its interior a *body*, and write $\{e_1, \dots, e_n\}$ for the canonical basis of $\mathbb{R}^n$. Given a body $K \subseteq \mathbb{R}^n$ and a set $A \subseteq [n] = \{1, \dots, n\}$ of d indices, we denote by K_A the orthogonal projection of K to the linear span of the vectors e_i, $i \in A$, and write $|K_A|$ for its d-dimensional Euclidean volume. (In particular, $K_{[n]} = K$.) The volumes $|K_A|$ can be viewed as a measure of the 'perimeter' of K. In 1949, Loomis and Whitney [14] (see also [1], [5, page 95] and [11, page 162]) proved the following isoperimetric inequality:

$$|K|^{n-1} \leq \prod_{i=1}^{n} |K_{[n]\setminus\{i\}}|. \tag{1}$$

Close to fifty years later, Bollobás and Thomason [4] proved the following *Box Theorem* showing that for the *set* of projection volumes $|K_A|$, $A \subseteq [n]$, the solution of the isoperimetric problem is a *box*, i.e., a rectangular parallelepiped whose sides are parallel to the coordinate axes.

Theorem 1. *Given a body $K \subseteq \mathbb{R}^n$, there is a box $B \subseteq \mathbb{R}^n$ with $|K| = |B|$ and $|K_A| \geq |B_A|$ for every $A \subseteq [n]$.* □

This theorem is equivalent to the assertion that there exist constants $k_i \geq 0$ such that

$$|K| = \prod_{i=1}^{n} k_i \quad \text{and} \quad |K_A| \geq \prod_{i \in A} k_i \text{ for all } A \subseteq [n]. \tag{2}$$

An immediate consequence of Theorem 1 is that, if the volume of a box can be bounded in terms of the volumes of a certain collection of projections, then the same bound will be valid for all bodies. In particular, the Loomis-Whitney Inequality (1) is an immediate consequence of the Box Theorem. In fact, the Box Theorem was deduced from the Uniform Cover Inequality, which is an even more obvious extension of (1). To state this inequality, we call a multiset $\mathcal{A}$ of subsets of $[n]$ such that each element $i \in [n]$ is in at least k of the members of $\mathcal{A}$ a *k-cover* of $[n]$. A *k-uniform cover* or *uniform k-cover* is one in which every element is in precisely k members of $\mathcal{A}$. Thus the sets $[n] \setminus \{i\}$ appearing in the Loomis-Whitney inequality (1) form an $(n-1)$-uniform cover of $[n]$. The Uniform Cover Inequality states that if K is a body in $\mathbb{R}^n$ and $\mathcal{A}$ is a k-uniform cover of $[n]$ then

$$|K|^k \leq \prod_{A \in \mathcal{A}} |K_A|. \tag{3}$$

Clearly, the Uniform Cover Inequality is a trivial consequence of the Box Theorem. Uniformity *is* needed for (3) to hold: if $\mathcal{A}$ is not k-uniform, then (3) does not hold for every body K, not even if $\mathcal{A}$ is a k-cover. Indeed, if $|K_A| < 1$ for some A, then we can add an arbitrary number of copies of A to $\mathcal{A}$, making the right hand side of (3) arbitrarily small.

The inequality (3) can in fact be strengthened slightly. Assume that A and B are disjoint subsets of $[n]$, and suppose that $B = \{i_1, i_2, \ldots, i_r\}$. Write $K(x_{i_1}, x_{i_2}, \ldots, x_{i_r})$ for the $(n-r)$-dimensional body consisting of those points of K for which the i_j coordinate is equal to x_{i_j}, $j = 1, \ldots, r$. Thus $K(x_{i_1}, x_{i_2}, \ldots, x_{i_r})$ is the intersection of K with a certain $(n-r)$-dimensional hyperplane. Define the quantity $|K_{A|B}|$ by

$$|K_{A|B}| = \sup_{x_{i_1}, \ldots, x_{i_r}} |K(x_{i_1}, \ldots, x_{i_r})_A|.$$

In other words, $|K_{A|B}|$ is the largest projection onto the A coordinates of a slice of K obtained by fixing the B coordinates. It is clear that for any $B' \supseteq B$ disjoint from A we have

$$|K_{A|B'}| \leq |K_{A|B}| \leq |K_A|. \tag{4}$$

Theorem 2. *If K is a body in $\mathbb{R}^n$ and $\mathcal{A}$ is a k-uniform cover of $[n]$ then*

$$|K|^k \le \prod_{A\in\mathcal{A}} |K_{A|A_*}|, \tag{5}$$

where $A_* = \{i \mid 1 \le i < a\}$ with $a = \min(A)$ being the smallest element of A.

Proof. The proof follows that of the proof of (3). We use induction on the dimension n, the case $n = 0$ being trivial (as all volumes are equal to 1 when $K \neq \emptyset$). Write $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_{1'}$ where $\mathcal{A}_1$ consists of those $A \in \mathcal{A}$ with $1 \in A$ and $\mathcal{A}_{1'}$ consists of those $A \in \mathcal{A}$ with $1 \notin A$. Now if we write $K(x_1)$ for a $(n-1)$-dimensional slice of K obtained by fixing the first coordinate, we have by induction

$$\begin{aligned}
|K(x_1)|^k &\le \prod_{A\in\mathcal{A}_{1'}} |K(x_1)_{A\,|\,A_*\setminus\{1\}}| \prod_{A\in\mathcal{A}_1} |K(x_1)_{A\setminus\{1\}\,|\,(A\setminus\{1\})_*\setminus\{1\}}| \\
&\le \prod_{A\in\mathcal{A}_{1'}} |K_{A|A_*}| \prod_{A\in\mathcal{A}_1} |K(x_1)_{A\setminus\{1\}}|,
\end{aligned}$$

