Entropy and additive combinatorics

Marcel K. Goh

Department of Mathematics and Statistics, McGill University

Abstract. These expository notes give a gentle introduction to the notion of entropy as it is used in additive combinatorics, moving at a leisurely pace through the entropic analogues of Plünnecke's theorem and the Balog–Szemerédi–Gowers theorem before tackling the recent proof of the polynomial Freiman–Ruzsa conjecture by W. T. Gowers, B. Green, F. Manners, and T. Tao. Effort has been put into making this document as self-contained as possible, and extra proof details have been supplied in the hope that these notes may be accessible to the average graduate student or enterprising undergraduate.

1. The Khintchine–Shannon axioms

Let X be a discrete random variable. Its entropy $\mathbf{H}\{X\}$ is a real number (or ∞) that measures the "information content" of X. For example, if X is a constant random variable, then $\mathbf{H}\{X\}$ should be zero (we do not gain any information from knowing the value of X), and if X is uniformly distributed on $\{0,1\}^n$, then $\mathbf{H}\{X\}$ should be proportional to n, since X is determined by n bits of information. It satisfies the following axioms, which are sometimes called the Khinchine–Shannon axioms.

a) (*Invariance.*) If X takes values in A, Y takes values in B, $\phi : A \to B$ is a bijection, and $\mathbf{P}\{Y = \phi(a)\} = \mathbf{P}\{X = a\}$ for all $a \in A$, then $\mathbf{H}\{X\} = \mathbf{H}\{Y\}$.

b) (*Extensibility.*) If X takes values in A and Y takes values in B for a set B such that $A \subseteq B$, and furthermore $\mathbf{P}\{Y = a\} = \mathbf{P}\{X = a\}$ for all $a \in A$, then $\mathbf{H}\{X\} = \mathbf{H}\{Y\}$.

c) (*Continuity.*) The quantity $\mathbf{H}\{X\}$ depends continuously on the probabilities $\mathbf{P}\{X = a\}$.

d) (*Maximisation.*) Over all possible random variables X taking values in a finite set A, the quantity $\mathbf{H}\{X\}$ is maximised for the uniform distribution.

e) (*Additivity.*) For X taking values in A and Y taking values in B, we have the formula

$$\mathbf{H}\{X, Y\} = \mathbf{H}\{X\} + \mathbf{H}\{Y \mid X\},$$

where $\mathbf{H}\{X, Y\} = \mathbf{H}\big\{(X, Y)\big\}$ and

$$\mathbf{H}\{Y \mid X\} = \sum_{x \in A} \mathbf{P}\{X = x\}\, \mathbf{H}\{Y \mid X = x\}.$$

The continuity axiom presupposes a topology on the set of all distributions on the target space of X. This topology will only be important here and there, so we will not give a complete description of it now, but simply introduce the relevant properties as needed later on.

We shall take it on faith that there really exists a function on random variables satisfying these axioms. (It is very possible you have met the formula for entropy somewhere on your travels, but you will not find it anywhere in these notes.) In fact, the axioms only define entropy up to a multiplicative constant, so we shall add the following axiom.

f) (*Normalisation.*) If X is uniformly distributed on $\{0,1\}$, then $\mathbf{H}\{X\} = 1$.

Notationally, we would expect that $\mathbf{H}\{Y \mid X\} = \mathbf{H}\{Y\}$ if X and Y are independent. This is the first proposition we will carefully prove, using only the axioms. The axiomatic description above, as well as many of the proofs in this section, are transcribed (with some adaptations here and there) from lectures given by W. T. Gowers.

Proposition 1.1. *Let X and Y be independent random variables. Then $\mathbf{H}\{Y \mid X\} = \mathbf{H}\{Y\}$ and consequently $\mathbf{H}\{X,Y\} = \mathbf{H}\{X\} + \mathbf{H}\{Y\}$.*

Proof. Suppose X takes values in a finite set A. Then for all $x \in A$, the distribution of Y and Y given that $X = x$ is the same, so

$$\mathbf{H}\{Y \mid X\} = \sum_{x \in A} \mathbf{P}\{X = x\}\, \mathbf{H}\{Y \mid X = x\} = \sum_{x \in A} \mathbf{P}\{X = x\}\, \mathbf{H}\{Y\} = \mathbf{H}\{Y\}.$$

The second version of the statement follows from the additivity axiom. ∎

We will sometimes use the notation X^n to denote the vector $(X_1, \dots, X_n)$ where the X_i are independent copies of the random variable X. We have the following three corollaries, which are each proved by induction. The second also requires the normalisation axiom, and the third is often known as the *chain rule*.

Corollary 1.2. *We have $\mathbf{H}\{X^n\} = n\,\mathbf{H}\{X\}$.* ∎

Corollary 1.3. *If X is uniformly distributed on a set of size 2^n, then*

$$\mathbf{H}\{X\} = n. \quad ∎$$

Corollary 1.4 (*Chain rule*). *Let $X_1, \dots, X_n$ be random variables. Then*

$$\mathbf{H}\{X_1, \dots, X_n\} = \mathbf{H}\{X_1\} + \mathbf{H}\{X_2 \mid X_1\} + \cdots + \mathbf{H}\{X_n \mid X_1, \dots, X_{n-1}\}. \quad ∎$$

Next we establish the intuitive statement that the entropy of a uniform random variable supported on a set A is at most the entropy of a uniform random variable supported on a superset B of A.

Proposition 1.5. *Let $A \subseteq B$ with B finite, let X be uniformly distributed on A, and let Y be uniformly distributed on B. Then $\mathbf{H}\{X\} \leq \mathbf{H}\{Y\}$, with equality if and only if $A = B$.*

Proof. By the extensibility axiom, $\mathbf{H}\{X\}$ is not affected if we regard X as a function taking values in B. Then by the maximisation axiom, $\mathbf{H}\{X\} \leq \mathbf{H}\{Y\}$, since Y is uniform on B.

If $A = B$, then it is clear that $\mathbf{H}\{X\} = \mathbf{H}\{Y\}$, since X and Y are the same random variable.

On the other hand, say $|A| = m$ and $|B| = n$ with $m < n$. If $m = 1$, then by the previous proposition we have $\mathbf{H}\{X\} = 0$, and by normalisation and invariance, $\mathbf{H}\{Y\} = 1$. When $m \geq 2$, pick k such that $m^k \leq n^{k-1}$, so that $|A^k| \leq |B^{k-1}|$. Then by Corollary 1.2 and the inequality we showed in the first paragraph of this proof, we have

$$n\,\mathbf{H}\{X\} = \mathbf{H}\{X^n\} \leq \mathbf{H}\{X^{n-1}\} = (n-1)\,\mathbf{H}\{Y\},$$

whence $\mathbf{H}\{X\} < \mathbf{H}\{Y\}$. ∎

The thread connecting entropy and additive combinatorics is rather a precarious one. As noted by I. Ruzsa [4], there are scenarios in which combinatorial inequalities and their corresponding entropy ones are equivalent, scenarios in which they both hold but their equivalence cannot be established, and further scenarios where an inequality holds in one world but cannot (or has not) been proven in the other.

The "dictionary" that allows us to translate statements between cardinality inequalities and entropy inequalities is based, in part, on the following observation. (In this and the rest of the notes, lg denotes the base-2 logarithm. This base can be changed by modifying the normalisation axiom.)

Proposition 1.6. *Let X be a uniform random variable on a finite set A. Then*

$$\mathbf{H}\{X\} = \lg |A|.$$

Proof. For any positive integer n we can let X^n denote a tuple of independent copies of X; Corollary 1.2 tells us $\mathbf{H}\{X^n\} = n\,\mathbf{H}\{X\}$. Let m be such that $2^m \leq |A|^n \leq 2^{m+1}$ so that

$$\frac{m}{n} \leq \lg |A| \leq \frac{(m+1)}{n}.$$

Let Y be uniform on a set of size 2^m, and let Z be uniform on a set of size 2^{m+1}. Then by Corollary 1.3 we have $\mathbf{H}\{Y\} = m$ and $\mathbf{H}\{Z\} = (m+1)$. Then by Proposition 1.5 we have

$$\frac{m}{n} \leq \mathbf{H}\{X\} \leq \frac{(m+1)}{n}.$$

In other words, $\mathbf{H}\{X\}$ satisfies the same bounds as $\lg |A|$. Taking n large, we can make these bounds arbitrarily tight, proving the claim. ∎

The maximisation axiom gives the following corollary.

Corollary 1.7. *Let X be a random variable supported on a finite set A. Then*

$$\mathbf{H}\{X\} \leq \lg |A|. \quad \blacksquare$$

Hence the entropy $\mathbf{H}\{X\}$ is at most the exponential of the size of its support. As we will see, simply replacing (logarithms of) cardinalities with entropies, we at useful "entropic analogues" of combinatorial statements. But first, we need more lemmas.

If Y is a random variable such that $Y = f(X)$ for some random variable Y and some function f, then we say that Y is *determined by* X or X *determines* Y. We want to show that $\mathbf{H}\{Y\} \leq \mathbf{H}\{X\}$, which reflects the idea that we get more information from X than from Y. This, rather annoyingly, seems to require a couple of steps.

Lemma 1.8. *If $Y = f(X)$ then $\mathbf{H}\{X\} = \mathbf{H}\{Y\} + \mathbf{H}\{X \mid Y\}$.*

Proof. There is a a bijection between values x taken by X and values $\big(x, f(x)\big)$ taken by (X, Y), so we have

$$\mathbf{H}\{X\} = \mathbf{H}\{X, Y\} = \mathbf{H}\{Y\} + \mathbf{H}\{X \mid Y\}$$

by invariance and additivity. $\blacksquare$

We are now done if we can show that entropy is nonnegative. This is a corollary of the following lemma, whose proof is a modification of one due to S. Eberhard.

Proposition 1.9. *Let X be a discrete random variable supported on a set A and let*

$$a^* = \arg\max_{a \in A} \mathbf{P}\{X = a\}.$$

Then

$$\mathbf{P}\{X = a^*\} \geq 2^{-\mathbf{H}\{X\}}.$$

Proof. First we will work in the case where there exists n such that $\mathbf{P}\{X = a\}$ is a multiple of $1/n$ for all $a \in A$. Let Y be uniformly distributed on $[n]$ and let $\{E_a\}_{a \in A}$ be a partition of $[n]$ such that $|E_a| = n\,\mathbf{P}\{X = a\}$ for all $a \in A$, and let $Z = a$ if $Y \in E_a$. This definition makes Z and X identically distributed, so $\mathbf{H}\{Z\} = \mathbf{H}\{X\}$ by the invariance axiom.

For every $a \in A$, the conditional entropy $\mathbf{H}\{Y \mid Z = a\}$ is uniformly distributed on a set of size $|E_a|$. From our choice of a^* we have $|E_{a^*}| \geq |E_a|$ for all $a \in A$. Hence by Proposition 1.6, we have

$$\mathbf{H}\{Y \mid Z\} = \sum_{a \in A} \mathbf{P}\{X = a\}\, \mathbf{H}\{Y \mid X = a\} = \sum_{a \in A} \mathbf{P}\{X = a\} \log |E_a| \leq \log |E_{a^*}|.$$

Since Z is determined by Y, we have $\mathbf{H}\{Y\} = \mathbf{H}\{Z\} + \mathbf{H}\{Y \mid Z\}$ by the previous lemma. So by another invocation of Proposition 1.6, we have

$$\begin{aligned}\mathbf{H}\{Z\} &= \mathbf{H}\{Y\} - \mathbf{H}\{Y \mid Z\} \\ &\geq \log n - \log |E_{a^*}| \\ &\geq \log\left(\frac{n}{|E_{a^*}|}\right) \\ &= \log\left(\frac{1}{\mathbf{P}\{X = a^*\}}\right),\end{aligned}$$

and hence $2^{-\mathbf{H}\{X\}} = 2^{-\mathbf{H}\{Z\}} \leq \mathbf{P}\{X = a^*\}$.

The general case follows from the continuity axiom. ∎

This proof came dangerously close to deriving the formula for entropy, but we will not need any such formula, so we will refrain from mentioning it. From the fact that $\mathbf{P}\{X = a^*\} \leq 1$ we can immediately conclude that entropy is nonnegative.

Corollary 1.10. *Let X be a discrete random variable taking values in a finite set A. Then $\mathbf{H}\{X\} \geq 0$.* ∎

This observation completes the proof that a random variable has a smaller entropy than one by which it is determined.

Corollary 1.11. *If $Y = f(X)$ then $\mathbf{H}\{X\} \geq \mathbf{H}\{Y\}$.* ∎

Next we show that a random variable has zero entropy if and only if it is constant. This reflects the idea that the variables from which we get no information are those which take the same value no matter what.

Proposition 1.12. *Let X be a discrete random variable. Then $\mathbf{H}\{X\} = 0$ if and only if it takes exactly one value.*

Proof. First suppose that X takes only one value. Let a be the value of X such that $\mathbf{P}\{X = a\} = 1$. Then (X, X) equals (a, a) with probability 1 as well, so $\mathbf{H}\{X\} = \mathbf{H}\{X, X\}$ by the invariance axiom. But it can easily be checked that X and (X, X) are independent (we have

$$\mathbf{P}\big\{X = a, (X, X) = (a, a)\big\} = \mathbf{P}\{X = a\}\,\mathbf{P}\big\{(X, X) = (a, a)\big\}$$

for instance), so $\mathbf{H}\{X, X\} = 2\,\mathbf{H}\{X\}$. Thus we conclude that $\mathbf{H}\{X\} = 0$.

Now suppose that X takes more than one value; let A be the set of a such that $\mathbf{P}\{X = a\} > 0$ and let $\alpha = \max_{a \in A} \mathbf{P}\{X = a\}$. For all n let X^n denote the tuple of n independent copies of X; the maximum probability of any particular value (in A^n) that X^n takes is α^n. But $\alpha < 1$ since X takes more than one value, so for any $\epsilon > 0$ we can find n such that $\alpha^n < \epsilon$. This means that we can partition A^n into two disjoint sets E and F such that $\mathbf{P}\{X^n \in E\}$ and $\mathbf{P}\{X^n \in F\}$ are both in the range $[1/2 - \epsilon, 1/2 + \epsilon]$.

