Entropy versions of additive inequalities

Alberto Espuny Díaz [1,3]

School of Mathematics
University of Birmingham
Birmingham, United Kingdom

Oriol Serra [2,4]

Department of Mathematics
Universitat Politècnica de Catalunya
Barcelona, Spain

Abstract

The connection between inequalities in additive combinatorics and analogous versions in terms of the entropy of random variables has been extensively explored over the past few years. This paper extends a device introduced by Ruzsa in his seminal work introducing this correspondence. This extension provides a toolbox for establishing the equivalence between sumset inequalities and their entropic versions. It supplies simpler proofs of known results and opens a path for obtaining new ones. Some new examples in nonabelian groups and with nonlinear functions of random variables illustrate the power of the device.

Keywords: Additive Combinatorics, Shannon entropy, additive inequalities.

[1] Partially supported by the Spanish MECD grant 2015/COLAB/00069.

[2] Supported by the Spanish MECD projects MTM2014-54745-P and MDM-2014-0445.

[3] Email: `axe673@bham.ac.uk`

[4] Email: `oriol.serra@upc.edu`

1 Introduction

In recent years, several authors realized that there exist certain analogies between many of the cardinality inequalities in additive combinatorics that have been developed over the years and some entropy inequalities. These analogies appear, for instance, with many important sumset inequalities such as the Plünnecke-Ruzsa inequalities, or with traditional entropy results such as Shearer's inequality. In the past decade, several papers exploring these analogies have appeared and many insightful results have been produced. The seminal work of Ruzsa [8] on this topic was extended by Balister and Bollobás [1], Kontoyiannis and Madiman [3], Madiman, Marcus and Tetali [4], Madiman and Tetali [5] or Tao [10], among many others. All these papers present different techniques with which the analogy between set cardinality inequalities and entropy inequalities can be studied. These techniques are used to obtain many new results, especially in the form of entropy inequalities.

Let X be a discrete random variable taking values $x_1, x_2, \dots, x_n$ with probabilities $p_1, p_2, \dots, p_n$, respectively. The Shannon entropy of X is defined as

$$\mathbf{H}(X) \coloneqq \sum_{i=1}^{n} p_i \log \frac{1}{p_i}.$$

This is a concave function, and Jensen's inequality gives

$$\mathbf{H}(X) \leq \log n,$$

where n is the cardinality of the range of X. Moreover, equality holds if and only if X is uniformly distributed over its range. This is the key property which allows one to translate entropy inequalities into combinatorial ones. From this perspective, entropy inequalities can be seen as generalizations of their combinatorial counterparts. One of the first examples in the literature is the translation of the classical inequality of Han,

$$(n-1)\mathbf{H}(X_1, \dots, X_n) \leq \sum_{i=1}^{n} \mathbf{H}(X_1, \dots, X_{i-1}, X_{i+1}, \dots, X_n),$$

which provides a simple direct proof of the inequality of Loomis and Whitney

$$|A|^{n-1} \leq \prod_{i=1}^{n} |A_i|,$$

where $A \subset E_1 \times \cdots \times E_n$ and A_i denotes the projection of A to the i-th coordinate hyperplane. This example opened the path to obtaining combinatorial inequalities from entropy ones, and most of the work in the area was devoted to proving entropy versions inspired by combinatorial ones.

In this sense, Ruzsa [8] introduced a device to walk the path backwards and obtain entropy inequalities from combinatorial ones, by establishing in fact the equivalence between the two versions. In his paper he restricted the device to linear functions in abelian groups.

The main goal of this paper is to extend the device of Ruzsa to arbitrary functions. By doing so we obtain a more flexible tool which allows us to give new combinatorial proofs of entropy inequalities and also to obtain new ones, particularly in the setting of nonabelian groups and also involving nonlinear functions.