where we have used (4) and the obvious inequality $|K(x_1)_{A|B}| \le |K_{A|B\cup\{1\}}|$. Now $|K_A| = \int |K(x_1)_{A\setminus\{1\}}|\,dx_1$ for any A with $1 \in A$. Thus

$$\begin{aligned}
|K| &= \int |K(x_1)|\,dx_1 \\
&\le \int \prod_{A\in\mathcal{A}_{1'}} |K_{A|A_*}|^{1/k} \prod_{A\in\mathcal{A}_1} |K(x_1)_{A\setminus\{1\}}|^{1/k}\,dx_1 \\
&\le \prod_{A\in\mathcal{A}_{1'}} |K_{A|A_*}|^{1/k} \int \prod_{A\in\mathcal{A}_1} |K(x_1)_{A\setminus\{1\}}|^{1/k}\,dx_1 \\
&\le \prod_{A\in\mathcal{A}_{1'}} |K_{A|A_*}|^{1/k} \prod_{A\in\mathcal{A}_1} \left(\int |K(x_1)_{A\setminus\{1\}}|\,dx_1\right)^{1/k} \\
&\le \prod_{A\in\mathcal{A}_{1'}} |K_{A|A_*}|^{1/k} \prod_{A\in\mathcal{A}_1} |K_A|^{1/k} = \prod_{A\in\mathcal{A}} |K_{A|A_*}|^{1/k},
\end{aligned}$$

where the second to last line follows from the Hölder inequality, using the fact that $\mathcal{A}$ is a uniform k-cover so that $|\mathcal{A}_1| = k$. The result now follows on taking kth powers. □

Following the proof of Theorem 1 given in [4], one can now strengthen Theorem 1 to the following.

Theorem 3. *Given a body $K \subseteq \mathbb{R}^n$, there exists constants $k_i \ge 0$ such that*

$$|K| = \prod_{i=1}^{n} k_i \quad \textit{and} \quad |K_{A|A_*}| \ge \prod_{i\in A} k_i \textit{ for all } A \subseteq [n]. \tag{6}$$

□

By identifying a lattice point $\mathbf{z} \in \mathbb{Z}^n$ with the unit cube $Q_{\mathbf{z}} \subseteq \mathbb{R}^n$ with centre $\mathbf{z}$, (3) implies that if S is a finite subset of $\mathbb{Z}^n$ and S_A is the projection of S to the subspace spanned by $\{e_i \colon i \in A\}$, then for every uniform k-cover $\mathcal{A}$ of $[n]$ we have

$$|S|^k \le \prod_{A\in\mathcal{A}} |S_{A|A_*}| \le \prod_{A\in\mathcal{A}} |S_A|. \tag{7}$$

In fact, for these inequalities we do not have to demand that the k-cover $\mathcal{A} = \{A_i\}$ is uniform: if $A' \subseteq A$ then $|S_{A'|A'_*}| \le |S_{A'|A_*}| \le |S_{A|A_*}|$; therefore, by removing elements from the sets A_i so as to obtain a *uniform* k-cover $\mathcal{A}' = \{A'_i\}$ with $A'_i \subseteq A_i$, we have $|S|^k \le \prod_j |S_{A'_j|(A'_j)_*}| \le \prod_j |S_{A_j|(A_j)_*}|$.

3 Entropy Inequalities

Let us turn to some entropy inequalities related to the projection inequalities above. As usual, we write $H(X)$ for the entropy of a random variable X; in particular, if X is a discrete random variable, then

$$H(X) = -\sum_k \mathbb{P}(X = k) \log_2 \mathbb{P}(X = k).$$

It is easily seen that if X takes n values then $H(X) \le \log_2 n$, with equality if and only if X is uniformly distributed, i.e., takes every value with probability $1/n$. If X and Y are two discrete random variables, then the entropy of X conditional on Y is

$$H(X \mid Y) = -\sum_{k,l} \mathbb{P}(X = k, Y = l) \log_2 \mathbb{P}(X = k \mid Y = l).$$

The entropy satisfies the following basic inequalities:

$$H(X,Y) = H(X \mid Y) + H(Y), \tag{8}$$

$$0 \le H(X \mid Y) \le H(X), \tag{9}$$

$$H(X \mid Y, Z) \le H(X \mid Y), \tag{10}$$

where, for example, we write $H(X,Y)$ for the entropy of the joint variable (X,Y).

Analogously to our notation concerning projections, given a sequence $X = (X_1, \dots, X_n)$ of n random variables, for $A \subseteq [n]$ we write $X_A = (X_i)_{i \in A}$. In 1978 Shearer proved the following analogue of (3) for entropy (the result was first published in [6]). Since $H(X_A)$ is a monotone increasing function of A, in this inequality it makes no difference whether we take $\mathcal{A}$ to be a k-cover or uniform k-cover.

Theorem 4. *If $\mathcal{A}$ is a uniform k-cover of $[n]$ then*

$$kH(X) \le \sum_{A \in \mathcal{A}} H(X_A). \tag{11}$$

A little earlier Han [12] had proved the 'Loomis-Whitney' form of Theorem 4: $(n-1)H(X) \le \sum_i H(X_{[n]\setminus\{i\}})$. The first non-trivial case of this inequality is $2H(X,Y,Z) \le H(X,Y) + H(X,Z) + H(Y,Z)$. Curiously, in [6] it is remarked that this special case can be proved analogously to what we stated as Theorem 4, and so can the case when $\mathcal{A}$ is the collection of all k-subsets of $[n]$.