Let Y be the random variable taking the value 0 if $X^n \in E$ and 1 if $X^n \in F$. Then by Corollary 1.2, $\mathbf{H}\{X^n\} = n\,\mathbf{H}\{X\}$, and since X^n determines Y,

$$\mathbf{H}\{X^n\} = \mathbf{H}\{Y\} + \mathbf{H}\{X^n \mid Y\} \geq \mathbf{H}\{Y\}.$$

But $\mathbf{H}\{Y\} > 0$ for ϵ small enough, the normalisation and continuity axioms. So $\mathbf{H}\{X\} \geq \mathbf{H}\{Y\}/n > 0$ as well. ∎

Mutual information. For random variables X and Y, the *mutual information* $\mathbf{I}\{X : Y\}$ is defined by the equivalent formulas

$$\begin{aligned}\mathbf{I}\{X : Y\} &= \mathbf{H}\{X\} + \mathbf{H}\{Y\} - \mathbf{H}\{X, Y\} \\ &= \mathbf{H}\{X\} - \mathbf{H}\{X \mid Y\} \\ &= \mathbf{H}\{Y\} - \mathbf{H}\{Y \mid X\}.\end{aligned}$$

It measures, roughly speaking, how much information one can get from one variable by looking at the other one. From the formula it is clear that $\mathbf{I}\{X : Y\} = 0$ if X and Y are independent. In general, we still have the inequality $\mathbf{I}\{X : Y\} \geq 0$, which a corollary of the following lemma, which expresses the intuitive fact that conditioning cannot increase the entropy of a random variable. The proof, which is due to C. West, uses an argument similar to the one we used to prove Proposition 1.9.

Proposition 1.13. *Let X and Y be discrete random variables. Then*

$$\mathbf{H}\{X \mid Y\} \leq \mathbf{H}\{X\}.$$

Proof. Let A be the support of X and B be the support of Y. First we consider the case that X is uniform on A (so A is finite). Then by the definition of conditional entropy,

$$\mathbf{H}\{X \mid Y\} = \sum_{b \in B} \mathbf{P}\{Y = b\}\,\mathbf{H}\{X \mid Y = b\}.$$

But for each b, the random variable $(X|Y = b)$ takes values in A, so its entropy is bounded above by $\mathbf{H}\{X\}$ by the maximisation axiom. Hence $\mathbf{H}\{X|Y\} \leq \mathbf{H}\{X\}$.

Next, suppose that A and B are both finite and suppose further that $\mathbf{P}\{Y = b\}$ is rational for all b. Then there is an integer n and integers $\{m_b\}_{b \in B}$ such that $\mathbf{P}\{Y = b\} = m_b/n$ for all $b \in B$. Now partition $[n]$ into sets $\{E_b\}_{b \in B}$, where $|E_b| = m_b$ for all $b \in B$. We define a random variable Z by sampling uniformly at random from E_b if $Y = b$, and doing so independently of $(X \mid Y = b)$. The result is a random variable Z that is uniform on $[n]$ and which is independent of $X \mid Y$. Furthermore, since Z determines Y, we have $\mathbf{H}\{Z\} = \mathbf{H}\{Y, Z\}$ by the invariance axiom. Hence

$$\begin{aligned}\mathbf{H}\{X \mid Y\} &= \mathbf{H}\{X \mid Y, Z\} \\ &= \mathbf{H}\{X, Y, Z\} - \mathbf{H}\{Y, Z\} \\ &= \mathbf{H}\{X, Z\} - \mathbf{H}\{Z\} \\ &= \mathbf{H}\{X \mid Z\} \\ &\leq \mathbf{H}\{X\},\end{aligned}$$

where the inequality on the last line follows from the previous paragraph.

The general case follows from the continuity axiom and the fact that any discrete random variable, regarded as a vector in $l_1(\mathbf{R})$, by vectors with finite support, and these in turn can be approximated by vectors of finite support and rational coordinates. ∎

By the additivity axiom, the previous proposition is equivalent to

$$\mathbf{H}\{X, Y\} \leq \mathbf{H}\{X\} + \mathbf{H}\{Y\}.$$

and we shall use this to prove the following submodularity inequality.

Proposition 1.14 (*Submodularity*). *Suppose X, Y, Z, and W are random variables such that (Z, W) determines X, Z determines Y, and W also determines Y. Then*

$$\mathbf{H}\{X\} + \mathbf{H}\{Y\} \leq \mathbf{H}\{Z\} + \mathbf{H}\{W\}.$$

Proof. The hypotheses give the three inequalities

$$\mathbf{H}\{X\} \leq \mathbf{H}\{Z, W\}, \quad \mathbf{H}\{Y\} \leq \mathbf{H}\{Z\}, \quad \text{and} \quad \mathbf{H}\{Y\} \leq \mathbf{H}\{W\}.$$

From this we see that

$$2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\} \leq 2\,\mathbf{H}\{Z, W\} + \mathbf{H}\{Z\} + \mathbf{H}\{W\}.$$

But then since conditioning does not increase entropy, we have

$$2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\} \leq 2\,\mathbf{H}\{Z\} + 2\,\mathbf{H}\{W\},$$

whence dividing both sides by 2 completes the proof. ∎

The submodularity inequality is often stated in terms of a triple of random variables in terms of the *conditional mutual information*, which is defined by

$$\begin{aligned}\mathbf{I}\{X : Y \mid Z\} &= \sum_{z \in C} \mathbf{P}\{Z = z\}\, \mathbf{I}\big\{(X \mid Z = z) : (Y \mid Z = z)\big\} \\ &= \mathbf{H}\{X \mid Z\} - \mathbf{H}\{X \mid Y, Z\} \\ &= \mathbf{H}\{Y \mid Z\} - \mathbf{H}\{Y \mid X, Z\}.\end{aligned}$$

Proposition 1.15. *Let X, Y, and Z be discrete random variables. Then*

$$\mathbf{H}\{X, Y, Z\} + \mathbf{H}\{Z\} \leq \mathbf{H}\{X, Z\} + \mathbf{H}\{Y, Z\},$$

which is equivalent to

$$\mathbf{I}\{X : Y \mid Z\} \geq 0.$$

Equality holds if and only if X and Y are independent conditional on Z.

Proof. It is clear from the additivity axiom that both sides in the first inequality are equal to $\mathbf{H}\{X\} + \mathbf{H}\{Y\} + 2\,\mathbf{H}\{Z\}$ if and only if X and Y are independent conditional on Z.

To prove the first inequality, note that (X, Y, Z) is jointly determined by (X, Z) and (Y, Z), and Z is determined by both (X, Z) and (Y, Z) separately, then apply the previous proposition. Now by the definition of conditional mutual information,

$$\begin{aligned}\mathbf{I}\{X : Y \mid Z\} &= \sum_{z \in C} \mathbf{P}\{Z = z\}\, \mathbf{I}\big\{(X \mid Z = z) : (Y \mid Z = z)\big\} \\ &= \mathbf{H}\{X \mid Z\} + \mathbf{H}\{Y \mid Z\} - \mathbf{H}\{X, Y \mid Z\} \\ &= \mathbf{H}\{X, Z\} - 2\,\mathbf{H}\{Z\} + \mathbf{H}\{Y, Z\} - \mathbf{H}\{X, Y, Z\} + \mathbf{H}\{Z\} \\ &= \mathbf{H}\{X, Z\} + \mathbf{H}\{Y, Z\} - \mathbf{H}\{X, Y, Z\} - \mathbf{H}\{Z\},\end{aligned}$$

and this proves that the first statement is equivalent to the second. ∎

2. Group-valued random variables

Now we will examine the case where the random variables in question take values in an abelian group G, meaning we can take sums $X + Y$ and differences $X - Y$ of them. Note that if we condition on Y, then the values taken by $X + Y$ are in bijection with values taken by X. This leads to the following proposition.

Proposition 2.1. *Let X and Y be random variables each taking finitely many values in an abelian group G. We have*

$$\max\big(\mathbf{H}\{X\}, \mathbf{H}\{Y\}\big) - \mathbf{I}\{X : Y\} \le \mathbf{H}\{X \pm Y\}.$$

Furthermore, for any random variable Z, we have the conditional version

$$\max\big(\mathbf{H}\{X \mid Z\}, \mathbf{H}\{Y \mid Z\}\big) - \mathbf{I}\{X : Y \mid Z\} \le \mathbf{H}\{X \pm Y \mid Z\}$$

of the same statement.

Proof. Since conditioning does not increase entropy, we have

$$\mathbf{H}\{X \pm Y\} \ge \mathbf{H}\{X \pm Y \mid Y\},$$

and since the probabilities $\mathbf{P}\{X + Y = z \mid Y = y\} = \mathbf{P}\{X = z - y \mid Y = y\}$ for all $z \in G$, by invariance we have

$$\mathbf{H}\{X \pm Y\} \ge \mathbf{H}\{X \mid Y\} = \mathbf{H}\{X\} - \mathbf{I}\{X : Y\}.$$

Repeating the same argument but exchanging the rôles of X and Y, we get

$$\mathbf{H}\{X \pm Y\} \ge \mathbf{H}\{Y \mid X\} = \mathbf{H}\{Y\} - \mathbf{I}\{X : Y\},$$

so

$$\mathbf{H}\{X \pm Y\} \ge \max\big(\mathbf{H}\{X\}, \mathbf{H}\{Y\}\big) - \mathbf{I}\{X : Y\}.$$

Now let Z be any random variable with finite support.

$$\begin{aligned}\mathbf{H}\{X \pm Y \mid Z\} &= \sum_{z \in G} \mathbf{P}\{Z = z\}\, \mathbf{H}\{X \pm Y \mid Z = z\} \\ &\geq \Big(\max\big(\mathbf{H}\{X \mid Z\}, \mathbf{H}\{Y \mid Z\}\big) - \mathbf{I}\{X : Y \mid Z\}\Big) \sum_{z \in G} \mathbf{P}\{Z = z\} \\ &= \max\big(\mathbf{H}\{X \mid Z\}, \mathbf{H}\{Y \mid Z\}\big) - \mathbf{I}\{X : Y \mid Z\},\end{aligned}$$

which completes the proof. ∎

Corollary 2.2. *If X and Y are independent, then*

$$\max\big(\mathbf{H}\{X\}, \mathbf{H}\{Y\}\big) \leq \mathbf{H}\{X \pm Y\}.$$

Proof. The mutual information $\mathbf{I}\{X : Y\}$ is zero whenever X and Y are independent. ∎

Entropic Ruzsa distance. In additive combinatorics, whenever we have two finite subsets A and B of the same abelian group, we can compute the Ruzsa distance

$$d(A, B) = \lg \frac{|A - B|}{\sqrt{|A| \cdot |B|}}$$

between them. (This satisfies all the axioms of a metric except the one requiring $d(A, A) = 0$ for all sets A.) This "distance" is called the *Ruzsa distance* as it was first defined by I. Ruzsa [1].

The entropic analogue of the Ruzsa distance is defined as follows. For finitely supported random variables X and Y taking values in the same abelian group, we let X' and Y' be independent copies of X and Y, respectively, and define the *entropic Ruzsa distance* by

$$\mathbf{d}\{X, Y\} = \mathbf{H}\{X' - Y'\} - \frac{\mathbf{H}\{X'\}}{2} - \frac{\mathbf{H}\{Y'\}}{2}.$$

This definition, first established by T. Tao [5], only depends on the individual distributions of X and Y and does not require them to have the same sample space. When X and Y are independent, the entropic Ruzsa distance enjoys some nice properties.

Proposition 2.3. *Let X and Y be independent taking values in the same group. Then*

$$\big|\mathbf{H}\{X\} - \mathbf{H}\{Y\}\big| \leq 2\,\mathbf{d}\{X, Y\}$$

and

$$\max\big(\mathbf{H}\{X - Y\} - \mathbf{H}\{X\}, \mathbf{H}\{X - Y\} - \mathbf{H}\{Y\}\big) \leq 2\,\mathbf{d}\{X, Y\}$$

Proof. By independence we have

$$2\,\mathbf{d}\{X, Y\} = 2\,\mathbf{H}\{X - Y\} - \mathbf{H}\{X\} - \mathbf{H}\{Y\}.$$

Corollary 2.2 tells us that $2\,\mathbf{H}\{X-Y\} \geq 2\,\mathbf{H}\{X\}$, so this is at least $\mathbf{H}\{X\} - \mathbf{H}\{Y\}$. The same corollary also says that $2\,\mathbf{H}\{X-Y\} \geq 2\,\mathbf{H}\{Y\}$, which allows us to add the absolute value bars to the inequality.

The second inequality is proved using the same corollary, but only applying it to one of the $\mathbf{H}\{X-Y\}$ terms. ∎

Similarly to the Ruzsa distance on sets, we don't necessarily have $\mathbf{d}\{X,X\} = 0$, but we do have the triangle inequality, which shall now prove.

Proposition 2.4. *Let X, Y, and Z be random variables with finite support in the same abelian group. Then*

$$\mathbf{d}\{X,Z\} \leq \mathbf{d}\{X,Y\} + \mathbf{d}\{Y,Z\},$$

which is equivalent to

$$\mathbf{H}\{X'-Z'\} \leq \mathbf{H}\{X'-Y'\} + \mathbf{H}\{Y'-Z'\} - \mathbf{H}\{Y'\}$$

for X', Y', and Z' independent and distributed as X, Y, and Z, respectively.

Proof. That the two statements are equivalent is easily obtained by expanding the definition of entropic Ruzsa distance and cancelling some terms. So without loss of generality, we may assume that that X, Y, and Z are independent and just prove the second statement.