Given any function $f\colon \mathcal{X} \to \mathcal{Y}$, we denote by f^k the function $f^k\colon \mathcal{X}^k \to \mathcal{Y}^k$ induced on the k-fold cartesian power $\mathcal{X}^k$, namely,

$$f^k(x_1, \ldots, x_k) = (f(x_1), \ldots, f(x_k))$$

for $x_1, \ldots, x_k \in \mathcal{X}$. The main result of the paper is the following.

Lemma 1.1 *Let $f, f_1, \ldots, f_n$ be functions defined over a set $\mathcal{X}$. Let $\alpha_1, \ldots, \alpha_n$ be real numbers. If for all positive k and every finite set $A \subseteq \mathcal{X}^k$, we have that*

$$|f^k(A)| \leq \prod_{i=1}^{n} \left|f_i^k(A)\right|^{\alpha_i},$$

then, for every discrete random variable X taking values in $\mathcal{X}$, the entropy of $f(X)$ satisfies that

$$\mathbf{H}(f(X)) \leq \sum_{i=1}^{n} \alpha_i \mathbf{H}(f_i(X))$$

whenever $\mathbf{H}(f_i(X))$ is finite for every $i \in [n]$.

The proof of Lemma 1.1 is based mainly in the following idea. The goal is to construct sets that “represent” random variables in some way. To this end, given an $\mathcal{X}$-random variable X, consider a vector of suitable length whose entries are possible outcomes of X, and such that they are repeated in such a way that we recover the distribution of X by choosing an entry of the vector uniformly at random. Then, consider the set of all such vectors. This set is what we call the *Ruzsa set* of X. One can show that the cardinality of the Ruzsa set is asymptotically determined by $\mathbf{H}(X)$ as the length of the vectors tends to infinity. Then, the main part of the proof consists in showing that, given a function f, the Ruzsa set of $f(X)$ is the same as the result of applying f^k to the Ruzsa set of X. This, together with the former, means that one can derive entropic inequalities by applying cardinality inequalities to Ruzsa sets and letting the length of the vectors go to infinity.

Lemma 1.1 is complemented by the following partial converse, which can be obtained from the concavity of the entropy function.

Lemma 1.2 *Let $f, f_1, \dots, f_n$ be any functions defined over a set $\mathcal{X}$. Let $\alpha_1, \dots, \alpha_n$ be positive real numbers. If the inequality*

$$\mathbf{H}(f(X)) \leq \sum_{i=1}^{n} \alpha_i \mathbf{H}(f_i(X))$$

holds for every random variable X with suport in a finite set $A \subseteq \mathcal{X}$, then we have that

$$|f(A)| \leq \prod_{i=1}^{n} |f_i(A)|^{\alpha_i}.$$

2 Applications

As was already mentioned, Lemma 1.1 can be used as a black box to prove many of the entropy inequalities which are usually obtained directly from entropy properties, so displaying their combinatorial nature. The following is an important example.

Theorem 2.1 *Let G be a (not necessarily commutative) group. The two following statements hold and are equivalent:*

(i) *For any A, B and C finite non-empty sets in G,*

$$|B - C| \leq \frac{|B - A||A - C|}{|A|}.$$

(ii) *For any X, Y and Z discrete random variables taking values in G such that X is independent from (Y, Z) and the entropies of X, $Y - X$ and $X - Z$ are finite, the entropy of $Y - Z$ is also finite and satisfies*

$$\mathbf{H}(Y - Z) \leq \mathbf{H}(Y - X) + \mathbf{H}(X - Z) - \mathbf{H}(X).$$

Both statements in Theorem 2.1 are well-known. The first one is Ruzsa's triangle inequality (see e.g. [7]) and the second one is its analogous entropy version, proved by Tao [10] and Ruzsa [8] independently by using entropy methods. The equivalence between the two is provided by the combination of Lemma 1.1 and Lemma 1.2.