Some years after the publication of [4] it was noted that Theorem 4 implies Theorem 1. In fact, the reverse implication is also easy: this follows from the fact that if $p_1, \dots, p_n$ are fixed 'probabilities' with $\sum p_i = 1$ and Np_i is an integer for every i, then the number of sequences of length N with Np_i terms equal to i is $2^{(1+o(1))H(X)N}$, where X is a random variable with $\mathbb{P}(X = i) = p_i$. Given random variables $X_1, \dots, X_n$, we may assume that X_i takes values in $V_i \subseteq \mathbb{Z}$, so that $X = (X_1, \dots, X_n)$ takes values in $V = V_1 \times \cdots \times V_n$, and there is an integer d such that $d\,\mathbb{P}(X = v)$ is an integer for every $v \in V$. Let N be a multiple of d, and let $S \subseteq V^N \subseteq \mathbb{Z}^{nN}$ be the set of all sequences in which v occurs precisely $N\,\mathbb{P}(X = v)$ times. For $A \subseteq [n]$, write $\tilde{A} \subseteq [nN]$ for the set of all coordinates of $V^N \subseteq \mathbb{Z}^{nN}$ that correspond to one of the factors V_i, $i \in A$. Then $S_{\tilde{A}}$ is the set of sequences in V_A^N where each value

$v \in V_A$ occurs $N\,\mathbb{P}(X_A = v)$ times. If $\mathcal{A}$ is a k uniform cover of $[n]$ then $\tilde{\mathcal{A}} = \{\tilde{A} : A \in \mathcal{A}\}$ is a k uniform cover of $[nN]$ and so by Theorem 1

$$|S|^k \leq \prod_{A \in \mathcal{A}} |S_{\tilde{A}}|.$$

Thus

$$2^{k(1+o(1))H(X)N} \leq \prod_{A \in \mathcal{A}} 2^{(1+o(1))H(X_A)N}$$

and Theorem 4 follows by letting $N \to \infty$.

Recently, Madiman and Tetali [18], [19] strengthened Theorem 4 by replacing the entropies $H(X_A)$ by certain conditional entropies; furthermore, they also gave lower bounds for $H(X)$.

Theorem 5. *Let $X = (X_i)_1^n$ be a sequence of random variables with $H(X)$ finite, and $\mathcal{A}$ a uniform k-cover of $[n]$. For $A \subseteq [n]$ with minimal element $a \geq 1$ and maximal element b, set $A_* = \{1, \dots, a-1\}$ and $A^* = \{i \notin A : 1 \leq i \leq b-1\}$. Then*

$$\sum_{A \in \mathcal{A}} H(X_A \mid X_{A^*}) \leq kH(X) \leq \sum_{A \in \mathcal{A}} H(X_A \mid X_{A_*}). \qquad \square$$

It should be noted that Theorem 5 does *not* follow from Shearer's Inequality, Theorem 4.

Trivially, in the lower bound $\mathcal{A}$ may be replaced by a k-packing or a fractional k-packing, and in the upper bound it may be replaced by a k-cover or a fractional k-cover, with the obvious definitions.

4 New Entropy Inequalities

Since, as shown in [4], the Box Theorem follows from the Uniform Cover Inequality (3), one has a Box Theorem type strengthening of Shearer's Inequality; in fact, there is a similar strengthening of Theorem 5 as well.

Theorem 6. *Let $X = (X_i)_1^n$ be a sequence of random variables with $H(X)$ finite. Then there are non-negative constants $h_1, \dots, h_n$ such that $H(X) = \sum_i^n h_i$ and*

$$H(X_A \mid X_{A^*}) \leq \sum_{i \in A} h_i \leq H(X_A \mid X_{A_*}) \quad \text{for all} \quad A \subseteq [n].$$

Proof. We may take $h_i = H(X_i \mid X_{[i-1]})$; to prove the inequalities, we inductively apply properties (8–10). □

Note that this theorem implies the Box Theorem in its strong form (Theorem 3). In particular, it gives a way to calculate the constants k_i explicitly rather than just proving their existence. Indeed, if we let $X = (X_1, \ldots, X_n)$ be any continuous random variable taking values in K, then $H(X) \le \log |K|$, with equality if X is uniformly distributed. (Here we are using the continuous variable version of entropy. However, by discretizing the space, we can in fact reduce it to the discrete case considered above.) Now $H(X_A \mid X_B)$, $B = \{i_1, \ldots, i_r\}$, is just a weighted average of the values of the entropies $H(Y)$, where Y is X_A conditioned on some specified value of $X_B = (x_{i_1}, \ldots, x_{i_r})$. But Y is then supported on $K(x_{i_1}, \ldots, x_{i_r})_A$, so $H(Y) \le \log |K(x_{i_1}, \ldots, x_{i_r})_A| \le \log |K_{A|B}|$. Since this holds for all choices of X_B we have $H(X_A \mid X_B) \le \log |K_{A|B}|$. Now Theorem 3 follows from the second inequality in Theorem 6 by taking X to be uniform on K.

Although Theorem 5 does not follow from Theorem 4 (Shearer's Inequality), as we shall see now, it does follow from a result which is extremely easy to prove but is still a considerable extension of Shearer's Inequality and a generalization of the submodularity of the entropy. Before we state this new inequality, we shall recall a consequence of the basic entropy inequalities, and introduce a partial order on the collection of multisets of subsets of $[n]$.

First, from (10) and (8) one can deduce that $H(X_A)$ is a *submodular* function of the set A: if $A, B \subseteq [n]$ then

$$H(X_{A\cup B}) + H(X_{A\cap B}) \le H(X_A) + H(X_B). \tag{12}$$

To see this, note that by (10) we have

$$H(X_{B\setminus A} \mid X_A) \le H(X_{B\setminus A} \mid X_{A\cap B});$$

using (8) to expand the first and last terms, we get

$$H(X_{A\cup B}) - H(X_A) \le H(X_B) - H(X_{A\cap B}),$$

which is (12).