By submodularity, we have $\mathbf{I}\big\{(X-Y:Z) \mid X-Z\big\} \geq 0$, so

$$\begin{aligned} 0 &\leq \mathbf{I}\big\{(X-Y:Z) \mid X-Z\big\} \\ &\leq \mathbf{H}\{X-Y \mid X-Z\} + \mathbf{H}\{Z \mid X-Z\} - \mathbf{H}\{X-Y, Z \mid X-Z\} \\ &\leq \mathbf{H}\{X-Y, X-Z\} + \mathbf{H}\{Z, X-Z\} - \mathbf{H}\{X-Y, Z, X-Z\} - \mathbf{H}\{X-Z\}. \end{aligned} \tag{1}$$

Now, since the values $(x-y, x-z)$ taken by $(X-Y, X-Z)$ are in bijection with values $(x-z, y-z)$ taken by $(X-Z, Y-Z)$ via the map $(v,w) \mapsto (w, w-v)$, by the invariance axiom we have

$$\mathbf{H}\{X-Y, X-Z\} = \mathbf{H}\{X-Z, Y-Z\} \leq \mathbf{H}\{X-Z\} + \mathbf{H}\{Y-Z\}.$$

Similar invocations of the invariance axiom give

$$\mathbf{H}\{Z, X-Z\} = \mathbf{H}\{X,Z\}$$

and

$$\mathbf{H}\{X-Y, Z, X-Z\} = \mathbf{H}\{X,Y,Z\} = \mathbf{H}\{X,Z\} + \mathbf{H}\{Y\},$$

where in the latter statement the second equality follows from the fact that (X,Y) and Z are independent. Substituting these three inequalities into (1), we have

$$0 \leq \mathbf{H}\{X-Y\} + \mathbf{H}\{Y-Z\} + \mathbf{H}\{X,Z\} - \mathbf{H}\{X,Z\} + \mathbf{H}\{Y\} - \mathbf{H}\{X-Z\},$$

whence

$$\mathbf{H}\{X - Z\} \le \mathbf{H}\{X - Y\} + \mathbf{H}\{Y - Z\} + \mathbf{H}\{Y\},$$

which completes the proof. ∎

We also define a conditional version of the entropic Ruzsa distance. If X and Y are G-valued random variables with finite support and Z and W are any random variables with finite supports A and B respectively, then we define

$$\mathbf{d}\{X \mid Z; Y \mid W\} = \sum_{z \in A} \sum_{w \in B} \mathbf{P}\{Z = z\}\, \mathbf{P}\{W = w\}\, \mathbf{d}\big\{(X \mid Z = z); (Y \mid W = w)\big\}.$$

If (X', Z') and (Y', W') are independent copies of (X, Z) and (Y, W) respectively, then this distance is also given by the formula

$$\mathbf{d}\{X \mid Z; Y \mid W\} = \mathbf{H}\{X' - Y' \mid Z', W'\} - \frac{\mathbf{H}\{X' \mid Z'\}}{2} - \frac{\mathbf{H}\{Y' \mid W'\}}{2}.$$

The following inequality relates the conditional and unconditional versions of Ruzsa distance.

Lemma 2.5 ([2], *Lemma 5.1*). *Let (X, Z), and (Y, W) be random variables, with X and Y taking values in the same abelian group. Then*

$$\mathbf{d}\{X \mid Z; Y \mid W\} \le \mathbf{d}\{X, Y\} + \frac{1}{2}\,\mathbf{I}\{X : Z\} + \frac{1}{2}\,\mathbf{I}\{Y : W\}.$$

Proof. Letting (X', Z') and (Y', W') be independent copies of the given random variables, we expand

$$\begin{aligned}
\mathbf{d}\{X \mid Z; Y \mid W\} &= \mathbf{H}\{X' - Y' \mid Z', W'\} - \frac{\mathbf{H}\{X' \mid Z'\}}{2} - \frac{\mathbf{H}\{Y' \mid W'\}}{2} \\
&\le \mathbf{H}\{X' - Y'\} - \frac{\mathbf{H}\{X', Z'\} - \mathbf{H}\{Z'\}}{2} \\
&\qquad - \frac{\mathbf{H}\{Y', W'\} - \mathbf{H}\{W'\}}{2} \\
&= \mathbf{d}\{X - Y\} + \frac{\mathbf{H}\{X'\} + \mathbf{H}\{Z'\} - \mathbf{H}\{X', Z'\}}{2} \\
&\qquad + \frac{\mathbf{H}\{Y'\} + \mathbf{H}\{W'\} - \mathbf{H}\{Y', W'\}}{2} \\
&= \mathbf{d}\{X - Y\} + \frac{1}{2}\,\mathbf{I}\{X : Z\} + \frac{1}{2}\,\mathbf{I}\{Y : W\},
\end{aligned}$$

by independence and the fact that conditioning does not increase entropy. ∎

The sum-difference inequality. The Ruzsa triangle inequality bounds the size of a difference set by passing through a different subset. There is another inequality that relates the size of a sumset with the size of a difference set. If A and B are nonempty finite subsets of an abelian group G, then

$$|A + B| \le \frac{|A - B|^3}{|A||B|}.$$

If we replace cardinalities by exponentials of entropies, then we obtain the statement of the following proposition.

Proposition 2.6. *Let X and Y be independent random variables taking values in the same abelian group. Then*

$$\mathbf{H}\{X+Y\} \le 3\,\mathbf{H}\{X-Y\} - \mathbf{H}\{X\} - \mathbf{H}\{Y\}.$$

Before we proceed to the proof, we establish the following definition, which will be needed later in these notes as well. Let X and Y be random variables (not necessarily independent). We say that X_1 and Y_1 are *conditionally independent trials of X and Y relative to Z* if for all z in the range of Z, the random variables distributed as $(X_1 \mid Z = z)$ and $(Y_1 \mid Z = z)$ are independent, $(X_1 \mid Z = z)$ has the same distribution as $(X \mid Z = z)$, and similarly for Y_1 and Y. In particular, if $X = Y$ and X_1 and X_2 are conditionally independent trials of X relative to Z, we have

$$\mathbf{H}\{X_1, X_2 \mid Z\} = \mathbf{H}\{X_1 \mid Z\} + \mathbf{H}\{X_2 \mid Z\} = 2\,\mathbf{H}\{X \mid Z\},$$

by additivity and independence. From this we obtain

$$\mathbf{H}\{X_1, X_2, Z\} = 2\,\mathbf{H}\{X \mid Z\} + \mathbf{H}\{Z\} = 2\,\mathbf{H}\{X, Z\} - \mathbf{H}\{Z\}. \tag{2}$$

It is also important to observe that (X_1, Z) and (X_2, Z) both have the same distributions as (X, Z).

Proof. Let (X_1, Y_1) and (X_2, Y_2) be conditionally independent trials of (X, Y) relative to $X - Y$. Since (X, Y) determines $X - Y$, we have $X_1 - Y_1 = X - Y = X_2 - Y_2$. Let (X_3, Y_3) be another trial of (X, Y) independent of (X_1, X_2, Y_1, Y_2). Then

$$X_3 + Y_3 = X_3 + Y_3 + X_1 - Y_1 - X_2 + Y_2 = (X_3 - Y_2) - (X_1 - Y_3) + X_2 + Y_1,$$

so $(X_3 - Y_2, X_1 - Y_3, X_2, Y_1)$ and (X_3, Y_3) each determine $X_3 + Y_3$. On the other hand, $(X_3 - Y_2, X_1 - Y_3, X_2, Y_1)$ and (X_3, Y_3) together determine the sextuple $(X_1, X_2, X_3, Y_1, Y_2, Y_3)$, so by the submodularity inequality, we have

$$\begin{aligned} \mathbf{H}\{X_1, X_2, X_3, Y_1, Y_2, Y_3\} + \mathbf{H}\{X_3 + Y_3\} \\ \le \mathbf{H}\{X_3 - Y_2, X_1 - Y_3, X_2, Y_1\} + \mathbf{H}\{X_3, Y_3\}. \end{aligned} \tag{3}$$

We have

$$\mathbf{H}\{X_3, Y_3\} = \mathbf{H}\{X, Y\} = \mathbf{H}\{X\} + \mathbf{H}\{Y\}$$

by independence of X and Y, and since (X_3, Y_3) and (X_1, X_2, Y_1, Y_2) are independent, we have

$$\begin{aligned} \mathbf{H}\{X_1, X_2, X_3, Y_1, Y_2, Y_3\} &= \mathbf{H}\{X_1, X_2, Y_1, Y_2\} + \mathbf{H}\{X_3, Y_3\} \\ &= \mathbf{H}\{X_1, Y_1, X_2, Y_2, X - Y\} + \mathbf{H}\{X\} + \mathbf{H}\{Y\} \\ &= 2\,\mathbf{H}\{X, Y, X - Y\} - \mathbf{H}\{X - Y\} + \mathbf{H}\{X\} + \mathbf{H}\{Y\} \\ &= 2\,\mathbf{H}\{X, Y\} - \mathbf{H}\{X - Y\} + \mathbf{H}\{X\} + \mathbf{H}\{Y\} \\ &= 3\,\mathbf{H}\{X\} + 3\,\mathbf{H}\{Y\} - \mathbf{H}\{X - Y\}, \end{aligned}$$

where in the third line we applied (2). On the other hand,

$$\mathbf{H}\{X_3 + Y_3\} = \mathbf{H}\{X + Y\}$$

and

$$\begin{aligned}\mathbf{H}\{X_3 - Y_2, &X_1 - Y_3, X_2, Y_1\} \\ &\le \mathbf{H}\{X_3 - Y_2\} + \mathbf{H}\{X_1 - Y_3\} + \mathbf{H}\{X_2\} + \mathbf{H}\{Y_1\} \\ &= 2\,\mathbf{H}\{X - Y\} + \mathbf{H}\{X\} + \mathbf{H}\{Y\}.\end{aligned}$$

Substituting everything into (3) yields

$$\begin{aligned}3\,\mathbf{H}\{X\} + 3\,\mathbf{H}\{Y\} &- \mathbf{H}\{X - Y\} + \mathbf{H}\{X + Y\} \\ &\le 2\,\mathbf{H}\{X - Y\} + 2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\},\end{aligned}$$

and the desired inequality follows upon rearrangement of terms. ∎

The entropic sum-difference inequality can also be stated in terms of entropic Ruzsa distances; independence is not necessary here because independent trials are baked into the definition of entropic Ruzsa distance.

Corollary 2.7. *Let X and Y be discrete random variables taking values in the same abelian group. Then*

$$\mathbf{d}\{X, -Y\} \le 3\,\mathbf{d}\{X, Y\}. \quad ∎$$

3. The Plünnecke–Ruzsa inequality

In additive combinatorics, one of the most useful sumset inequalities is the following.

Theorem 3.1 (*Plünnecke–Ruzsa inequality*). *Let A and B be finite subsets of an abelian group and suppose that $|A + B| \le K|A|$ for some constant K. Then for any integers $r, s \ge 0$, not both zero, we have $|rB - sB| \le K^{r+s}|A|$.* ∎

In this section we will develop an entropic analogue of this statement, in which sets are replaced by random variables of finite support and cardinality is replaced with the exponential of entropy. First, a technical lemma.

Lemma 3.2 ([2], *Lemma A.1*). *Let X, Y, and Z be independent random variables taking values in a common abelian group. Then*

$$\mathbf{H}\{X + Y + Z\} - \mathbf{H}\{X + Y\} \le \mathbf{H}\{Y + Z\} - \mathbf{H}\{Y\}.$$

Proof. By submodularity, the quantity $\mathbf{I}\{X : Z \mid X + Y + Z\}$ is nonnegative, so we have

$$\begin{aligned}0 &\le \mathbf{I}\{X : Z \mid X + Y + Z\} \\ &= \mathbf{H}\{X, X + Y + Z\} + \mathbf{H}\{Z, X + Y + Z\} \\ &\qquad - \mathbf{H}\{X, Z, X + Y + Z\} - \mathbf{H}\{X + Y + Z\}.\end{aligned}$$

Since X, Y, and Z are independent, we have

$$\mathbf{H}\{X, X+Y+Z\} = \mathbf{H}\{X, Y+Z\} = \mathbf{H}\{X\} + \mathbf{H}\{Y+Z\},$$

where in the first equality we use invariance. By similar reasoning we have

$$\mathbf{H}\{Z, X+Y+Z\} = \mathbf{H}\{Z\} + \mathbf{H}\{X+Y\}$$

and

$$\mathbf{H}\{X, Z, X+Y+Z\} = \mathbf{H}\{X\} + \mathbf{H}\{Y\} + \mathbf{H}\{Z\}.$$

Plugging these three identities into the inequality above yields

$$\begin{aligned} 0 &\le \mathbf{H}\{X\} + \mathbf{H}\{Y+Z\} + \mathbf{H}\{Z\} + \mathbf{H}\{X+Y\} \\ &\qquad - \mathbf{H}\{X\} - \mathbf{H}\{Y\} - \mathbf{H}\{Z\} - \mathbf{H}\{X+Y+Z\} \\ &= \mathbf{H}\{Y+Z\} + \mathbf{H}\{X+Y\} - \mathbf{H}\{Z\} - \mathbf{H}\{X+Y+Z\}, \end{aligned}$$

whence the claim follows upon rearranging. ∎

From here we are not far from proving the entropic Plünnecke–Ruzsa inequality, a result of T. Tao.

Theorem 3.3. *Let $X, Y_1, \dots, Y_m$ be independent random variables of finite entropy taking values in an abelian group G, such that*

$$\mathbf{H}\{X+Y_i\} \le \mathbf{H}\{X\} + \log K_i$$

for all $1 \le i \le m$ and some scalars $K_1, \dots, K_m \ge 1$. Then

$$\mathbf{H}\{X+Y_1+\cdots+Y_m\} \le \mathbf{H}\{X\} + \log(K_1 \cdots K_m).$$

Proof. We prove the claim by induction on m. If $m = 1$, then we are done by hypothesis. Now suppose that $\mathbf{H}\{X+Y_1+\cdots+Y_{m-1}\} \le \mathbf{H}\{X\}+\log(K_1 \cdots K_{m-1})$. Then by the previous lemma, the induction hypothesis, and the hypothesis on $\mathbf{H}\{X+Y_m\}$, we have

$$\begin{aligned} \mathbf{H}\{Y_1+\cdots+Y_{m-1}+X+Y_m\} &\le \mathbf{H}\{Y_1+\cdots+Y_{m-1}+X\} \\ &\qquad + \mathbf{H}\{X+Y_m\} - \mathbf{H}\{X\} \\ &\le \mathbf{H}\{X\} + \log(K_1 \cdots K_{m-1}) + \log K_m \\ &\le \mathbf{H}\{X\} + \log(K_1 \cdots K_m), \end{aligned}$$

which is what we sought to prove. ∎

We can make this look bit more like the version of the Plünnecke–Ruzsa inequality above by using the triangle inequality.

