Information-theoretic Inequalities in Additive Combinatorics

(Invited Paper)

Mokshay Madiman
Department of Statistics
Yale University
24 Hillhouse Avenue
New Haven, CT 06511, USA
Email: mokshay.madiman@yale.edu

Adam W. Marcus
Department of Mathematics
Yale University
PO Box 208283
New Haven, CT 06520, USA
Email: adam.marcus@yale.edu

Prasad Tetali
School of Mathematics and
School of Computer Science
Georgia Institute of Technology
Atlanta, GA 30332-0160, USA
Email: tetali@math.gatech.edu

***Abstract*—We review recent developments that adopt an information-theoretic approach to the study of sumset inequalities in additive combinatorics.**

Keywords: Sumsets, additive combinatorics, entropy inequalities, set cardinality inequalities.

I. Introduction

Additive combinatorics (see, e.g., [1]) is the theory of additive structures in sets equipped with a group structure (and possibly other structure that interacts with the group structure). The prototypical example is the study of additive structure in the integers, as distinct from the multiplicative structure that underlies prime factorization and much of classical number theory. There have been several major developments and a lot of high-profile mathematical activity in additive combinatorics in recent years, with the most famous example being the celebrated Green-Tao theorem on the existence of arbitrarily long arithmetic progressions within the primes.

An important collection of tools in the study of additive combinatorics is a variety of sumset inequalities. Here, by "sumset" is meant sets such as $A + B = \{a + b : a \in A, b \in B\}$, where A, B are finite sets in some group G, and by "sumset inequality" is meant inequalities for the cardinalities of sumsets under a variety of conditions. Examples of sumset inequalities that have left a deep mark on the field include the Cauchy-Davenport inequality, Freiman's theorem, the Balog-Szemerédi-Gowers theorem, and the Plünnecke-Ruzsa inequalities.

A natural connection between entropy and set cardinalities arises from the fact that the entropy of the uniform distribution on a finite set A is just $\log |A|$, and this is the maximum entropy of any distribution supported on A. The idea that this observation may be useful for formulating entropy inequalities for sums that are analogous to sumset inequalities appears to have been first explored by Imre Ruzsa, and has in the last couple of years been fleshed out independently and in different directions by Ruzsa [2], the authors of this note [3] and Tao [4]. In particular, in [3], we both present new inequalities for entropies of sums of independent random variables, and prove a number of (some old and some new) sumset inequalities. Among the results we prove are two conjectures of Ruzsa– one partially and one fully. The goal of this note is to summarize the contributions of [3], with discussion of other related work when relevant.

Without further ado, let us state two illustrative results, one each for entropy and set cardinality (for definitions used in the statements, see the beginning of Section II).

Theorem 1. *Let $Z_1, \ldots, Z_n$ be independent discrete random variables taking values in the Abelian group $(\mathcal{G}, +)$, and let $\mathcal{C}$ be an r–regular hypergraph on $[n]$. Then*

$$H(Z_1 + \cdots + Z_n) \leq \frac{1}{r} \sum_{s \in \mathcal{C}} H\Big(\sum_{i \in s} Z_i \Big).$$

Theorem 2. *Let $X_1, \ldots, X_n$ be finite subsets of the Abelian group $(\mathcal{G}, +)$, and let $\mathcal{C}$ be an r–regular hypergraph on $[n]$. Then*

$$|X_1 + \cdots + X_n| \leq \prod_{s \in \mathcal{C}} \Big| \sum_{i \in s} X_i \Big|^{\frac{1}{r}}.$$

This set inequality (and others for sums in non-Abelian groups, projections, and polynomial functions) are obtained in [3] as a corollary of inequalities for cardinalities of compound sets. Here a compound set means a set obtained by varying each argument of a function of several variables over a set associated with that argument, where all the sets are subsets of an appropriate algebraic structure so that the function is well defined. In other words, for subsets $X_1, \ldots, X_k$ of some ambient space $\mathcal{X}$, consider

$$f(X_1, \ldots, X_k) = \{f(x_1, \ldots, x_k) : x_1 \in X_1, \ldots, x_k \in X_k\}.$$

When the ambient space is a group, the only operation available is the sum, and all compound sets are sumsets. When the ambient space is a ring, one may consider compound sets built from polynomials. For particular ambient spaces, such as Euclidean space, the class of functions available is extremely broad and therefore so is the class of compound sets that can be considered.