Second, let $\mathcal{M}_{n,m}$ be the family of multisets $\mathcal{A}$ of non-empty subsets of $[n]$ with a total of m elements (i.e., $m = \sum_{A\in\mathcal{A}} |A|$). Given a multiset $\mathcal{A} = \{A_1, \ldots, A_\ell\} \in \mathcal{M}_{n,m}$ with non-nested sets A_i and A_j (thus neither

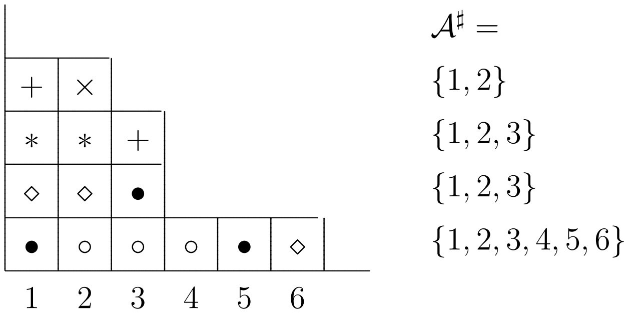

Figure 1: The minimal compression $\mathcal{A}^\sharp$ of $\mathcal{A} = \{\{2,3,4\},\{1,3,5\},\{1,2,6\},\{1,2\},\{1,3\},\{2\}\}$.

$A_i \subseteq A_j$ nor $A_j \subseteq A_i$ holds), let $\mathcal{A}' = \mathcal{A}_{(ij)}$ be obtained from $\mathcal{A}$ by replacing A_i and A_j by $A_i \cap A_j$ and $A_i \cup A_j$, keeping only $A_i \cup A_j$ if $A_i \cap A_j = \emptyset$. (If A_i and A_j are nested then replacing (A_i, A_j) by $(A_i \cap A_j, A_i \cup A_j)$ does not change $\mathcal{A}$.) We call $\mathcal{A}'$ an *elementary compression* of $\mathcal{A}$. Also, we call the result of a sequence of elementary compressions a *compression*.

Let us define a partial order on $\mathcal{M}_{n,m}$ by setting $\mathcal{A} > \mathcal{B}$ if $\mathcal{B}$ is a compression of $\mathcal{A}$. That ‘ $>'$ defines a partial order on $\mathcal{M}_{n,m}$ follows from the fact that if $\mathcal{A}'$ is an elementary compression of $\mathcal{A}$ then

$$\sum_{A\in\mathcal{A}} |A|^2 < \sum_{A\in\mathcal{A}'} |A|^2.$$

Note that for every multiset $\mathcal{A} \in \mathcal{M}_{n,m}$ there is a unique minimal multiset $\mathcal{A}^\sharp$ dominated by $\mathcal{A}$ consisting of the sets

$$A_j^\sharp = \{i \in [n] \mid i \text{ lies in at least } j \text{ of the sets } A \in \mathcal{A}\}.$$

Equivalently, $\mathcal{A}^\sharp$ is the unique multiset that is totally ordered by inclusion and has the same multiset union as $\mathcal{A}$. If we renumber $[n]$ in such a way that each $A_j^\sharp$ is an initial segment, then the $A_j^\sharp$s are just the rows of the Young tableaux associated with the set system $\mathcal{A}$, as in Figure 1. In particular, if $\mathcal{A}$ is a k-uniform cover, then $\mathcal{A}^\sharp = k\{[n]\}$.

Here is then our essentially trivial but general entropy inequality.

Theorem 7. *Let $X = (X_i)_1^n$ be a sequence of random variables with $H(X)$ finite, and let $\mathcal{A}$ and $\mathcal{B}$ be finite multisets of subsets of $[n]$. If $\mathcal{A} > \mathcal{B}$ then*

$$\sum_{A\in\mathcal{A}} H(X_A) \ge \sum_{B\in\mathcal{B}} H(X_B). \tag{13}$$

Proof. All we have to check is that (13) holds if $\mathcal{B}$ is an elementary compression of $\mathcal{A}$, i.e., if $\mathcal{B} = \mathcal{A}_{(ij)}$ for some i and j, where $\mathcal{A} = \{A_1, \dots, A_\ell\}$. But then (13) is equivalent to

$$H(X_{A_i}) + H(X_{A_j}) \geq H(X_{A_i \cap A_j}) + H(X_{A_i \cup A_j}),$$

which holds by (12), the submodularity of the entropy. □

An 'abstract' version of Theorem 7 is the triviality that if $f\colon \mathcal{P}(n) \to \mathbb{P}$ is a submodular function and $\mathcal{A}$ and $\mathcal{B}$ are finite multisets of subsets of $[n]$ with $\mathcal{A} > \mathcal{B}$ then

$$\sum_{A \in \mathcal{A}} f(A) \geq \sum_{B \in \mathcal{B}} f(B). \tag{14}$$

In particular, taking for $\mathcal{B}$ the minimal multiset $\mathcal{A}^\sharp$ dominated by $\mathcal{A}$, we have

$$\sum_{A \in \mathcal{A}} f(A) \geq \sum_{A \in \mathcal{A}^\sharp} f(A). \tag{15}$$

Not surprisingly, the method of proof of Theorem 7 is far from new: anyone working with submodular functions is bound to use it. In particular, László Lovász used it in a competition for undergraduates in Hungary in 1968 (see [15]). Later, Lovász used it in [16] and [17]; also, the quickest proof of the classical Cauchy-Davenport theorem [7] uses precisely the same idea (see [2]). Concerning abstract submodular functions, inequality (15) was noted by Frank in [8], where he also gives several applications of this method he calls *'Lovász's uncrossing method'*.