Corollary 3.4 (*Entropic Plunnecke–Ruzsa inequality*). *Let X and Y be random variables with $\mathbf{H}\{X+Y\} \leq \mathbf{H}\{X\} + \log K$. Then for any $r, s \geq 0$ not both zero, we have*

$$\mathbf{H}\{Y_1 + \cdots + Y_r - Z_1 - \cdots - Z_s\} \leq \mathbf{H}\{X\} + (r+s)\log K,$$

where $Y_1, \ldots, Y_r, Z_1, \ldots, Z_s$ are independent copies of Y.

Proof. By the entropic Ruzsa triangle inequality, we have

$$\begin{aligned}\mathbf{H}\{Y_1 + \cdots + Y_r - Z_1 - \cdots - Z_s\} \leq \\ \mathbf{H}\{Y_1 + \cdots + Y_r + X\} + \mathbf{H}\{-X - Z_1 - \cdots - Z_s\} - \mathbf{H}\{-X\}.\end{aligned}$$

The values of $-X$ are in bijection with values of X, and the values of $-X - Z_1 - \cdots - Z_s$ are in bijection with the values of $X + Z_1 + \cdots + Z_s$ (with the same probabilities in both cases), so by the invariance axiom, we have

$$\begin{aligned}\mathbf{H}\{Y_1 + \cdots + Y_r - Z_1 - \cdots - Z_s\} \leq \\ \mathbf{H}\{X + Y_1 + \cdots + Y_r\} + \mathbf{H}\{X + Z_1 + \cdots + Z_s\} - \mathbf{H}\{X\},\end{aligned}$$

and we can apply the the previous theorem twice to get

$$\begin{aligned}\mathbf{H}\{Y_1 + \cdots + Y_r - Z_1 - \cdots - Z_s\} &\leq \mathbf{H}\{X\} + \log(K^r) + \log(K^s), \\ &= \mathbf{H}\{X\} + (r+s)\log K. \quad \blacksquare\end{aligned}$$

We conclude this section by recording two consequences of Lemma 3.2 and will be of use in the proof of the polynomial Freiman–Ruzsa theorem.

Lemma 3.5 ([2], *Lemma 5.2*). *Let X, Y, and Z be random variables taking values in an abelian group, with Y and Z independent. Then*

$$\begin{aligned}\mathbf{d}\{X, Y - Z\} - \mathbf{d}\{X, Y\} &\leq \frac{\mathbf{H}\{Y - Z\} - \mathbf{H}\{Y\}}{2} \\ &= \frac{1}{2}\,\mathbf{d}\{Y, Z\} + \frac{1}{4}\,\mathbf{H}\{Z\} - \frac{1}{4}\,\mathbf{H}\{Y\}\end{aligned}$$

and

$$\begin{aligned}\mathbf{d}\{X; Y \mid Y - Z\} - \mathbf{d}\{X, Y\} &\leq \frac{\mathbf{H}\{Y - Z\} - \mathbf{H}\{Z\}}{2} \\ &= \frac{1}{2}\,\mathbf{d}\{Y, Z\} + \frac{1}{4}\,\mathbf{H}\{Y\} - \frac{1}{4}\,\mathbf{H}\{Z\}.\end{aligned}$$

Proof. For the first inequality, let X' be a copy of X that is independent of (Y, Z), so that

$$\begin{aligned}\mathbf{d}\{X, Y - Z\} - \mathbf{d}\{X, Y\} &= \mathbf{H}\{X' - Y + Z\} - \frac{1}{2}\,\mathbf{H}\{X'\} - \frac{1}{2}\,\mathbf{H}\{Y - Z\} \\ &\qquad - \mathbf{H}\{X' - Y\} + \frac{1}{2}\,\mathbf{H}\{X'\} + \frac{1}{2}\,\mathbf{H}\{Y\}. \\ &= \mathbf{H}\{X' - Y + Z\} - \frac{1}{2}\,\mathbf{H}\{Y - Z\} \\ &\qquad - \mathbf{H}\{X' - Y\} + \frac{1}{2}\,\mathbf{H}\{Y\}.\end{aligned}$$

Lemma 3.2 with X' in place of X and $-Y$ in place of Y yields

$$\mathbf{H}\{X' - Y + Z\} - \mathbf{H}\{X' - Y\} \le \mathbf{H}\{Y - Z\} - \mathbf{H}\{Y\},$$

so

$$\mathbf{d}\{X, Y - Z\} - \mathbf{d}\{X, Y\} \le \frac{1}{2}\,\mathbf{H}\{Y - Z\} - \frac{1}{2}\,\mathbf{H}\{Y\},$$

as desired. The alternate version of the statement follows directly from the definition of Ruzsa distance, since Y and Z are independent.

For the other inequality, we apply Lemma 2.5 with $W = Y - Z$ (and the variable Z in that lemma set to anything that is independent of X) to obtain

$$\begin{aligned}\mathbf{d}\{X; Y \mid Y - Z\} - \mathbf{d}\{X, Y\} &\le \frac{1}{2}\,\mathbf{I}\{Y : Y - Z\} \\ &= \frac{1}{2}\big(\mathbf{H}\{Y\} + \mathbf{H}\{Y - Z\} - \mathbf{H}\{Y, Y - Z\}\big) \\ &= \frac{1}{2}\big(\mathbf{H}\{Y\} + \mathbf{H}\{Y - Z\} - \mathbf{H}\{Y, Z\}\big) \\ &= \frac{1}{2}\big(\mathbf{H}\{Y - Z\} - \mathbf{H}\{Z\}\big),\end{aligned}$$

where in the last line we used independence of Y and Z. Independence also gives the alternative version of the right-hand-side expression. ∎

Changing variables in the first inequality and then adding the second inequality gives us the following.

Lemma 3.6 ([2], *Lemma 7.1*). *Let X, Y, Z, and W be random variables taking values in the same abelian group, with Y, Z, and W independent. Then*

$$\begin{aligned}&\mathbf{d}\{X; Y - Z \mid Y - Z - W\} - \mathbf{d}\{X, Y\} \\ &\quad\le \frac{1}{2}\big(\mathbf{H}\{Y - Z - W\} + \mathbf{H}\{Y - Z\} - \mathbf{H}\{Y\} - \mathbf{H}\{W\}\big).\end{aligned}$$

Proof. The second inequality of Lemma 3.5 (setting $Y \leftarrow Y - Z$ and $Z \leftarrow W$) yields

$$\mathbf{d}\{X; Y - Z \mid Y - Z - W\} - \mathbf{d}\{X, Y - Z\} \le \frac{\mathbf{H}\{Y - Z - W\} - \mathbf{H}\{W\}}{2},$$

and the first inequality (without any change of variables) is

$$\mathbf{d}\{X, Y - Z\} - \mathbf{d}\{X, Y\} \le \frac{\mathbf{H}\{Y - Z\} - \mathbf{H}\{Y\}}{2}.$$

Adding these two inequalities proves the lemma. ∎

4. The Balog–Szemerédi–Gowers theorem

If A and B are subsets of the same abelian group such that $|A+B|$ is much smaller than $|A| \cdot |B|$, then it stands to reason that there must be a lot of redundancy in $A + B$; that is, many elements of $A + B$ can be expressed as $a + b$ in lots of different ways. To capture this notion, we can define the *additive energy* between two sets A and B to be

$$E(A, B) = \big|\big\{(a, a', b, b') \in A \times A \times B \times B : a + b = a' + b'\big\}\big|.$$

We now define an entropic version of $E(A, B)$. Let X and Y be discrete random variables taking values in the same abelian group. Let (X_1, Y_1) and (X_2, Y_2) be conditionally independent trials of (X, Y) relative to $X + Y$. These $X_1 + Y_1 = X_2 + Y_2 = A + B$. The *entropic additive energy* between X and Y is

$$\mathbf{A}\{X, Y\} = \mathbf{H}\{X_1, Y_1, X_2, Y_2\}.$$

This definition makes clear the analogy between this value and the additive energy of sets, but by conditional independence and the fact that (X_1, Y_1, X_2, Y_2) determines $X_1 + Y_1 = X + Y$, we can rewrite

$$\begin{aligned}\mathbf{A}\{X, Y\} &= \mathbf{H}\{X_1, Y_1, X_2, Y_2, X + Y\} \\ &= 2\,\mathbf{H}\{X, Y, X + Y\} - \mathbf{H}\{X + Y\} \\ &= 2\,\mathbf{H}\{X, Y\} - \mathbf{H}\{X + Y\},\end{aligned}$$

where in the second equality we applied (2). This formula is something we will use often, as it makes no direct mention of the variables (X_1, Y_1) and (X_2, Y_2).

Quantifying the idea that small sumset must imply large additive energy, we have the following proposition.

Proposition 4.1. *Let A and B be finite subsets of an abelian group. If $|A+B| \leq K|A|^{1/2}|B|^{1/2}$ for some constant K, then we have*

$$E(A, B) \geq \frac{1}{K}|A|^{3/2}|B|^{3/2}. \quad \blacksquare$$

Somewhat surprisingly, if we convert these statements into their entropic analogues in the naïve way, as we've been doing, the implication goes the other way! However, we have a weak equivalence (with worse constants in one direction) under the further assumption that the random variables in question are not too dependent.

Proposition 4.2. *Let X and Y be discrete random variables taking values in the same abelian group. If*

$$\mathbf{A}\{X, Y\} \geq \frac{3}{2}\,\mathbf{H}\{X\} + \frac{3}{2}\,\mathbf{H}\{Y\} - \log K, \tag{4}$$

for some constant K, then

$$\mathbf{H}\{X+Y\} \le \frac{1}{2}\,\mathbf{H}\{X\} + \frac{1}{2}\,\mathbf{H}\{Y\} + \log K. \tag{5}$$

If one adds the further assumption that $\mathbf{H}\{X,Y\} \ge \mathbf{H}\{X\}+\mathbf{H}\{Y\}-C$, then (5) *implies* (4) *with a worse constant, namely, we may only conclude*

$$\mathbf{A}\{X,Y\} \ge \frac{3}{2}\,\mathbf{H}\{X\} + \frac{3}{2}\,\mathbf{H}\{Y\} - \log K - 2C. \tag{6}$$

In particular, if X and Y are independent, then we can recover (4) *from* (5).

Proof. Assuming the lower bound on the additive energy, we have

$$2\,\mathbf{H}\{X,Y\} - \mathbf{H}\{X+Y\} \ge \frac{3}{2}\,\mathbf{H}\{X\} + \frac{3}{2}\,\mathbf{H}\{Y\} - \log K,$$

so

$$\begin{aligned}\mathbf{H}\{X+Y\} &\le 2\,\mathbf{H}\{X,Y\} - \frac{3}{2}\,\mathbf{H}\{X\} - \frac{3}{2}\,\mathbf{H}\{Y\} + \log K\\ &\le 2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\} - \frac{3}{2}\,\mathbf{H}\{X\} - \frac{3}{2}\,\mathbf{H}\{Y\} + \log K\\ &= \frac{1}{2}\,\mathbf{H}\{X\} + \frac{1}{2}\,\mathbf{H}\{Y\} + \log K.\end{aligned}$$

On the other hand, assuming this upper bound on $\mathbf{H}\{X+Y\}$, we have

$$\begin{aligned}\mathbf{A}\{X,Y\} &= 2\,\mathbf{H}\{X,Y\} - \mathbf{H}\{X+Y\}\\ &\ge 2\,\mathbf{H}\{X,Y\} - \frac{1}{2}\,\mathbf{H}\{X\} - \frac{1}{2}\,\mathbf{H}\{Y\} - \log K,\end{aligned}$$

and if $2\,\mathbf{H}\{X,Y\} \ge 2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\} - 2C$, then (6) follows directly. ∎

The fact that the implication goes the "wrong" way may seem somewhat baffling. Let us get to the bottom of this. If A and B are subsets of a finite abelian group, we can let X and Y be the uniform distributions on A and B, respectively. We have been operating under the belief that $\mathbf{H}\{X+Y\}$ should correspond (up to taking powers or logarithms) to the size of $A+B$. But this is not true, since X and Y may be given a joint distribution that is not uniform on $A \times B$, even if its marginals are uniform on A and B.

For example, let A and B be subsets of G and consider any regular bipartite graph H on the vertex set $A \cup B$. Let (X,Y) be defined by sampling an edge from H uniformly at random, letting X be its endpoint in A and Y be its endpoint in B. Since the graph is regular, X is uniform on A and Y is uniform on B, but $X+Y$ can only take values $a+b$ where (a,b) is an edge of H. In other words, $X+Y$ samples from the *partial sumset*

$$A +_H B = \big\{a+b : (a,b) \in E(H)\big\},$$

where elements that are represented more times as the sum of edge endpoints are given a greater weight. The way to properly recover the ordinary sumset $A + B$ is to let H be all of $A \times B$, in which case X and Y are independent. The extra assumption we added in Proposition 4.2 is analogous to stipulating that $|H| \geq K|A| \cdot |B|$, so that the resulting X and Y are "nearly" independent.

Simply put, in the entropic setting the bound (4) is stronger than (5) because the entropy $\mathbf{H}\{X, Y\}$ of the joint distribution appears in the formula for $\mathbf{A}\{X, Y\}$, whereas (5) says nothing whatsoever about this joint distribution.

The converse to Proposition 4.1 does not hold in general; that is, large additive energy does not necessarily imply a small sumset. However, there does exist a partial converse, which says that if sets A and B have a large additive energy, then there are dense subsets $A' \subseteq A$ and $B' \subseteq B$ such that the sumset $|A + B|$ is small. This is the celebrated Balog–Szemerédi–Gowers theorem.

Theorem 4.3 (*Balog–Szemerédi–Gowers theorem*). *Let A be a finite subset of an abelian group with $E(A, B) \geq c|A|^{3/2}|B|^{3/2}$. Then there are subsets $A' \subseteq A$ and $B' \subseteq B$ with $|A'| \geq c'|A|$ and $|B'| \geq c''|B|$ such that*

$$|A' + B'| \leq C|A|^{1/2}|B|^{1/2},$$

where c', c'', and C depend only on c.