In proving cardinality inequalities for compound sets, [3] introduces a framework of "partition-determined functions",

and develop general results for that framework that yield the above results as corollaries. Since developing the necessary terminology and notation for partition-determined functions will take us too far afield for our present purposes, we only discuss in this note the specializations of our results to sums. The sumset inequalities we obtain in this fashion generalize some of the results of Gyarmati, Matolcsi, and Ruzsa [5], by combining an idea of "representative elements" used in [5] with entropy inequalities developed in [6]. Independently of our work, recent papers by by Balister and Bollobás [7] and Gyarmati, Matolcsi, and Ruzsa [8] have proved results that overlap with ours using different techniques.

Section II discusses entropy inequalities for sums of independent discrete random variables taking values in an Abelian group. Section III discusses analogous sumset inequalities. Interestingly, the proof of the latter do not directly follow from the former, and the reason for this is explained in Section IV. Section V describes inequalities for sumsets in non-Abelian groups, motivated by (and partially resolving) a conjecture of Ruzsa [9]. We conclude with some discussion in Section VI.

II. Entropy inequalities for sums in Abelian groups

Let $[n]$ be the index set $\{1, 2, \ldots, n\}$. Let $\mathcal{C}$ be a collection of subsets of $[n]$. For any index i in $[n]$, define the *degree* of i in $\mathcal{C}$ as $r(i) = |\{t \in \mathcal{C} : i \in t\}|$. A function $\alpha : \mathcal{C} \to \mathbb{R}_+$, is called a *fractional covering*, if for each $i \in [n]$, we have $\sum_{s \in \mathcal{C}: i \in s} \alpha_s \geq 1$. If α satisfies the equalities $\sum_{s \in \mathcal{C}: i \in s} \alpha_s = 1$ for each $i \in [n]$, it is called a *fractional partition*. If the degree of every index i in $\mathcal{C}$ is exactly r, $\mathcal{C}$ is called an r-regular hypergraph, and $\alpha_s = 1/r$ for every $s \in \mathcal{C}$ constitutes a fractional partition using $\mathcal{C}$.

Then a general result for partition-determined functions in [3] yields the following corollary for sums, recovering a result of [10].

Theorem 3. [Entropy of Sums in Abelian Groups] *Let $Z_1, \ldots, Z_n$ be independent discrete random variables taking values in the Abelian group $\mathcal{G}$, and let*

$$Z_s^+ = \sum_{i \in s} Z_i.$$

Then:

1) *The set function $f(s) = H(Z_s^+)$ is submodular.*
2) *For any fractional covering α using any collection $\mathcal{C}$ of subsets of $[n]$,*

$$H(Z_1 + \cdots + Z_n) \leq \sum_{s \in \mathcal{C}} \alpha_s H\big(Z_s^+\big).$$

Note that Theorem 1 is just the second part of Theorem 3 written for the special case of regular hypergraphs.

In addition, the first part of Theorem 3 resolves affirmatively (and in fact strengthens) "Entropy Conjecture 3" in the recent paper of Ruzsa [2]. Indeed, the latter conjecture stated that

$$H(X) + H(Y + Z) \leq H(X + Y) + H(X + Z),$$

whereas the submodularity assertion of Theorem 3 combined with the fact that entropy can only increase on summation implies that

$$\begin{aligned} H(X) + H(Y + Z) &\leq H(X) + H(X + Y + Z) \\ &\leq H(X + Y) + H(X + Z). \end{aligned}$$

III. Sumset inequalities for Abelian groups

Theorem 2 is a special case of the following inequality.