However, it is somewhat surprising that, as far as we know, the more general (and just as trivial) inequality (14) has not been noted before, although it does have applications that do not follow from (15).

Returning to entropy inequalities, we dignify the special case $\mathcal{B} = \mathcal{A}^\sharp$ of Theorem 7 corresponding to (15) by calling it a theorem. So far this form of the inequality seems to be most likely to be used.

Theorem 8. *Let $X = (X_i)_1^n$ be a sequence of random variables with $H(X)$ finite, and let $\mathcal{A}$ be a finite multiset of subsets of $[n]$. Then*

$$\sum_{A \in \mathcal{A}^\sharp} H(X_A) \leq \sum_{A \in \mathcal{A}} H(X_A). \qquad \square$$

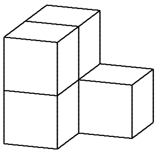

Figure 2: A body K made up of five unit cubes. Coordinate '1' is horizontal.

Let us illustrate Theorem 7 with a simple example: as $\{\{1,2\},\{1,3\},\{4\}\} > \{\{1,2,3\},\{1,4\}\}$,

$$H(X_1,X_2)+H(X_1,X_3)+H(X_4) \geq H(X_1,X_2,X_3)+H(X_1,X_4).$$

Also, let us point out that even Theorem 8 is stronger than Theorem 5, the Madiman-Tetali inequality.

Proof of Theorem 8 $\Rightarrow$ Theorem 5. Since $H(X_A \mid X_B) = H(X_{A\cup B})-H(X_B)$, the upper bound inequality is

$$kH(X)+\sum_{A\in\mathcal{A}} H(X_{A_*}) \leq \sum_{A\in\mathcal{A}} H(X_{A\cup A_*}),$$

which follows from the fact that the multiset $\mathcal{C}_1 = \{A_* : A \in \mathcal{A}\} \cup k\{[n]\}$ is totally ordered and has the same multiset union as $\mathcal{C}_2 = \{A \cup A_* : A \in \mathcal{A}\}$, so $\mathcal{C}_1 = \mathcal{C}_2^\sharp$. Similarly, the lower bound inequality is equivalent to

$$\sum_{A\in\mathcal{A}} H(X_{A\cup A^*}) \leq \sum_{A\in\mathcal{A}} H(X_{A^*})+kH(X).$$

which follows from the fact that the multiset $\mathcal{C}_3 = \{A\cup A^* : A \in \mathcal{A}\}$ is totally ordered and has the same multiset union as $\mathcal{C}_4 = \{A^* : A \in \mathcal{A}\} \cup k\{[n]\}$, so $\mathcal{C}_3 = \mathcal{C}_4^\sharp$. □

The inequality corresponding to Theorem 8 in terms of projections of bodies is false. For example, consider the set K in Figure 2. Then $|K| = 5$, $|K_{\{1\}}| = 2$, but $|K_{\{1,2\}}| = |K_{\{1,3\}}| = 3$, so

$$|K_{\{1,2,3\}}||K_{\{1\}}| > |K_{\{1,2\}}||K_{\{1,3\}}|.$$

Nevertheless, Theorem 2 is a natural analogue of Theorem 5.

5 Sumsets

Let $S_1, \ldots, S_n$ be finite sets in a commutative semigroup with sum

$$S = S_1 + \cdots + S_n = \{s_1 + \cdots + s_n : s_i \in S_i \text{ for every } i\}.$$

For $A \subseteq [n]$ set $S_A = \sum_{i \in A} S_i$, so that $S_{[n]} = S$. We shall think of S as an n-dimensional body in $\mathbb{R}^n$ and S_A as its canonical projection into the subspace spanned by $\{e_i : i \in A\}$. Gyarmati, Matolcsi and Ruzsa [10] proved the analogue of the Loomis-Whitney inequality in this context. In fact, the analogue of the Uniform Cover inequality and Box Theorem are just as easy to show.

To see this, put an arbitrary linear order on each of the sets S_i: for simplicity, we may as well just put an order on the entire semigroup and take each S_i with the induced order. For each $A = \{i_1, \ldots, i_r\} \subseteq [n]$ define an embedding φ_A of S_A into the Cartesian product $\prod_{i \in A} S_i$ by mapping $s \in S_A$ to the lexicographically least element $(s_{i_1}, \ldots, s_{i_r})$ of $\prod_{i \in A} S_i$ with coordinates summing to s. (In fact, there are many other orders we could choose instead of the lexicographic order: all we need is that the assertions below hold for these orders.) Strictly speaking, instead of φ_A we should write $\varphi_{\{S_i : i \in A\}}$, since the map φ_A depends on the sets S_i, $i \in A$, and not only on their sum $S_A = \sum_{i \in A} S_i$ or the index set A. As shown by Gyarmati, Matolcsi and Ruzsa [10], the projection of $S' = \varphi_{[n]}(S_{[n]})$ into $\prod_{i \in A} S_i$ is contained in $\varphi_A(S_A)$. To see this, note that if $(s_1, \ldots, s_n) \in S'$ then any projection $(s_{i_1}, \ldots, s_{i_r})$ is lexicographically minimal with the same sum, since if $s_{i_1} + \cdots + s_{i_r} = s'_{i_1} + \cdots + s'_{i_r}$ with $(s'_{i_1}, \ldots, s'_{i_r}) < (s_{i_1}, \ldots, s_{i_r})$, then $s_1 + \cdots + s_n = s'_1 + \cdots + s'_n$ and $(s'_1, \ldots, s'_n) < (s_1, \ldots, s_n)$ where $s'_i = s_i$ if $i \notin A$. Thus $|(S')_A| \le |\varphi_A(S_A)| = |S_A|$. Now the following result is immediate from Theorem 1 applied to S'.