In the entropy setting, the operation on random variables that corresponds to taking subsets is conditioning. (As a sanity check, recall that conditioning never increases entropy, just as taking subsets never increases cardinality.) The Balog–Szemerédi–Gowers theorem gives us subsets between which we can take a *bona fide* sumset, so its entropic analogue should return conditionings X' and Y' of X and Y relative to some random variable Z, such that

i) X' and Y' are conditionally independent relative to Z;
ii) the entropies $\mathbf{H}\{X' \mid Z\}$ and $\mathbf{H}\{Y' \mid Z\}$ are not too small compared to their unconditioned analogues; and
iii) $\mathbf{H}\{X' + Y' \mid Z\}$ is small.

In fact, the conditioning we shall perform is exactly the one used to define additive energy.

First, we need a lemma, which we state separately since it will also be used later in the proof of the polynomial Freiman–Ruzsa theorem.

Lemma 4.4 ([2], *Lemma A.2*). *Let X and Y be discrete random variables taking values in the same abelian group. Let (X_1, Y_1) and (X_2, Y_2) be conditionally independent trials of (X, Y) relative to $X + Y$. Then we have*

$$\max\big(\mathbf{H}\{X_1 - X_2\}, \mathbf{H}\{X_1 - Y_2\}\big) \leq \mathbf{H}\{X + Y\} + 2\,\mathbf{I}\{X : Y\}.$$

The right-hand side of this expression can also be written $2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\} - \mathbf{A}\{X, Y\}$.

Proof. First we perform the proof for $X_1 - Y_2$. Submodularity gives us

$$\mathbf{H}\{X_1, Y_1, X_1 - Y_2\} + \mathbf{H}\{X_1 - Y_2\} \leq \mathbf{H}\{X_1, X_1 - Y_2\} + \mathbf{H}\{Y_1, X_1 - Y_2\}.$$

Since $X_1 + Y_1 = X + Y = X_2 + Y_2$, given $(X_1, Y_1, X_1 - Y_2)$ we can recover the values of X_2 and Y_2. So $(X_1, Y_1, X_1 - Y_2)$ and (X_1, Y_1, X_2, Y_2) determine each other and hence

$$\mathbf{H}\{X_1, Y_1, X_1 - Y_2\} = \mathbf{H}\{X_1, Y_1, X_2, Y_2\} = 2\,\mathbf{H}\{X, Y\} - \mathbf{H}\{X + Y\}.$$

On the other side of the inequality, we have

$$\mathbf{H}\{X_1, X_1 - Y_2\} = \mathbf{H}\{X_1, Y_2\} \le \mathbf{H}\{X\} + \mathbf{H}\{Y\},$$

and similarly

$$\mathbf{H}\{Y_1, X_1 - Y_2\} = \mathbf{H}\{Y_1, X_2 - Y_1\} = \mathbf{H}\{X_2, Y_1\} \le \mathbf{H}\{X\} + \mathbf{H}\{Y\}.$$

Therefore,

$$\begin{aligned}\mathbf{H}\{X_1 - Y_2\} &\le \mathbf{H}\{X + Y\} - 2\,\mathbf{H}\{X, Y\} + 2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\} \\ &= \mathbf{H}\{X + Y\} - \mathbf{I}\{X : Y\}.\end{aligned}$$

The same holds with Y_2 replaced by X_2. ∎

Theorem 4.5 (*Entropic Balog–Szemerédi–Gowers theorem*). *Let X and Y be discrete random variables taking values in the same abelian group, and suppose that*

$$\mathbf{A}\{X, Y\} \ge \frac{3}{2}\,\mathbf{H}\{X\} + \frac{3}{2}\,\mathbf{H}\{Y\} + \log K$$

for some constant K. Then letting (X_1, Y_1) and (X_2, Y_2) be conditionally independent trials of (X, Y) relative to $X + Y$, we have

$$\mathbf{H}\{X_1 \mid X + Y\} \ge \mathbf{H}\{X\} - 2\log K$$

and

$$\mathbf{H}\{Y_2 \mid X + Y\} \ge \mathbf{H}\{Y\} - 2\log K.$$

Furthermore, the variables X_1 and Y_2 are conditionally independent relative to $X + Y$, and we have

$$\mathbf{H}\{X_1 + Y_2 \mid X + Y\} \le \frac{1}{2}\,\mathbf{H}\{X\} + \frac{1}{2}\,\mathbf{H}\{Y\} + \log K.$$

Proof. Using the coupling $X + Y = X_1 + Y_1 = X_2 + Y_2$, we have

$$\begin{aligned}\mathbf{H}\{X_1 \mid X + Y\} &= \mathbf{H}\{X_1, X_1 + Y_1\} - \mathbf{H}\{X + Y\} \\ &= \mathbf{H}\{X, Y\} - \mathbf{H}\{X + Y\} \\ &= \mathbf{A}\{X, Y\} - \mathbf{H}\{X, Y\} \\ &\ge \frac{3}{2}\,\mathbf{H}\{X\} + \frac{3}{2}\,\mathbf{H}\{Y\} - \log K - \mathbf{H}\{X, Y\} \\ &\ge \frac{1}{2}\,\mathbf{H}\{X\} + \frac{1}{2}\,\mathbf{H}\{Y\} - \log K.\end{aligned}$$

where in the fourth line we used the hypothesis on $\mathbf{A}\{X, Y\}$, and in the last line we observed that $\mathbf{H}\{X\} + \mathbf{H}\{Y\} - \mathbf{H}\{X, Y\} \geq 0$. The bound

$$\mathbf{H}\{Y_2, X + Y\} \geq \frac{1}{2}\,\mathbf{H}\{X\} + \frac{1}{2}\,\mathbf{H}\{Y\} - \log K$$

is shown in the exact same way; only the first step differs.

Now, taking the sum of both these bounds, we arrive at

$$\mathbf{H}\{X_1 \mid X + Y\} + \mathbf{H}\{Y_2 \mid X + Y\} \geq \mathbf{H}\{X\} + \mathbf{H}\{Y\} - 2\log K.$$

From this one deduces

$$\begin{aligned}\mathbf{H}\{X_1 \mid X + Y\} &\geq \mathbf{H}\{X\} + \mathbf{H}\{Y_2\} - \mathbf{H}\{Y_2 \mid X + Y\} - 2\log K \\ &\geq \mathbf{H}\{X\} - 2\log K.\end{aligned}$$

The corresponding lower bound on $\mathbf{H}\{Y_2 \mid X + Y\}$ is proved similarly.

It remains to prove the upper bound on $\mathbf{H}\{X_1 + Y_2 \mid X + Y\}$. Note that $(X_1, Y_2, X + Y)$ and $(X_1 - X_2, X + Y)$ jointly determine $(X_1, X_2, X + Y)$. Then given $X_1 - X_2$ and $X + Y$ we can calculate

$$X_1 + Y_2 = X_1 - X_2 + X_2 + Y_2 = (X_1 - X_2) + (X + Y),$$

so $(X_1, Y_2, X+Y)$ and $(X_1-X_2, X+Y)$ each separately determine $(X_1+Y_2, X+Y)$. Hence the submodularity inequality yields

$$\mathbf{H}\{X_1, X_2, X+Y\}+\mathbf{H}\{X_1+Y_2, X+Y\} \leq \mathbf{H}\{X_1, Y_2, X+Y\}+\mathbf{H}\{X_1-X_2, X+Y\}.$$

From $(X_1, X_2, X + Y)$ we can calculate $Y_1 = X + Y - X_1$ and $Y_2 = X + Y - X_2$, so this triple and the triple (X_1, X_2, Y_1, Y_2) determine each other. So the first term above is simply the additive energy between X and Y; that is

$$\mathbf{H}\{X_1, X_2, X + Y\} = \mathbf{H}\{X_1, X_2, Y_1, Y_2\} = 2\,\mathbf{H}\{X, Y\} - \mathbf{H}\{X + Y\}.$$

Now since X_1 and Y_2 are conditionally independent relative to $X + Y$, we have

$$\begin{aligned}\mathbf{H}\{X_1, Y_2, X + Y\} &= \mathbf{H}\{X_1, X + Y\} + \mathbf{H}\{Y_2, X + Y\} - \mathbf{H}\{X + Y\} \\ &= \mathbf{H}\{X, X + Y\} + \mathbf{H}\{Y, X + Y\} - \mathbf{H}\{X + Y\} \\ &= 2\,\mathbf{H}\{X, Y\} - \mathbf{H}\{X + Y\}\end{aligned}$$

For the last term above we split

$$\mathbf{H}\{X_1 - X_2, X + Y\} = \mathbf{H}\{X_1 - X_2 \mid X + Y\} - \mathbf{H}\{X + Y\}.$$

Putting everything together, we obtain

$$\begin{aligned}2\,\mathbf{H}\{X, Y\} - \mathbf{H}\{X+Y\} + \mathbf{H}\{X_1 + Y_2, X + Y\} \\ \leq 2\,\mathbf{H}\{X, Y\} + \mathbf{H}\{X_1 - X_2 \mid X + Y\} - 2\,\mathbf{H}\{X + Y\},\end{aligned}$$

so that

$$\mathbf{H}\{X_1 + Y_2 \mid X + Y\} \le \mathbf{H}\{X_1 - X_2 \mid X + Y\} - 2\,\mathbf{H}\{X + Y\} \le \mathbf{H}\{X_1 - X_2\}.$$

The previous lemma then gives

$$\mathbf{H}\{X_1 + Y_2 \mid X + Y\} \le 2\,\mathbf{H}\{X\} + 2\,\mathbf{H}\{Y\} - \mathbf{A}\{X, Y\},$$

and from our lower bound on $\mathbf{A}\{X, Y\}$, we conclude that

$$\mathbf{H}\{X_1 + Y_2 \mid X + Y\} \le \frac{1}{2}\,\mathbf{H}\{X\} + \frac{1}{2}\,\mathbf{H}\{Y\} + \log K,$$

which is what we wanted to show. ∎

Let us also extract the specific consequence of Lemma 4.4 that we need for the polynomial Freiman–Ruzsa theorem.

Lemma 4.6. *Let X and Y be G-valued random variables, let $Z = X + Y$, and let C denote the support of Z. Then*

$$\sum_{z \in C} \mathbf{P}\{Z = z\}\,\mathbf{d}\big\{(X \mid Z = z); (Y \mid Z = z)\big\} \le 2\,\mathbf{I}\{X : Y\} + 2\,\mathbf{H}\{Z\} - \mathbf{H}\{X, Y\}.$$

Proof. Let (X_1, Y_1) and (X_2, Y_2) be as in the proof of Lemma 4.4. Note that X_1 and Y_2 are conditionally independent copies of X and Y relative to Z, so it suffices to show that

$$\mathbf{H}\{X_1 - Y_2 \mid Z\} - \frac{1}{2}\,\mathbf{H}\{X_1 \mid Z\} - \frac{1}{2}\,\mathbf{H}\{Y_2 \mid Z\} \le 2\,\mathbf{I}\{X : Y\} + 2\,\mathbf{H}\{Z\} - \mathbf{H}\{X, Y\}. \tag{7}$$

Lemma 4.4 tells us that $\mathbf{H}\{X_1 - Y_2\} \le \mathbf{H}\{Z\} + 2\,\mathbf{I}\{X : Y\}$, so that

$$\mathbf{H}\{X_1 - Y_2 \mid Z\} \le \mathbf{H}\{Z\} + 2\,\mathbf{I}\{X : Y\}, \tag{8}$$

since conditioning does not increase entropy. Then we expand

$$\mathbf{H}\{X_1 \mid Z\} = \mathbf{H}\{X_1, X_1 + Y_1\} - \mathbf{H}\{Z\} = \mathbf{H}\{X, Y\} - \mathbf{H}\{Z\}, \tag{9}$$

and $\mathbf{H}\{Y_2 \mid Z\}$ is equal to this quantity as well. Subtracting (9) from (8) gives us (7), which completes the proof. ∎

5. The Freiman–Ruzsa theorem

Both Plünnecke's theorem and the Balog–Szemerédi–Gowers theorem have to do with the ratio $|A + A|/|A|$ of a finite set A. Let us now give this ratio a name.

It is called the *doubling constant* of A. Plünnecke's theorem says that if the doubling constant is at most K, then the ratio $|rA - sA|/|A|$ is at most K^{r+s}. The Balog–Szemerédi–Gowers theorem, on the other hand, says that sets A with large additive energy contain large subsets A and A'' such that $|A' + A''|/|A|$ is bounded from above by some constant.

It is natural to ask whether there is some way to characterise the structure of sets A with small doubling constant (in some asymptotic sense). One reason why A might have $|A+A| \leq K|A|$ is if A is contained in some subgroup H, since subgroups are closed under addition. If H is not much larger than A, then this would explain why the doubling constant of A is small. It turns out that the converse of this statement holds; that is, if A has small doubling constant, then it has high density as a subset of a subgroup. This is called the Freiman–Ruzsa theorem.

Theorem 5.1 (*Freiman–Ruzsa theorem*). *Let $A \subseteq G = (\mathbf{Z}/r\mathbf{Z})^n$ and suppose that $|A + A| \leq K|A|$. Then there is a subgroup H of G such that $A \subseteq H$ and $|H| \leq K'|A|$, where K' depends only on K and r.*

But subspaces may not be a very efficient way of covering sets with small doubling. One can easily cook up examples where there is enough linear independence in A so that the smallest subspace containing A has size exponential in $|A|$, but A still has small doubling.

On the other hand, in the extreme case that the doubling constant of A is 1, then the structure of A is completely determined, as shown by the following proposition.

Proposition 5.2. *Let A be a finite subset of an abelian group G. If $|A + A| = |A|$, then A is a coset of a subgroup H.*

Proof. First we assume that A contains 0. Then

$$A = A + \{0\} \subseteq A + A,$$

but since $|A + A| = |A|$ this implies that $A + A = A$. Now let $H = \{h \in G : A + h = A$. The above observation shows that $A \subseteq H$. But if $h \in H$, then $A = A + h$ contains the element $0 + h = h$, so in fact $H = A$.

If $x, y \in H$, then $A + x = A$ and $A + y = A$, so subtracting y from both sides of the second identity we obtain $A - y = A$, and adding x to both sides of this, we get

$$A + x - y = A + x = A,$$

so $x - y \in H$. This along with the fact that $0 \in H$ shows that H is a subgroup of G.