Theorem 4. *Let $X_1, \ldots, X_n$ be finite subsets of the Abelian group $(\mathcal{G}, +)$, and let $\mathcal{C}$ be an r–regular hypergraph on $[n]$. Then*

$$|X_1 + \cdots + X_n| \leq \prod_{s \in \mathcal{C}} |X_s^+|^{\alpha_s},$$

where

$$X_s^+ = \sum_{i \in s} X_i$$

is well defined by commutativity of addition.

This may be paraphrased as saying that the logarithms of sumset cardinalities are fractionally subadditive. It is classical (and recently reviewed in [6]) that fractional subadditivity is weaker than submodularity. Thus it is natural to ask if the logarithms of sumset cardinalities are submodular. However, this is not true in general, as observed by Ruzsa. Indeed, log-submodularity of sumset cardinality would imply that $|kA|$ is a log concave function of k, which is not the case. In fact, if $|A| = n$ and $|2A| = m$, then $|3A|$ can be anywhere between cm and $C \min(m^{3/2}, m^3/n^2)$.

By a slightly more elaborate deployment of the general framework used in [3] to prove Theorem 4, one can also obtain the following result.

Theorem 5. *Let $A, X_1, X_2, \ldots, X_n \subset \mathcal{G}$, where $(\mathcal{G}, +)$ is an Abelian group. Let α be any fractional partition on $[n]$ using the collection $\mathcal{C}$ of subsets of $[n]$. Then, for any $D \subseteq X_{[n]}^+$,*

$$|A + D|^c \leq |D|^{c-1} \prod_{s \in \mathcal{C}} |A + X_s^+|^{\alpha_s},$$

where $c = \sum_{s \in \mathcal{C}} \alpha_s$.

By applying Theorem 5 to an r-regular hypergraph $\mathcal{C}$, for which $\alpha_s = \frac{1}{r}$ gives a fractional partition, we obtain that for any $D \subseteq X_1 + \ldots + X_n$,

$$|A + D|^{|\mathcal{C}|} \leq |D|^{|\mathcal{C}| - r} \prod_{s \in \mathcal{C}} \big|A + X_s^+\big|. \tag{1}$$

Note that when $\mathcal{C} = \mathcal{C}_1$ is the collection of singleton sets, (1) reduces to Theorem 1.5 of Gyarmati, Matolcsi, and Ruzsa [5], namely

$$|A + D|^k \leq |D|^{k-1} \prod_{i=1}^{k} |A + X_i|.$$

When $\mathcal{C} = \mathcal{C}_{k-1}$ is the collection of leave-one-out sets, (1) resolves a conjecture stated in [5], namely if $\overline{X_i} = X_1 + \cdots + X_{i-1} + X_{i+1} + \cdots + X_k$ for $i = 1, \ldots, k$,

$$|A + D|^k \leq |D| \prod_{i=1}^{k} |A + \overline{X_i}|.$$

Some of these corollaries are also discussed by Balister and Bollobás [7], but our methods are independent of theirs.

IV. Remarks on the connection between the entropy and set inequalities

Suppose $X_1, \ldots, X_n$ are finite sets in some ambient Abelian group G. Let $Z_1, \ldots, Z_n$ be random variables supported by the sets $X_1, \ldots, X_n$ respectively, and note that for any subset $s \subset [n]$, the random variable Z_s^+ is supported on the compound set X_s^+. Thus the left hand side of Theorem 1 has the bound

$$H(Z_s^+) \leq \log |X_s^+|, \tag{2}$$

while its right hand side has the bound

$$\sum_{s \in \mathcal{C}} \alpha_s H\big(Z_s^+\big) \leq \sum_{s \in \mathcal{C}} \alpha_s \log \big|X_s^+\big|.$$

Theorem 2 says that these bounds themselves are ordered. This would be implied by Theorem 1 if we could find a product distribution on $(X_1 \times \ldots \times X_n)$ that made the sum uniformly distributed on its range, since then (2) would simply hold with equality.