Theorem 9. *There are constants $\lambda_1, \ldots, \lambda_n \ge 0$ such that*

$$|S| = \prod_1^n \lambda_i \quad \textit{and} \quad |S_A| \ge \prod_{i \in A} \lambda_i \quad \textit{for all} \quad A \subseteq [n].$$

In particular, if $\mathcal{A}$ is a uniform k-cover of $[n]$ then

$$|S|^k \le \prod_{A \in \mathcal{A}} |S_A|. \qquad \square$$

It is worth noting that one can prove a lower bound on $|S|$ which is additive in the $|S_A|$ in the case when the sets S_i lie in a torsion free abelian group. This generalizes Theorem 1.1 of [10].

Theorem 10. *If the sets S_i lie in a torsion-free abelian group then there are subsets $S'_i \subseteq S_i$ of cardinality at most 2 such that for any uniform k-cover $\mathcal{A}$ of $[n]$ we have*

$$k(|S|-1) \geq k(|S'|-1) \geq \sum_{A\in\mathcal{A}} (|S_A|-1),$$

where S' is the set of sums $s_1 + \cdots + s_k \in S$ such that $\{i : s_i \notin S'_i\} \subseteq A$ for some $A \in \mathcal{A}$.

Proof. We first note that any torsion-free abelian group can be given an ordering compatible with addition.

Pack a $k \times n$ grid with the sets $A \in \mathcal{A}$ in the obvious manner: each $A = \{j_1, \ldots, j_r\}$ is packed as a set of pairs $A' = \{(i_1, j_1), \ldots, (i_r, j_r)\}$ so that the A', $A \in \mathcal{A}$, are disjoint and cover the whole of $[k] \times [n]$. The i_ks are otherwise arbitrarily chosen.

We may assume without loss of generality that the minimum elements of S_i are all equal to 0. Let a_i be the maximum element of S_i. The set S'_i will be chosen to be $\{0, a_i\}$. For convenience write $a_T = \sum_{i\in T} a_i$. We shall mark k copies of $S' - \{0\}$ as follows.

Process each element of $[k] \times [n]$ in the lexicographic order — i.e.,

$$(1,1), \ldots, (1,n), (2,1), \ldots, (2,n), \ldots\ldots, (k,n).$$

Suppose we are processing (i,j). Then $(i,j) = (i_t, j_t)$ for some $A \in \mathcal{A}$. In the i'th copy of $S' - \{0\}$, mark all the elements that are in

$$a_{[j]-A} + S_A \cap (a_{[j-1]}, a_{[j]}].$$

Note that all elements of $a_{[j]-A} + S_A$ lie in S' (indeed in $S_{A\cup[j]} \cap S'$), and, subtracting $a_{[j]-A}$, the number of elements marked is equal to the number of elements of S_A that lie in the interval

$$(a_{[j-1]\cap A}, a_{[j]\cap A}].$$

(Note by assumption $j \in A$ so $[j] - A = [j-1] - A$). Now it is clear that for distinct (i,j), distinct elements are marked (since they all lie in the i'th copy of $S' \cap (a_{[j-1]}, a_{[j]}]$ and these sets are distinct), so at most $k(|S'|-1)$ elements are marked in total. (The element $0 \in S$ is not included in any of the intervals $(a_{[j-1]}, a_{[j]}]$.) However, every element in $S_A - \{0\}$ lies in some interval $(a_{[j-1]\cap A}, a_{[j]\cap A}]$ for some $j \in A$, so results in some element being marked. Since it is clear that $|S| \geq |S'|$, the result follows. □

Corollary 11. *If the sets S_i lie in a torsion-free abelian group then there exists constants σ_i such that*

$$|S| - 1 = \sum_{i=1}^{n} \sigma_i \quad \textit{and} \quad |S_A| - 1 \leq \sum_{i \in A} \sigma_i \textit{ for all } A \subseteq [n]. \qquad \square$$

Theorem 10 fails for groups with torsion when, for example, all S_i are equal to some non-trivial finite subgroup. If we insist that $|S|$ is smaller than the order of the smallest non-trivial subgroup then we have the famous Cauchy–Davenport theorem, which can be written in the following form.

Theorem 12. *If $S_1, \dots, S_n$ are non-empty subsets of $\mathbb{Z}_p$ and $S = S_1 + \dots + S_n$, then either $|S| \geq p$ or*

$$|S| - 1 \geq \sum_i (|S_i| - 1). \qquad \square$$

Theorem 12 is the analogue of Corollary 11 for the 1-uniform cover $\mathcal{A} = \{\{1\}, \dots, \{n\}\}$, and can be extended to all finite (even non-abelian) groups as is shown in [13] and [21] (see also [3]).