Thus in the case where A contains 0, we see that A is actually a subgroup of G, and in the general case we may translate A without changing $|A + A|$, so A is the translate of a subgroup; that is, A is a coset of H. ∎

So in the case that A is a union of not too many cosets, we also expect $A+A$ to be quite small, since each individual coset does not grow under addition (the only possible growth comes from additions between cosets).

It is believed that the converse of this statement holds and yields a more efficient version of the Freiman–Ruzsa theorem; that is, if A has doubling constant bounded above by K, then A is contained in a union of cosets, where the number of cosets needed can be taken to be no more than polynomial in the doubling constant. This conjecture is due to K. Marton.

Conjecture 5.3. *Let $A \subseteq \big(\mathbf{Z}/r\mathbf{Z}\big)^n$ have $|A+A| \leq K|A|$ for some constant K. Then there is a subgroup H with $|H| \leq |A|$ such that A is contained in a union of K^C cosets of H, where C is a constant that can depend on r but not on n or K.*

In 2023, the first special case of this conjecture was proved by W. T. Gowers, B. Green, F. Manners, and T. Tao.

Theorem 5.4 (*Gowers–Green–Manners–Tao,* 2023). *There is a constant C such that the following holds. Let $A \subseteq \mathbf{F}_2^n$ have $|A+A| \leq K|A|$ for some constant K. Then there is a subgroup H with $|H| \leq |A|$ such that A is contained in a union of $2K^C$ cosets of H.*

The reason we need the factor of two is that if A is almost all of a subgroup, then K is very close to 1 and the largest subgroup of $\mathbf{F}_2^n$ has cardinality just slightly above half that of A. That creates the possibility that K^C is much less than 2, making it impossible to satisfy the statement of the conjecture. Adding the factor of two solves this issue because if $A \subseteq H_0$ with $|A| \geq (1-\epsilon)|H_0|$, then we may take a subgroup H of H_0 of index 2 and cover A by two cosets of H.

The proof of this theorem will occupy the remainder of these notes. As a first step, we shall show that Theorem 5.4 follows from an entirely information-theoretic statement. In this and the rest of the notes, for A finite we use the notation U_A to denote the random variable that is uniform on A.

Theorem 5.5. *There is a constant C' such that the following holds. Let $G = \mathbf{F}_2^n$ and let X_1^* and X_2^* be random variables taking values in G. There is some subgroup $H \subseteq G$ such that*

$$\mathbf{d}\{X_1^*, U_H\} + \mathbf{d}\{X_2^*, U_H\} \leq C'\,\mathbf{d}\{X_1^*, X_2^*\}.$$

In the proof that this theorem implies the previous one, we shall need the Ruzsa covering lemma, whose proof is short enough that we include it for completeness.

Lemma 5.6 (*Ruzsa covering lemma*). *Let G be an abelian group and let A and B be finite subsets of G. Then there is a set $L \subseteq A$ of size at most $|A+B|/|B|$ such that $A \subseteq L + B - B$.*

Proof. Let $L = \{a_1, \dots, a_l\}$ be a maximal subset of A with the property that the sets $a_i + B$ are disjoint. For all $a \in A$ there is i such that $(a+B) \cap (a_i + B) \neq \emptyset$,

otherwise we could add a to L and contradict maximality. Equivalently, we can find b and b' such that $a+b=a_i+b'$ so $a \in a_i+B-B$. Hence $A \subseteq L+B-B$. Lastly, $|L|\cdot|B| = |K+B| \leq |A+B|$. ∎

We now show that if Theorem 5.5 holds for some constant C', then Theorem 5.4 holds with $C = C'+1$.

Proof of Theorem 5.4 assuming Theorem 5.5. Let $A \subseteq \mathbf{F}_2^n$ and $K \geq 1$ such that $|A+A| \leq K|A|$. Let U_A be the uniform distribution on A, so that $\mathbf{H}\{U_A\} = \lg|A|$. Letting U_A' and U_A'' be two independent copies of U_A, the sum $U_A' - U_A''$ belongs to $|A+A|$, so $\mathbf{H}\{U_A'+U_A''\} \leq \lg|A+A|$, and we have

$$\mathbf{d}\{U_A, U_A\} = \mathbf{H}\{U_A'+U_A''\} - \mathbf{H}\{U_A\} \leq \lg|A+A| - \lg|A| \leq \lg K,$$

since $U_A'+U_A'' = U_A'-U_A''$ in $\mathbf{F}_2^n$. By Theorem 5.5, we can find $H \subseteq G$ such that

$$2\,\mathbf{d}\{U_A, U_H\} \leq C' \lg K,$$

where U_H is taken to be uniform on H independently of U_A. Independence and the definition of Ruzsa distance let us conclude

$$\mathbf{H}\{U_A - U_H\} \leq \frac{\lg|A| + \lg|H| + C'\lg K}{2},$$

from which applying Proposition 1.9 allows us to procure some $x_0 \in \mathbf{F}_2^n$ such that

$$\mathbf{P}\{U_A - U_H = x_0\} \geq \frac{1}{|A|^{1/2}|H|^{1/2}K^{C'/2}}.$$

Since U_A and U_H are uniformly and independently random on A and H respectively, we rewrite this as

$$\frac{\big|A \cap (H+x_0)\big|}{|A|\cdot|H|} \geq \frac{1}{|A|^{1/2}|H|^{1/2}K^{C'/2}},$$

whence

$$\big|A \cap (H+x_0)\big| \geq \frac{|A|^{1/2}|H|^{1/2}}{K^{C'/2}}.$$

By the Ruzsa covering lemma applied to the sets A and $A \cap (H+x_0)$, there is a set $L \subseteq A$ of size

$$\begin{aligned}
|L| &\leq \frac{\big|A + (A \cap (H+x_0))\big|}{\big|A \cap (H+x_0)\big|} \\
&\leq \frac{|A+A|}{\big|A \cap (H+x_0)\big|} \\
&\leq \frac{K|A|\cdot K^{C'/2}}{|A|^{1/2}|H|^{1/2}} \\
&= \frac{K^{C'/2+1}|A|^{1/2}}{|H|^{1/2}}
\end{aligned}$$

such that

$$A \subseteq L + \big(A \cap (H + x_0)\big) - \big(A \cap (H + x_0)\big) \subseteq L + H.$$

Proposition 2.3 tells us that

$$\big|\lg|H| - \lg|A|\big| = \big|\mathbf{H}\{U_H\} - \mathbf{H}\{H_A\}\big| \leq 2\,\mathbf{d}\{U_A, U_H\} \leq C' \lg K,$$

so

$$\max\left(\frac{|H|}{|A|}, \frac{|A|}{|H|}\right) \leq K^{C'}$$

and we have

$$|L| \leq K^{C'+1}.$$

If $|H| \leq |A|$ then we are done (and the factor of 2 is unnecessary in the theorem statement), since A is a subset of $|L|$ cosets of H.

If not, then we can pick a subgroup H' of H with $|A|/2 \leq |H'| \leq |A|$ by iteratively taking hyperplanes until the condition is met. The group H may be covered by $|H|/|H'| \leq 2|H|/|A|$ translates of H', so A can be covered by

$$\frac{2|H|}{|A|}|L| \leq \frac{2|H|}{|A|} \cdot \frac{K^{C'/2+1}|A|^{1/2}}{|H|^{1/2}} = 2K^{C'/2+1}\frac{|H|^{1/2}}{|A|^{1/2}} \leq 2K^{C'+1}$$

translates of H'. ∎

6. A compactness argument

We have reduced the proof of the Freiman–Ruzsa theorem to proving the information-theoretic Theorem 5.5. In this section, we shall reduce the proof to a different statement via a compactness argument. First we recall some basic topological properties of probability measures.

Fix a finite set G of size n. Ordering the elements of G from 1 to n, one can associate to each probability distribution μ on G the vector

$$\big(\mu(\{1\}), \mu(\{2\}), \ldots, \mu(\{n\})\big)$$

in $\mathbf{R}^n$. (Starting now we will write μ_i instead of $\mu(\{i\})$.) Let $\mathcal{P}(G)$ denote the set of all probability distributions on G. This is a metric space under the *total variation distance* d_{TV}, where if $\mu = (\mu_1, \ldots, \mu_n)$ and $\nu = (\nu_1, \ldots, \nu_n)$ are two probability distributions on G, then

$$d_{\mathrm{TV}}(\mu, \nu) = \sum_{i=1}^{n} |\mu_i - \nu_i|.$$

Regarding the space as a subset of $\mathbf{R}^n$, this norm is equivalent to the L_1 norm; hence the total variation metric on $\mathcal{P}(G)$ is equivalent to the Euclidean metric on

$\mathbf{R}^n$. Since probabilities are in $[0,1]$, the space $\mathcal{P}(G)$ is a subset of the closed set $[0,1]^n$, and since $\mathcal{P}(G)$ is the inverse image of the set $\{1\}$ under the continuous function

$$(x_1, \ldots, x_n) \mapsto x_1 + \cdots + x_n,$$

we conclude that $\mathcal{P}(G)$ is a compact topological space. In passing, we note here that this is the topology under which the continuity axiom asserts that $\mathbf{H}\{X\}$ is continuous.

Next we show that the Ruzsa distance is also continuous with respect to this topology.

Proposition 6.1. *Let G be a finite abelian group. The Ruzsa distance $\mathbf{d}$, regarded as a functional from $\mathcal{P}(G) \times \mathcal{P}(G) \to \mathbf{R}$, is continuous.*

Let X and Y be random variables; without loss of generality, we may assume they are independent, so that

$$\mathbf{d}\{X, Y\} = \mathbf{H}\{X - Y\} - \frac{\mathbf{H}\{X\}}{2} - \frac{\mathbf{H}\{Y\}}{2}.$$

The entropy functional is continuous by axiom, so from here it suffices to show that the function $f : \mathcal{P}(G) \times \mathcal{P}(G) \to \mathcal{P}(G)$ mapping $(X, Y) \mapsto X - Y$ is continuous.

Since all norms on finite-dimensional spaces are equivalent, we may work with $\mathcal{P}(G) \times \mathcal{P}(G)$ and $\mathcal{P}(G)$ as normed spaces using the L_1 norm on $\mathbf{R}^{2n}$ and $\mathbf{R}^n$ respectively. Let $\mu_g = \mathbf{P}\{X = g\}$, $\nu_g = \mathbf{P}\{Y = g\}$, and $\xi_g = \mathbf{P}\{X - Y = g\}$ for all $g \in G$. Given another pair (X', Y') of random variables, define μ'_g, ν'_g, and ξ'_g accordingly. Now let $\epsilon > 0$ and suppose that the vectors (μ, ν) and (μ', ν') have $\left\|(\mu, \nu), (\mu', \nu')\right\|_1 < \delta$ for some $\delta > 0$ to be specified later; writing this out in full, we have

$$\sum_{g \in G} \left(|\mu_g - \mu'_g| + |\nu_g - \nu_g| \right) < \delta.$$

Observe that

$$\xi_g = \mathbf{P}\{X - Y = g\} = \sum_{h \in G} \mathbf{P}\{X = g + h\} \, \mathbf{P}\{Y = h\} = \sum_{h \in G} \mu_{g+h} \nu_h$$

and similarly for ξ'_g. From this we can expand and bound

$$\begin{aligned}
\|\xi, \xi'\|_1 &= \sum_{g \in G} |\xi_g - \xi'_g| \\
&= \sum_{g \in G} \Big| \sum_{h \in G} \mu_{g+h} \nu_h - \sum_{h \in G} \mu_{g+h} \nu_h \Big| \\
&\leq \sum_{g \in G} \sum_{h \in G} \big(|\mu_{g+h} \nu_h - \mu'_{g+h} \nu_h| + |\mu'_{g+h} \nu_h - \mu'_{g+h} \nu'_h| \big) \\
&= \sum_{h \in G} \nu_h \sum_{g \in G} |\mu_{g+h} - \mu'_{g+h}| + \sum_{g \in G} \mu_g \sum_{h \in G} |\nu_h - \nu'_h| \\
&= \sum_{g \in G} |\mu_g - \mu'_g| + \sum_{h \in G} |\nu_h - \nu'_h| \\
&< \delta,
\end{aligned}$$

so in fact we can set $\delta = \epsilon$. ∎

So much for topology. Returning to entropies, we now prove an entropic version of Proposition 5.2.

Proposition 6.2. *Let X be a random variable taking values in an abelian group G. Then $\mathbf{d}\{X, -X\} = 0$ if and only if X is uniformly distributed on a coset of a finite subgroup of G.*

Proof. If X is the uniform distribution on a coset $a + H$ of a finite subgroup H, then letting X' be an independent copy of X, the random variable $X + X'$ is uniform on the coset $a + a + H$, which has the same size as $|a + H|$, so $\mathbf{H}\{X + X'\} = \mathbf{H}\{X\}$, meaning that

$$\mathbf{d}\{X, -X\} = \mathbf{H}\{X + X'\} - \frac{\mathbf{H}\{X\}}{2} - \frac{\mathbf{H}\{X'\}}{2} = 0.$$

On the other hand, suppose that $\mathbf{H}\{X + X'\} = 2\,\mathbf{H}\{X\}$, where X' is an independent copy of X. Then

$$\begin{aligned}
\mathbf{H}\{X + X'\} &= 2\,\mathbf{H}\{X\} - \mathbf{H}\{X\} \\
&= \mathbf{H}\{X, X'\} - \mathbf{H}\{X'\} \\
&= \mathbf{H}\{X + X', X'\} - \mathbf{H}\{X'\} \\
&= \mathbf{H}\{X + X' \mid X'\},
\end{aligned}$$

so $X + X'$ is independent of X'.

Let A denote the support of X. We have shown that the distribution $(X + X' \mid X' = x)$ is the same regardless of our choice of $x \in A$. Then for all $x, y \in A$, the distribution of $X + x$ is the same as that of $X + y$, so the distribution of of $X + x - y$ is the same as that of X. This implies that A is finite, X is the

uniform distribution on A, and adding an element $h \in A - A$ to the set A simply permutes the set.

As in the proof of Proposition 5.2, let $H = \{h \in G : A + h = A\}$. We proved back then that H is a subgroup of G, and the conclusion of the previous paragraph asserts that $A - A \subseteq H$. But letting $h \in H$, the fact that $h + A = A$ implies that $h \in A - A$. We conclude that $A - A$ is a finite subgroup of G, whence A is a coset of a finite subgroup of G. ∎

From this, we are able to prove the special case of Theorem 5.5 in which the Ruzsa distance between the two given random variables is zero.