Interestingly, while it is in general not possible to find such product distributions, it is always possible to find a *joint* distribution (with dependence) that makes the sum uniformly distributed on its range.

Lemma 6. *There exists a joint distribution for $(Z_1, \ldots, Z_k)$ on $X_1 \times \ldots \times X_k$ that makes $Z_{[k]}^+$ uniformly distributed on $X_{[k]}^+$.*

As a result, the simple proof of Theorem 2 from Theorem 1 as outlined above (that one may hope for) fails because of the independence requirement of Theorem 1. Our method of proof of Theorem 2 in [3] instead uses, following Gyarmati, Matolcsi and Ruzsa [5], the uniform distribution on a set of "representatives" combined with joint entropy inequalities to address this problem.

V. Sumset inequalities for non-Abelian groups

A particularly interesting thrust of recent research in additive combinatorics is to generalize many of the known results to non-Abelian groups. For instance, given $X_1, \ldots, X_k$ subsets of a non-Abelian group $(\mathcal{G}, \circ)$, can we find similar bounds on $|X_1 \circ \ldots \circ X_k|$ as we did when the underlying group was Abelian? To see that, in fact, the same bounds cannot hold, consider the following example.

Example 1. Let $\mathcal{G} = \{e, R, R^2, F, RF, R^2F\}$ be the dihedral group on 6 elements, $S = \{e, F\}, T = \{R\}$, and $U = \{e, F\}$. Then it is *not* the case that $|S \circ T \circ U|^2 \leq |S \circ T||T \circ U||S \circ U|$.

Proof: On one hand, we have that $S \circ T = \{R, FR\}, S \circ U = \{e, F\}$, and $T \circ U = \{R, RF\}$, and so $|S \circ T||T \circ U||S \circ U| = 8$. On the other hand, $S \circ T \circ U = \{R, FR, RF, R^2\}$ and so $|S \circ T \circ U|^2 = 16$. ∎

The underlying reason that non-Abelian groups cannot be bounded in such a way is that, as in the example above, $|S \circ U|$ need not have any relation to $|S \circ T \circ U|$. So any bound will need to find some way to link the two. To do so, we use in [3] as a key ingredient a new joint entropy inequality. The idea is to compare the entropy of a collection of random variables to the sum of the entropies of pairs of random variables conditioned on all of the random variables falling in between the pair.

Lemma 7. *Let $Z = Z_1, Z_2, \ldots, Z_k$ be random variables, and define*

$$Z_{(i,j)} = \{Z_t : i < t < j\}$$

for all $1 \leq i < j \leq k$. Then, for $k \geq 2$,

$$(k-1)H(Z_1, Z_2, \ldots, Z_k) \leq \sum_{i=1}^{k} \sum_{j>i}^{k} H(Z_i, Z_j \mid Z_{(i,j)})$$

Proof: We prove this by induction on k. The base case (when $k = 2$) is trivial, so assume the hypothesis to be true for $k-1$ random variables and consider a collection of k random variables. The idea is to peel off all of the pairs that contain the random variable Z_k and then appeal to the induction hypothesis for the other $\binom{k-1}{2}$ pairs.