Theorem 13. *If $S_1, \dots, S_n$ are non-empty subsets of a finite group G and $S = S_1 \star \dots \star S_n$ ($\star$ denoting the group operation), then either $|S| \geq p$ or*

$$|S| - 1 \geq \sum_i (|S_i| - 1).$$

where p is the smallest prime dividing $|G|$. □

Unfortunately, Theorem 13 does not seem to generalize to more general covers. For example, if $S_1 = S_2 = S_3 = \{0, 1, 3, 5\} \subseteq \mathbb{Z}_{13}$ then $|S_1 + S_2| = |S_1 + S_3| = |S_2 + S_3| = 9$ and $|S_1 + S_2 + S_3| = 12$, so

$$2(|S_1 + S_2 + S_3| - 1) < (|S_1 + S_2| - 1) + (|S_1 + S_3| - 1) + (|S_2 + S_3| - 1).$$

6 Conjectures

The most obvious problems related to the results above concern general (not necessarily commutative) groups. In fact, Gyarmati, Matolcsi and Ruzsa [10], Problem 1.3, have already asked whether a suitable analogue of the inequality corresponding to the Loomis–Whitney inequality holds for all groups. Even more, before this problem was posed, Ruzsa [20], Theorem 5.1, had solved its first non-trivial case. It is not unreasonable to hope that the analogue of the Box Theorem (or Cover Inequality) holds as well, as does the extension of Corollary 11. To state these conjectures, given finite non-empty sets $S_1, \dots, S_n$ in a group G with operation $\star$ as above, and a set $A \subset [n]$, write N_A for the maximal number of elements in a product set obtained from $S_1 \star \cdots \star S_n$ by replacing each S_i, $i \notin A$, by a single element of S_i. Similarly, write n_A for the corresponding minimum.

Conjecture 14. *Let $S_1, \dots, S_n$ be non-empty finite subsets of a group. Set $S = S_1 \star \cdots \star S_n$, and let N_A be as above. Then there are constants $\lambda_1, \dots, \lambda_n > 0$ such that*

$$|S| = \prod_1^n \lambda_i \quad \text{and} \quad N_A \geq \prod_{i \in A} \lambda_i \quad \text{for all} \ \ A \subseteq [n].$$

The case $n = 2$ follows from the obvious inequality $|S_1 \star S_2| \leq |S_1||S_2|$, while the $n = 3$ case reduces to the corresponding uniform cover inequality for the cover $\mathcal{A} = \{\{1,2\},\{1,3\},\{2,3\}\}$. This is the case that was proved by Ruzsa [20], Theorem 5.1. Here prove the following extension of this result.

Theorem 15. *Let $S_1, \dots, S_n$ be non-empty finite subsets of a group. Set $S = S_1 \star \cdots \star S_n$, and let N_A be defined as above. Suppose $\mathcal{A}$ is a uniform k-cover of $[n]$ and $<'$ is an ordering of $[n]$ such that for every $A \in \mathcal{A}$ and every $i \in A$ and $j \notin A$ with $\min A < j < \max A$, we have $j <' i$. Then*

$$|S|^k \leq \prod_{A \in \mathcal{A}} N_A.$$

Proof. Before we prove this result in general, we shall illustrate the proof by considering the case $\mathcal{A} = \{\{1,2\},\{1,3\},\{2,3\}\}$, which was proved by Ruzsa [20]. Write $S' = \varphi_{\{1,2,3\}}(S)$. Then just as in the abelian case, $|S'_{\{1,2\}}| \leq |\varphi_{\{1,2\}}(S_1 \star S_2)| = N_{\{1,2\}}$, and $|S'_{\{2,3\}}| \leq |\varphi_{\{2,3\}}(S_2 \star S_3)| = N_{\{2,3\}}$. However, $s'_1 \star s'_3 = s_1 \star s_3$ does not necessarily imply that $s'_1 \star s_2 \star s'_3 = s_1 \star s_2 \star s_3$, so we cannot conclude that if $(s_1, s_2, s_3) \in S'$ then $(s_1, s_3) \in \varphi_{\{1,3\}}(S_1 \star S_3)$. But if $(s_1, s_2, s_3) \in S'$, $s_i \in S_i$, then $(s_1, s_2, s_3) \in \varphi_{\{1,2,3\}}(S_1 \star \{s_2\} \star S_3)$, so

$$|S'_{\{1,3\}|\{2\}}| \leq \max_{s_2 \in S_2} |\varphi_{\{1,2,3\}}(S_1 \star \{s_2\} \star S_3)| = \max_{s_2 \in S_2} |S_1 \star \{s_2\} \star S_3| = N_{\{1,3\}}.$$

The result now follows from Theorem 2 by using the uniform 2-cover $\mathcal{A} = \{\{1,2\}, \{1,3\}, \{2,3\}\}$ and swapping indices 1 and 2 throughout so that

$$|S|^2 = |S'|^2 \leq |S'_{\{1,2\}}||S'_{\{2,3\}}||S'_{\{1,3\}|\{2\}}| \leq N_{\{1,2\}} N_{\{2,3\}} N_{\{1,3\}}.$$

For the general case, write $S' = \varphi_{[n]}(S)$. If $A \in \mathcal{A}$, let $A_{\text{gap}} = \{j \notin A : \min A < j < \max A\}$ be the set of elements in the 'gaps' of A. If $(s_1, \dots, s_n) \in S'$, then

$$(s_i)_{i \in A \cup A_{\text{gap}}} \in \varphi_{A \cup A_{\text{gap}}}(\tilde{S}_{\min A} \star \cdots \star \tilde{S}_{\max A}),$$

where $\tilde{S}_i = S_i$ if $i \in A$ and $\tilde{S}_i = \{s_i\}$ otherwise. Thus

$$\begin{aligned} |S'_{A|A_{\text{gap}}}| &\leq \max_{s_i \in S_i,\, i \in A_{\text{gap}}} |\varphi_{A \cup A_{\text{gap}}}(\tilde{S}_{\min A} \star \cdots \star \tilde{S}_{\max A})| \\ &\leq \max_{s_i \in S_i,\, i \in A_{\text{gap}}} |\tilde{S}_{\min A} \star \cdots \star \tilde{S}_{\max A}| = N_A \end{aligned}$$