Lemma 6.3. *Let X_1 and X_2 be random variables taking values in $\mathbf{F}_2^n$ with $\mathbf{d}\{X_1, X_2\} = 0$. Then there exists a subgroup H of $\mathbf{F}_2^n$ such that*

$$\mathbf{d}\{X_1, U_H\} = \mathbf{d}\{X_2, U_H\} = 0.$$

Proof. The triangle inequality gives $\mathbf{d}\{X_1, X_1\} = 0$, and since $-x = x$ in $\mathbf{F}_2^n$, we also have $\mathbf{d}\{X_1, -X_1\} = 0$. By the previous proposition, there is a coset S of a subgroup H such that X_1 has the same distribution as U_S. So

$$\mathbf{d}\{X_1, U_H\} = \mathbf{H}\{U_S - U_H\} - \frac{1}{2}\mathbf{H}\{U_S\} - \frac{1}{2}\mathbf{H}\{U_H\} = 0,$$

since $U_S - U_H$ is uniform on S and $|H| = |S|$. Then by the triangle inequality once again, $\mathbf{d}\{X_2, U_H\} = 0$ as well. ∎

Recall that in the previous section, we reduced the proof of the polynomial Freiman–Ruzsa theorem to a statement involving some random variables X_1^* and X_2^*. For the remainder of these notes, we shall consider these variables to be fixed, and we also fix $\eta = 1/9$ for short. Defining the functional

$$\tau(X_1, X_2) = \mathbf{d}\{X_1, X_2\} - \eta\,\mathbf{d}\{X_1^*, X_1\} - \eta\,\mathbf{d}\{X_2^*, X_2\},$$

we can then reduce the proof of Theorem 5.5 (and consequently the entire proof) to showing the following proposition.

Proposition 6.4. *Let X_1 and X_2 be two $\mathbf{F}_2^n$-valued random variables with $\mathbf{d}\{X_1, X_2\} > 0$. Then there are $\mathbf{F}_2^n$-valued random variables X_1' and X_2' such that*

$$\tau\{X_1', X_2'\} < \tau\{X_1, X_2\}.$$

If we can prove this proposition, then we have Theorem 5.5 with a certain constant C' which shall be made explicit during the proof.

Proof of Theorem 5.5 assuming Proposition 6.4. We proved earlier that $\mathcal{P}(\mathbf{F}_2^n)$ is a compact topological space, thus so is its Cartesian product with itself. By Proposition 6.1, the Ruzsa distance functional is continuous on this space, which implies that τ is as well. Hence τ attains its infimum on the space $\mathcal{P}(\mathbf{F}_2^n)^2$; let X_1 and X_2 be two distributions such that $\tau\{X_1, X_2\}$ is minimal. By the

contrapositive of Proposition 6.4, we must have $\mathbf{d}\{X_1, X_2\} = 0$, so by Lemma 6.3, there exists a subgroup H of $\mathbf{F}_2^n$ such that

$$\mathbf{d}\{X_1, U_H\} = \mathbf{d}\{X_2, U_H\} = 0.$$

This means that

$$\mathbf{d}\{X_1^*, U_H\} \le \mathbf{d}\{X_1^*, X_1\} + \mathbf{d}\{X_1, U_H\} = \mathbf{d}\{X_1^*, X_1\}$$

and

$$\mathbf{d}\{X_1^*, X_1\} \le \mathbf{d}\{X_1^*, U_H\} + \mathbf{d}\{U_H, X_1\} = \mathbf{d}\{X_1^*, U_H\},$$

so $\mathbf{d}\{X_1^*, U_H\} = \mathbf{d}\{X_1^*, X_1\}$. The same holds for X_2^* and X_2. Then since $\mathbf{d}\{X_1, X_2\} = 0$ and by minimality of $\tau\{X_1, X_2\}$, we obtain

$$\begin{aligned}
\eta\,\mathbf{d}\{X_1^*, U_H\} + \eta\,\mathbf{d}\{X_2^*, U_H\} &= \eta\,\mathbf{d}\{X_1, U_H\} + \eta\,\mathbf{d}\{X_2, U_H\} \\
&= \tau\{X_1, X_2\} \\
&\le \tau\{X_1^*, X_2^*\} \\
&= \mathbf{d}\{X_2^*, X_1^*\} + \eta\,\mathbf{d}\{X_1^*, X_2^*\} + \eta\,\mathbf{d}\{X_2^*, X_1^*\} \\
&= (1 + 2\eta)\,\mathbf{d}\{X_1^*, X_2^*\}.
\end{aligned}$$

From our choice of $\eta = 1/9$ we can multiply both sides by 9 to get

$$\mathbf{d}\{X_1^*, U_H\} + \mathbf{d}\{X_2^*, U_H\} \le 11\,\mathbf{d}\{X_1^*, X_2^*\},$$

and furthermore, since

$$\big|\mathbf{d}\{X_1^*, U_H\} - \mathbf{d}\{X_2^*, U_H\}\big| \le \mathbf{d}\{X_1^*, X_2^*\}$$

by two invocations of the triangle inequality, neither $\mathbf{d}\{X_1^*, U_H\}$ nor $\mathbf{d}\{X_2^*, U_H\}$ can be greater than $6\,\mathbf{d}\{X_1^*, X_2^*\}$. ∎

For posterity, let us restate what we have just proved with the constants plugged in.

Theorem 5.5′. *Let X_1^* and X_2^* be random variables taking values in $\mathbf{F}_2^n$. There is some subgroup $H \subseteq \mathbf{F}_2^n$ such that*

$$\mathbf{d}\{X_1^*, U_H\} + \mathbf{d}\{X_2^*, U_H\} \le 11\,\mathbf{d}\{X_1^*, X_2^*\},$$

and furthermore,

$$\max\big(\mathbf{d}\{X_1^*, U_H\}, \mathbf{d}\{X_2^*, U_H\}\big) \le 6\,\mathbf{d}\{X_1^*, X_2^*\}.$$ ∎

7. Entropy distance under homomorphisms

Our goal has been reduced to showing that if X_1 and X_2 have $\mathbf{d}\{X_1, X_2\} > 0$, then there exist X_1' and X_2' with $\tau\{X_1', X_2'\} < \tau\{X_1, X_2\}$. Remember that the

fixed variables X_1^* and X_2^* still play a rôle in this statement, since they feature in the definition of the functional τ.

We have elected to first prove all the technical lemmas then stitch them into a complete proof. We employ this "bottom-up" approach not because it is necessarily more natural, but because the original proof is written in a rather "top-down" style, and the reader may find it enlightening to have both expositions at his or her disposal.

Having said this, the remainder of this section will be devoted to the study of how entropy behaves under homomorphism. This appeared as Proposition 1.4 in [3], and reproved in [2] with the explicit error term shown below.

Proposition 7.1 ([2], *Proposition 4.1*). *Let $\pi : G \to G'$ be a homomorphism of abelian groups and let Z_1 and Z_2 be G-valued random variables. Then*

$$\mathbf{d}\{Z_1, Z_2\} \geq \mathbf{d}\big\{\pi(Z_1), \pi(Z_2)\big\} + \mathbf{d}\big\{Z_1 \mid \pi(Z_1); Z_2 \mid \pi(Z_2)\big\}.$$

If Z_1 and Z_2 are assumed to be independent, then the two sides differ by

$$\mathbf{I}\big\{Z_1 - Z_2 : \big(\pi(Z_1), \pi(Z_2)\big) \,\big|\, \pi(Z_1 - Z_2)\big\}.$$

Proof. Let Z_1' and Z_2' be independent copies of Z_1 and Z_2. Then

$$\begin{aligned}
&\mathbf{d}\big\{Z_1|\pi(Z_1); Z_2 \mid \pi(Z_2)\big\} \\
&\quad= \mathbf{H}\big\{Z_1' - Z_2' \mid \pi(Z_1'), \pi(Z_2')\big\} - \frac{1}{2}\mathbf{H}\big\{Z_1' \mid \pi(Z_1')\big\} - \frac{1}{2}\mathbf{H}\big\{Z_2' \mid \pi(Z_2')\big\} \\
&\quad\leq \mathbf{H}\big\{Z_1' - Z_2' \mid \pi(Z_2')\big\} - \frac{1}{2}\mathbf{H}\big\{Z_1' \mid \pi(Z_1')\big\} - \frac{1}{2}\mathbf{H}\big\{Z_2' \mid \pi(Z_2')\big\} \\
&\quad= \mathbf{H}\big\{Z_1' - Z_2' \mid \pi(Z_1' - Z_2')\big\} - \frac{1}{2}\mathbf{H}\big\{Z_1' \mid \pi(Z_1')\big\} - \frac{1}{2}\mathbf{H}\big\{Z_2' \mid \pi(Z_2')\big\} \\
&\quad= \mathbf{H}\{Z_1' - Z_2'\} - \mathbf{H}\big\{\pi(Z_1' - Z_2')\big\} \\
&\qquad\qquad + \frac{1}{2}\mathbf{H}\{Z_1'\} - \frac{1}{2}\mathbf{H}\big\{\pi(Z_1')\big\} + \frac{1}{2}\mathbf{H}\{Z_2'\} - \frac{1}{2}\mathbf{H}\big\{\pi(Z_2')\big\} \\
&\quad= \mathbf{d}\{Z_1, Z_2\} - \mathbf{d}\big\{\pi(Z_1), \pi(Z_2)\big\},
\end{aligned}$$

where the inequality follows from submodularity, and in the second-last equality we used (three times) the identity

$$\mathbf{H}\big\{X \mid \pi(X)\big\} = \mathbf{H}\{X\} - \mathbf{H}\big\{\pi(X)\big\},$$

which holds for any G-valued random variable X, since X determines $\pi(X)$.

If Z_1 and Z_2 are independent, then the difference between the two sides is

$$\begin{aligned}
&\mathbf{H}\big\{Z_1 - Z_2 \mid \pi(Z_1 - Z_2)\big\} - \mathbf{H}\big\{Z_1 - Z_2 \mid \pi(Z_1), \pi(Z_2)\big\} \\
&\quad= \mathbf{H}\big\{Z_1 - Z_2 \mid \pi(Z_1 - Z_2)\big\} - \mathbf{H}\big\{Z_1 - Z_2 \mid \pi(Z_1), \pi(Z_2), \pi(Z_1 - Z_2)\big\}.
\end{aligned}$$

But one of the definitions of conditional mutual information is

$$\mathbf{I}\{X : Y \mid Z\} = \mathbf{H}\{X \mid Z\} - \mathbf{H}\{X \mid Y, Z\},$$

so letting $X = Z_1 - Z_2$, $Y = \big(\pi(Z_1), \pi(Z_2)\big)$, and $Z = \pi(Z_1 - Z_2)$, we see that the two sides of the inequality differ by exactly the conditional mutual information term claimed above. ∎

We will use this proposition in the following special case.

Corollary 7.2 ([2], *Corollary 4.2*). *Let G be an abelian group and let Y_1, Y_2, Y_3, and Y_4 be independent G-valued random variables. Then*

$$\begin{aligned}\mathbf{d}\{Y_1,Y_2\}+\mathbf{d}\{Y_3,Y_4\} &= \mathbf{d}\{Y_1-Y_3;\,Y_2-Y_4\}+\mathbf{d}\{Y_1\,|\,Y_1-Y_3;\;Y_2\,|\,Y_2-Y_4\}\\ &\quad+\mathbf{I}\{Y_1-Y_2:Y_2-Y_4\,|\,Y_1-Y_2-Y_3+Y_4\}.\end{aligned}$$

Proof. We apply this with the subtraction homomorphism $\pi : G\times G\to G$ given by $\pi(x,y)=x-y$. Then the previous proposition applied to $Z_1=(Y_1,Y_3)$ and $Z_2=(Y_2,Y_4)$ yields the equality

$$\begin{aligned}&\mathbf{d}\big\{(Y_1,Y_3);\;(Y_2,Y_4)\big\}\\ &\quad=\mathbf{d}\big\{Y_1-Y_3,Y_2-Y_4\big\}+\mathbf{d}\big\{(Y_1,Y_3)\,|\,Y_1-Y_3;\;(Y_2,Y_4)\,|\,Y_2-Y_4\big\}\\ &\qquad\quad+\mathbf{I}\big\{(Y_1-Y_2,Y_3-Y_4):(Y_1-Y_3,Y_2-Y_4)\,\big|\,Y_1-Y_2-Y_3-Y_4\big\}\\ &\quad=\mathbf{d}\big\{Y_1-Y_3,Y_2-Y_4\big\}+\mathbf{d}\big\{Y_1\,|\,Y_1-Y_3;\;Y_2\,|\,Y_2-Y_4\big\}.\\ &\qquad\quad+\mathbf{I}\big\{Y_1-Y_2:Y_2-Y_4\,\big|\,Y_1-Y_2-Y_3+Y_4\big\}\end{aligned}$$

where in the last line we used the fact that $(Y_1-Y_2,Y_1-Y_2-Y_3+Y_4)$ determines Y_3-Y_4 as well as the fact that $(Y_2-Y_4,Y_1-Y_2-Y_3+Y_4)$ determines Y_1-Y_3. But by independence,

$$\begin{aligned}\mathbf{d}\big\{(Y_1,Y_3);\;(Y_2,Y_4)\big\}&=\mathbf{H}\{Y_1-Y_2,Y_3-Y_4\}-\frac{\mathbf{H}\{Y_1,Y_3\}}{2}-\frac{\mathbf{H}\{Y_2,Y_4\}}{2}\\ &=\mathbf{d}\{Y_1,Y_2\}+\mathbf{d}\{Y_3,Y_4\},\end{aligned}$$

so the corollary is proved. ∎

8. Sums and fibres

In Section 5, we showed that to prove the polynomial Freiman–Ruzsa theorem, it suffices to produce, for every pair (X_1,X_2) of random variables taking values in $G=\mathbf{F}_2^n$ with $\mathbf{d}\{X_1,X_2\}>0$, a pair (X_1',X_2') with $\tau\{X_1',X_2'\}<\tau\{X_1,X_2\}$. (This was Proposition 6.4.) Recall that the definition of τ involves the "global" variables X_1^* and X_2^*, as well as the global constant $\eta=1/9$.