Fix the usual notation for open and closed intervals (even though we are only concerned about integers) and the convention that the interval $[1, 0]$ is the empty set. Using Shannon's chain rule for entropy, one can write

$$\begin{aligned} H(Z_{[1,k]}) &= H(Z_{[1,i]} \mid Z_{(i,k]}) + H(Z_{i+1}, Z_k \mid Z_{(i+1,k)}) \\ &\quad + H(Z_{(i+1,k)}) \\ &\leq H(Z_{[1,i]} \mid Z_{(i,k)}) + H(Z_{i+1}, Z_k \mid Z_{(i+1,k)}) \\ &\quad + H(Z_{(i+1,k)}) \end{aligned}$$

for any $0 \leq i \leq k-2$, where the inequality is because conditioning decreases entropy. Summing all of these inequalities, we get by relabeling the last sum

$$\begin{aligned} (k-1)H(Z_{[1,k]}) &\leq \sum_{i=0}^{k-2} H(Z_{i+1}, Z_k \mid Z_{(i+1,k)}) + \\ &\quad \sum_{i=1}^{k-2} H(Z_{[1,i]} \mid Z_{(i,k)}) + \sum_{i=1}^{k-1} H(Z_{(i,k)}) \\ &= \sum_{i=0}^{k-2} H(Z_{i+1}, Z_k \mid Z_{(i+1,k)}) + \sum_{i=1}^{k-2} H(Z_{[1,k)}), \end{aligned}$$

where in the last step we used the fact that $(k-1, k)$ is the empty set, and the chain rule. Thus

$$\begin{aligned} (k-1)H(Z_{[1,k]}) &\leq \sum_{i=1}^{k-1} H(Z_i, Z_k \mid Z_{(i,k)}) \\ &\quad + (k-2)H(Z_{[1,k-1]}), \end{aligned}$$

and the rest of the inequality follows from the induction hypothesis. ∎

This proof is on similar lines to the simplified proof of Shearer's Lemma due to Llewellyn and Radhakrishnan (see [11]). Similar ideas have also been used to find extended Shearer-type bounds (the most general of these appearing in [6]). However, Lemma 7 seems to be new (and is a considerable strengthening of the similar-looking Han's inequality for pairs).

Using Lemma 7 and some other non-trivial constructions, we can give some sumset inequalities for non-Abelian groups.

Theorem 8. *Let $X_1, X_2, \ldots, X_k$ be subsets of a non-Abelian group, and define*

$$A(i,j) = \max\{|X_i \circ x_{i+1} \circ \ldots \circ x_{j-1} \circ X_j| : x_{i+1} \in X_{i+1}, \ldots, x_{j-1} \in X_{j-1}\}$$

for all $1 \leq i < j \leq k$. Then, for $k \geq 2$,

$$|X_1 \circ X_2 \circ \ldots \circ X_k|^{k-1} \leq \prod_{1 \leq i < j \leq k} A(i,j)$$

The following corollary, which inspired Theorem 8, was originally proved by Ruzsa [9].

Corollary 9. *Let S, T, U be subsets of a non-Abelian group. Then*

$$|S \circ T \circ U|^2 \leq \max_{t \in T} |S \circ T||T \circ U||S \circ t \circ U|\,.$$

Curiously, the method seems to break down in other cases. For example, the following problem posed in [9] remains open.

Problem 1. *Let S, T, U, V be subsets of a non-Abelian group. Is it true that*

$$|S \circ T \circ U \circ V|^3 \leq \max_{t,u} |S \circ T \circ U||S \circ T \circ u \circ V| \\ |S \circ t \circ U \circ V||T \circ U \circ V|?$$

Observe that the corresponding entropy inequality is *not* true; indeed, if one chooses $Z_2 = Z_3$ and $Z_1 = Z_4 = 0$, then

$$\begin{aligned} H(Z_2) &= H(Z_1, Z_2, Z_3, Z_4) \\ &> \frac{1}{3}\big[H(Z_1, Z_2, Z_3) + H(Z_2, Z_3, Z_4) + \\ &\quad H(Z_1, Z_3, Z_4 \mid Z_2) + H(Z_1, Z_2, Z_4 \mid Z_3)\big] \\ &= \frac{2}{3} H(Z_2). \end{aligned}$$

VI. Discussion

Entropies of sums of random variables, even in the setting of independent summands, are not as well understood as joint entropies. In this note, we have reviewed some new developments in the understanding of the entropy of sums, for discrete random variables taking values in general groups. From the point of view of additive combinatorics, these inequalities are not only interesting in their own right, but the techniques they call on can be used to prove several interesting sumset inequalities. Although few proofs have been given in this note, they can be found in [3] and are rather elementary for anyone even slightly familiar with information theory.