The result now follows from Theorem 2 by first ordering the indices using $<'$ so that $A_{\text{gap}} \subseteq A_*$ for all $A \in \mathcal{A}$. Then $|S|^k = |S'|^k \leq \prod_{A \in \mathcal{A}} |S_{A|A_*}| \leq \prod_{A \in \mathcal{A}} |S_{A|A_{\text{gap}}}| \leq \prod_{A \in \mathcal{A}} N_A$. □

Theorem 15 applies to any cover, such as $\{\{1,2\}, \{2,3\}, \dots, \{n,1\}\}$, that has only one set A with a gap, as well as many others, for example the uniform 3-cover $\{\{1,2,3\}, \{2,3,4\}, \{3,4,5\}, \{1,2,4,5\}, \{1,5\}\}$. Unfortunately Theorem 15 does not apply to the cover $\{\{1,2,3\}, \{1,2,4\}, \{1,3,4\}, \{2,3,4\}\}$. Indeed, Theorem 2 does not seem to imply Conjecture 14 for $n \geq 4$.

Conjecture 16. *Let $S_1, \dots, S_n$ be non-empty finite subsets of a finite group, and let S and n_A be as above. Then there are constants σ_i such that for $|S| \leq cp$,*

$$|S| - 1 = \sum_{i=1}^{n} \sigma_i \quad \textit{and} \quad n_A - 1 \leq \sum_{i \in A} \sigma_i \textit{ for all } A \subseteq [n].$$

Here p is the smallest prime dividing $|G|$, and $c > 0$ is some absolute constant.

In conclusion, we should say that both these conjectures are rather tentative: we would not be amazed if they turned out to be false.

7 Acknowledgements

The initial results in Section 5 were proved after (and while) listening to I. Ruzsa's lecture in Tel Aviv in June, 2007; we are grateful to Professor Ruzsa for showing us his slides of this lecture, and for prepublication access to [10]. We are also grateful to András Frank and László Lovász who drew our attention to Lovász's use of compressions. Finally, we are grateful to Ehud Friedgut and Vsevolod Lev for pointing out errors in an earlier version of this paper.

References

[1] G.R. Allan, An inequality involving product measures, in *Radical Banach Algebras and Automatic Continuity* (J.M. Bachar et al., eds.), Lecture Notes in Mathematics **975**, Springer-Verlag, 1981, 277–279.

[2] N. Alon and M. Dubiner, Zero-sum sets of prescribed size, in *Combinatorics, Paul Erdős is eighty, Vol. 1,* pp. 33–50, Bolyai Soc. Math. Stud., János Bolyai Math. Soc., Budapest, 1993.

[3] P. Balister and J.P. Wheeler, The Erdős-Heilbronn problem for finite groups, *Acta Arithmetica*, **140** (2009), 105–118.

[4] B. Bollobás and A. Thomason, Projections of bodies and hereditary properties of hypergraphs, *Bull. London Math. Soc.* **27** (1995), 417–424.

[5] Yu.D. Burago and V.A. Zalgaller, *Geometric Inequalities*, Springer-Verlag, 1988, xiv+331pp.

[6] F.R.K. Chung, R.L. Graham, P. Frankl and J.B. Shearer, Some intersection theorems for ordered sets and graphs, *J. Combinatorial Theory A* **43** (1986), 23–37.

[7] H. Davenport, On the addition of residue classes, *J. London Math. Soc.* **10** (1935), 30–32.

[8] A. Frank, Edge-connection of graphs, digraphs, and hypergraphs, in *More Sets, Graphs and Numbers,* Bolyai Soc. Math. Stud. **15**, Springer, Berlin, 2006, pp. 93–141.

[9] K. Gyarmati, M. Matolcsi and I.Z. Ruzsa, Plünnecke's inequality for different summands, in *Building Bridges*, Bolyai Soc. Mathematical Studies **19**, ed. M. Grötschel, G. O. H. Katona, Springer-Bolyai 2008, 309–320.

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

[11] H. Hadwiger, *Vorlesungen über Inhalt, Oberfläche und Isoperimetrie*, Springer-Verlag, 1957, xiii+312pp.

[12] T.S. Han, Nonnegative entropy measures of multivariate symmetric correlations, *Information and Control* **36** (1978), 133–156.

[13] G. Károlyi, The Cauchy–Davenport theorem in group extensions, *L' Enseignement Mathématique* **51** (2005), 239–254.

[14] L.H. Loomis and H. Whitney, An inequality related to the isoperimetric inequality, *Bull. Amer. Math. Soc.* **55** (1949), 961–962.

[15] L. Lovász, Solution to Problem 11, see pp. 168–169 of 'Report on the 1968 Miklós Schweitzer Memorial Mathematical Competition' (in Hungarian), *Matematikai Lapok* **20** (1969), 145–171.

[16] L. Lovász, 2-matchings and 2-covers of hypergraphs, *Acta Math. Acad. Sci. Hungar.* **26** (1975), 433–444.

[17] L. Lovász, On two minimax theorems in graph, *J. Combinatorial Theory Ser. B* **21** (1976), 96–103.

[18] M. Madiman and P. Tetali, Sandwich bounds for joint entropy, *Proc. IEEE Intl Symp. Inform. Theory, Nice*, June, 2007.

[19] M. Madiman and P. Tetali, Information inequalities for joint distributions, with interpretations and applications, *IEEE Transactions on Information Theory* **56** (2010), 2699-2713.

[20] I.Z. Ruzsa, Cardinality questions about sumsets, in *Additive Combinatorics*, CRM Proc. Lecture Notes **43**, Amer. Math. Soc., Providence, RI, 2007, pp. 195–205.

[21] J.P. Wheeler, The Cauchy-Davenport theorem for finite groups, *preprint*, `http://www.msci.memphis.edu/preprint.html` (2006).