The second application considers random variables in non-commutative groups. Petridis [6], building on previous work by Tao [9], gave the first non-commutative version of the Plünnecke-Ruzsa inequalities. These inequalities provide upper bounds on the cardinality of iterated sumsets in terms of 2-fold sumsets and have been extensively used in applications. The inequality of Petridis we consider reads as follows. Let A be a finite set in a non-commutative group. Assume that $|A^2| \leq \alpha|A|$ and $|A^3| \leq \beta|A|$. Then, for any $n \geq 3$ and any choice of $\epsilon_i \in \{1, -1\}$, $i \in [n]$,

$$|A^{\epsilon_1} A^{\epsilon_2} A^{\epsilon_3} \cdots A^{\epsilon_n}| \leq \alpha^{n-1}\beta^{n-2}|A|.$$

The next result is an entropy analog for random variables taking values in a nonabelian group. Inequalities in this setting are scarce in the literature.

Theorem 2.2 *Let G be a group and $n \geq 3$. Let X be a discrete G-random variable. Let X_i, $i \in [n]$, be independent copies of X. Assume that the entropies of X, X_1X_2 and $X_1X_2X_3$ are all finite. Then, for every choice of $\epsilon_i \in \{1, -1\}$, the entropy of $X_1^{\epsilon_1} X_2^{\epsilon_2} \cdots X_n^{\epsilon_n}$ is finite and satisfies*

$$\mathbf{H}(X_1^{\epsilon_1} X_2^{\epsilon_2} X_3^{\epsilon_3} \cdots X_n^{\epsilon_n}) \leq (n-1)\mathbf{H}(X_1X_2) + (n-2)\mathbf{H}(X_1X_2X_3) - 2(n-2)\mathbf{H}(X).$$

As before, Theorem 2.2 can be proved using Lemma 1.1. However, in this case Lemma 1.2 cannot be applied, so the equivalence between the statements has not been proved.

3 Final remarks

The results shown in Section 2 are only some examples of the versatility of Lemma 1.1. The results that can be obtained from it are very varied. It is noteworthy that most of the entropy inequalities obtained by Kontoyiannis and Madiman [3], Madiman, Marcus and Tetali [4], Madiman and Tetali [5] or Tao [10], which use sometimes sophisticated entropy arguments, are obtained through Lemma 1.1 and their combinatorial counterparts in a straightforward, unified way. The examples in Section 2 illustrate how the device presented in this paper can be used to extend the entropy analogs of combinatorial inequalities in the nonabelian setting or involving non linear functions. Additional examples and complete proofs can be found in [2].

References

[1] Balister, P. and B. Bollobás, *Projections, entropy and sumsets*, Combinatorica **32** (2012), pp. 125–141.

[2] Espuny Díaz, A. and O. Serra, *Entropy versions of additive inequalities*, Manuscript (2018).

[3] Kontoyiannis, I. and M. Madiman, *Sumset and inverse sumset inequalities for differential entropy and mutual information*, IEEE Trans. Inform. Theory **60** (2014), pp. 4503–4514.

[4] Madiman, M., A. W. Marcus and P. Tetali, *Entropy and set cardinality inequalities for partition-determined functions*, Random Structures Algorithms **40** (2012), pp. 399–424.

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

[6] Petridis, G., *New proofs of Plünnecke-type estimates for product sets in groups*, Combinatorica **32** (2012), pp. 721–733.

[7] Ruzsa, I. Z., *Sums of finite sets*, in: *Number theory (New York, 1991–1995)*, Springer, New York, 1996 pp. 281–293.

[8] Ruzsa, I. Z., *Sumsets and entropy*, Random Structures Algorithms **34** (2009), pp. 1–10.

[9] Tao, T., *Product set estimates for non-commutative groups*, Combinatorica **28** (2008), pp. 547–594.

[10] Tao, T., *Sumset and inverse sumset theory for Shannon entropy*, Combin. Probab. Comput. **19** (2010), pp. 603–639.