While this note has focused on discrete random variables and finite sets, it is possible at least partially to develop analogues for continuous random variables and measurable sets (in the context of finite-dimensional linear spaces). However, these analogues cannot quite be of the same form because of the non-trivial differences between discrete and differential entropy. Some valid analogues of the results in this note are developed in [10], [12]; these have recently found interesting applications in convex geometry [13]. In a different direction, entropy power inequalities provide other ways to quantify using entropy the behavior of sums– the most general such inequalities known for sums of independent random vectors may be found in [14], [15].

Acknowledgments

We thank Imre Ruzsa for sharing the preprint [5], for helpful discussions, and for informing us of the independent and recent work of Balister–Bollobás. We also thank Béla Bollobás for promptly sending us the preprint [7], which contains results of independent interest.

References

[1] T. Tao and V. Vu, *Additive combinatorics*, ser. Cambridge Studies in Advanced Mathematics. Cambridge: Cambridge University Press, 2006, vol. 105.

[2] I. Z. Ruzsa, "Entropy and sumsets," *Random Struct. Alg.*, vol. 34, pp. 1–10, 2009.

[3] M. Madiman, A. Marcus, and P. Tetali, "Entropy and set cardinality inequalities for partition-determined functions, with applications to sumsets," *Preprint*, 2008. [Online]. Available: http://arxiv.org/abs/0901.0055

[4] T. Tao, "Sumset and inverse sumset theory for Shannon entropy," *Preprint*, 2009. [Online]. Available: arXiv:0906.4387v1

[5] K. Gyarmati, M. Matolcsi, and I. Ruzsa, "A superadditivity and submultiplicativity property for cardinalities of sumsets," *Preprint*, June 2007.

[6] M. Madiman and P. Tetali, "Information inequalities for joint distributions, with interpretations and applications," *Submitted*, 2007.

[7] P. Balister and B. Bollobás, "Projections, entropy, and sumsets," *Preprint*, October 2007.

[8] K. Gyarmati, M. Matolcsi, and I. Z. Ruzsa, "Plünnecke's inequality for different summands," *Preprint,* `arXiv:0810:1488v1`, 2008.

[9] I. Z. Ruzsa, "Cardinality questions about sumsets," in *Additive combinatorics*, ser. CRM Proc. Lecture Notes. Providence, RI: Amer. Math. Soc., 2007, vol. 43, pp. 195–205.

[10] M. Madiman, "On the entropy of sums," in *Proc. IEEE Inform. Theory Workshop*, Porto, Portugal, 2008.

[11] J. Radhakrishnan, "Entropy and counting," in *IIT Kharagpur Golden Jubilee Volume*, 2001. [Online]. Available: http://www.tcs.tifr.res.in/~jaikumar/mypage.html

[12] M. Madiman, "Determinant and trace inequalities for sums of positive-definite matrices," *Preprint*, 2008.

[13] S. G. Bobkov and M. Madiman, "Reverse Brunn-Minkowski and reverse entropy power inequalities for convex measures," *Preprint*, 2009.

[14] M. Madiman and A. Barron, "Generalized entropy power inequalities and monotonicity properties of information," *IEEE Trans. Inform. Theory*, vol. 53, no. 7, pp. 2317–2329, July 2007.

[15] M. Madiman and F. Ghassemi, "The entropy power of sums is fractionally superadditive," in *Proc. IEEE Intl. Symp. Inform. Theory*, Seoul, Korea, 2009.