To prove this statement, it is somewhat more convenient to work with its contrapositive. That is, we shall assume that (X_1,X_2) has $\mathbf{d}\{X_1,X_2\}=k$ and that this pair minimises τ; that is, $\tau\{X_1',X_2'\}\geq\tau\{X_1,X_2\}$ for all (X_1',X_2'). Without loss of generality we may further assume that X_1 and X_2 are independent. Our aim is to show that $k=0$. The first lemma we prove amounts to little more than expansion of definitions, but is worth writing explicitly nonetheless.

Lemma 8.1. *Suppose $(X_1,X_2)\in\mathcal{P}(G)^2$ is a minimiser of τ with $\mathbf{d}\{X_1,X_2\}=k$. Then for all $(X_1',X_2')\in\mathcal{P}(G)^2$,*

$$\mathbf{d}\{X_1',X_2'\}\geq k-\eta\big(\mathbf{d}\{X_1^*,X_1'\}-\mathbf{d}\{X_1^*,X_1\}+\mathbf{d}\{X_2^*,X_2'\}-\mathbf{d}\{X_2^*,X_2\}\big),$$

and for all $(X_1', X_2') \in \mathcal{P}(G)^2$ *and any discrete random variables* Y_1 *and* Y_2*, we have*

$$\begin{aligned}&\mathbf{d}\{X_1' \mid Y_1; X_2' \mid Y_2\}\\ &\quad\geq k - \eta\big(\mathbf{d}\{X_1^*; X_1' \mid Y_1\} - \mathbf{d}\{X_1^*, X_1\} + \mathbf{d}\{X_2^*; X_2' \mid Y_2\} - \mathbf{d}\{X_2^*, X_2\}\big).\end{aligned}$$

Proof. The first inequality follows directly from $\tau\{X_1', X_2'\} \geq \tau\{X_1, X_2\}$ and the definition of τ. For the second inequality, suppose that Y_1 has support A_1 and Y_2 has support A_2 and sum the inequalities $\tau\{X_1 \mid Y_1 = y_1;\ X_2 \mid Y_2 = y_2\} \geq \tau\{X_1, X_2\}$, weighted by $\mathbf{P}\{Y_1 = y_1, Y_2 = y_2\}$ for all $(y_1, y_2) \in A_1 \times A_2$. ∎

The task now is to investigate specific choices of (X_1', X_2') and see what information we can glean about a minimising pair (X_1, X_2). The next lemma proves our first inequality in this direction.

Lemma 8.2. *Let* $G = \mathbf{F}_2^n$ *and suppose* $(X_1, X_2) \in \mathcal{P}(G)^2$ *is a minimiser of* τ *with* $\mathbf{d}\{X_1, X_2\} = k$*. Then letting*

$$I_1 = \mathbf{I}\{X_1 + X_2 : \overline{X}_1 + X_2 \mid X_1 + X_2 + \overline{X}_1 + \overline{X}_2\},$$

where X_1 *and* $\overline{X}_1$ *be copies of* X_1 *and* X_2 *and* $\overline{X}_2$ *be copies of* X_2 *such that* X_1*,* $\overline{X}_1$*,* X_2*, and* $\overline{X}_2$ *are all independent, we have*

$$I_1 \leq 2\eta k.$$

Furthermore, we have

$$\mathbf{H}\{X_1 + X_2 + \overline{X}_1 + \overline{X}_2\} \leq \frac{1}{2}\,\mathbf{H}\{X_1\} + \frac{1}{2}\,\mathbf{H}\{X_2\} + (2+\eta)k - I_1.$$

Proof. Since

$$k = \mathbf{d}\{X_1, X_2\} = \mathbf{d}\{X_1, \overline{X}_2\} = \mathbf{d}\{X_2, \overline{X}_1\},$$

applying Lemma 3.5 four times gives us the inequalities

$$\begin{aligned}\mathbf{d}\{X_1^*; X_1 + \overline{X}_2\} - \mathbf{d}\{X_1^*, X_1\} &\leq \frac{1}{2}k + \frac{1}{4}\,\mathbf{H}\{X_2\} - \frac{1}{4}\,\mathbf{H}\{X_1\},\\ \mathbf{d}\{X_2^*; X_2 + \overline{X}_1\} - \mathbf{d}\{X_2^*, X_2\} &\leq \frac{1}{2}k + \frac{1}{4}\,\mathbf{H}\{X_1\} - \frac{1}{4}\,\mathbf{H}\{X_2\},\\ \mathbf{d}\{X_1^*; X_1 \mid X_1 - \overline{X}_2\} - \mathbf{d}\{X_1^*, X_1\} &\leq \frac{1}{2}k + \frac{1}{4}\,\mathbf{H}\{X_1\} - \frac{1}{4}\,\mathbf{H}\{X_2\},\end{aligned}$$

and

$$\mathbf{d}\{X_2^*; X_2 \mid X_2 - \overline{X}_1\} - \mathbf{d}\{X_2^*, X_2\} \leq \frac{1}{2}k + \frac{1}{4}\,\mathbf{H}\{X_2\} - \frac{1}{4}\,\mathbf{H}\{X_1\}.$$

By summing these four inequalites, we obtain

$$\begin{aligned} 2k \geq\ & \mathbf{d}\{X_1^*; X_1 + \overline{X}_2\} + \mathbf{d}\{X_2^*; X_2 + \overline{X}_1\} \\ & + \mathbf{d}\{X_1^*; X_1 \mid X_1 - \overline{X}_2\} + \mathbf{d}\{X_2^*; X_2 \mid X_2 - \overline{X}_1\} \\ & - 2\,\mathbf{d}\{X_1^*, X_1\} - 2\,\mathbf{d}\{X_2^*, X_2\}. \end{aligned} \tag{10}$$

On the other hand, Corollary 7.2 with (Y_1, Y_2, Y_3, Y_4) set to $(X_1, X_2, \overline{X}_2, \overline{X}_1)$ gives us

$$\begin{aligned} \mathbf{d}\{X_1, X_2\} + \mathbf{d}\{\overline{X}_2, \overline{X}_1\} = \mathbf{d}\{& X_1 - \overline{X}_2;\ X_2 - \overline{X}_1\} \\ & + \mathbf{d}\{X_1 \mid X_1 - \overline{X}_2;\ X_2 \mid X_2 - \overline{X}_1\} \\ & + \mathbf{I}\{X_1 - X_2 : X_2 - \overline{X}_1 \mid X_1 - X_2 - \overline{X}_2 + \overline{X}_1\}. \end{aligned}$$

which can be rewritten

$$\begin{aligned} 2k &= \mathbf{d}\{X_1 + \overline{X}_2;\ X_2 + \overline{X}_1\} + \mathbf{d}\{X_1 \mid X_1 + \overline{X}_2;\ X_2 \mid X_2 + \overline{X}_1\} \\ &\qquad + \mathbf{I}\{X_1 + X_2 : X_2 + \overline{X}_1 \mid X_1 + X_2 + \overline{X}_1 + \overline{X}_2\} \\ &= \mathbf{d}\{X_1 + \overline{X}_2;\ X_2 + \overline{X}_1\} + \mathbf{d}\{X_1 \mid X_1 + \overline{X}_2;\ X_2 \mid X_2 + \overline{X}_1\} + I_1 \end{aligned} \tag{11}$$

since $\mathbf{d}\{X_1, X_2\} = k$ and we are working in $G = \mathbf{F}_2^n$. By Lemma 8.1, we have

$$\begin{aligned} \mathbf{d}\{X_1 + \overline{X}_2; X_2 + \overline{X}_1\} \geq k - \eta\big(& \mathbf{d}\{X_1^*, X_1 + \overline{X}_2\} - \mathbf{d}\{X_1^*, X_1\} \\ & + \mathbf{d}\{X_2^*, X_2 + \overline{X}_1\} - \mathbf{d}\{X_2^*, X_2\}\big) \end{aligned} \tag{12}$$

and

$$\begin{aligned} \mathbf{d}\{X_1 \mid X_1 + \overline{X}_2; X_2 \mid X_2 + \overline{X}_1\} \geq k - \eta\big(& \mathbf{d}\{X_1^*; X_1 \mid X_1 + \overline{X}_2\} - \mathbf{d}\{X_1^*, X_1\} \\ & + \mathbf{d}\{X_2^*, X_2 \mid X_2 + \overline{X}_1\} - \mathbf{d}\{X_2^*, X_2\}\big). \end{aligned} \tag{13}$$

Substituting these two inequalities into (11) yields

$$\begin{aligned} 2k \geq I_1 + 2k - \eta\big(& \mathbf{d}\{X_1^*, X_1 + \overline{X}_2\} + \mathbf{d}\{X_1^*, X_1 \mid X_1 + \overline{X}_2\} \\ & + \mathbf{d}\{X_2^*, X_2 + \overline{X}_1\} + \mathbf{d}\{X_2^*, X_2 \mid X_2 + \overline{X}_2\}, \\ & - 2\,\mathbf{d}\{X_1^*, X_1\} - 2\,\mathbf{d}\{X_2^*, X_2\}\big) \end{aligned}$$

whence

$$\begin{aligned} I_1 \leq \eta\big(& \mathbf{d}\{X_1^*, X_1 + \overline{X}_2\} + \mathbf{d}\{X_1^*, X_1 \mid X_1 + \overline{X}_2\} \\ & + \mathbf{d}\{X_2^*, X_2 + \overline{X}_1\} + \mathbf{d}\{X_2^*, X_2 \mid X_2 + \overline{X}_2\}, \\ & - 2\,\mathbf{d}\{X_1^*, X_1\} - 2\{X_2^*, X_2\}\big) \end{aligned}$$

and we can substitute the inequality (10) to get $I_1 \leq 2\eta k$.

To prove the other claim, we substitute (13) into (11) to obtain

$$\begin{aligned} k &\geq I_1 + \mathbf{d}\{X_1 + \overline{X}_2; X_2 + \overline{X}_1\} - \eta\big(\mathbf{d}\{X_1^*; X_1 \mid X_1 + \overline{X}_2\} - \mathbf{d}\{X_1^*, X_1\} \\ &\qquad + \mathbf{d}\{X_2^*; X_2 \mid X_2 + \overline{X}_1\} - \mathbf{d}\{X_2^*, X_2\}\big) \\ &\geq I_1 + \mathbf{d}\{X_1 + \overline{X}_2; X_2 + \overline{X}_1\} - \eta k, \end{aligned}$$

where in the second line we have used two of the four inequalities begotten by Lemma 3.5. So $(1+\eta)k - I_1 \geq \mathbf{d}\{X_1 + \overline{X}_2; X_2 + \overline{X}_1\}$, which we expand to

$$\begin{aligned} (1+\eta)k - I_1 &\geq \mathbf{H}\{X_1 + X_2 + \overline{X}_1 + \overline{X}_2\} - \frac{1}{2}\mathbf{H}\{X_1 + \overline{X}_2\} - \frac{1}{2}\mathbf{H}\{X_2 + \overline{X}_1\} \\ &= \mathbf{H}\{X_1 + X_2 + \overline{X}_1 + \overline{X}_2\} - \mathbf{d}\{X_1, X_2\} - \frac{1}{2}\mathbf{H}\{X_1\} - \frac{1}{2}\mathbf{H}\{X_2\} \end{aligned}$$

by independence and the fact that subtraction is addition in G. Rearranging terms completes the proof. ∎

We can interpret this lemma as saying that if setting $X_1' = X_1 + \overline{X}_2$ and $X_2' = X_2 + \overline{X}_1$ gives the inequality $\tau\{X_1', X_2'\} \geq \tau\{X_1, X_2\}$, and furthermore if the same holds when setting (X_1', X_2') to any pair

$$\big(X_1 \mid X_1 + \overline{X}_2 = v_1, X_2 \mid X_2 + \overline{X}_1 = v_2\big)$$

of "fibres", where $(v_1, v_2) \in G^2$, then we have a bound on a certain mutual information quantity I_1.

Now we perform a similar analysis, where instead of taking sums $X_1 + \overline{X}_2$ and $X_2 + \overline{X}_1$ across variables, now we sum copies of the *same* variables.

Lemma 8.3. *Let $G = \mathbf{F}_2^n$ and suppose $(X_1, X_2) \in \mathcal{P}(G)^2$ is a minimiser of τ with $\mathbf{d}\{X_1, X_2\} = k$. Then letting*

$$I_2 = \mathbf{I}\{X_1 + X_2 : X_1 + \overline{X}_1 \mid X_1 + X_2 + \overline{X}_1 + \overline{X}_2\},$$

where X_1 and $\overline{X}_1$ be copies of X_1 and X_2 and $\overline{X}_2$ be copies of X_2 such that X_1, $\overline{X}_1$, X_2, and $\overline{X}_2$ are all independent, we have

$$I_2 \leq 2\eta k + \frac{2\eta(2\eta k - I_1)}{1-\eta}.$$

Proof. Applying Lemma 3.5 four times gives us the inequalities

$$\begin{aligned} \mathbf{d}\{X_1^*; X_1 + \overline{X}_1\} - \mathbf{d}\{X_1^*, X_1\} &\leq \frac{1}{2}\mathbf{d}\{X_1, X_1\}, \\ \mathbf{d}\{X_2^*; X_2 + \overline{X}_2\} - \mathbf{d}\{X_2^*, X_2\} &\leq \frac{1}{2}\mathbf{d}\{X_2, X_2\}, \\ \mathbf{d}\{X_1^*; X_1 \mid X_1 - \overline{X}_1\} - \mathbf{d}\{X_1^*, X_1\} &\leq \frac{1}{2}\mathbf{d}\{X_1, X_1\}, \end{aligned} \tag{14}$$

and

$$\mathbf{d}\{X_2^*; X_2 \mid X_2 - \overline{X}_2\} - \mathbf{d}\{X_2^*, X_2\} \le \frac{1}{2}\,\mathbf{d}\{X_2, X_2\}. \tag{15}$$

These inequalities and Lemma 8.1 together give

$$\begin{aligned}\mathbf{d}\{X_1 + \overline{X}_1; X_2 + \overline{X}_2\} &\ge k - \eta\big(\mathbf{d}\{X_1^*, X_1 + \overline{X}_1\} - \mathbf{d}\{X_1^*, X_1\}\\ &\qquad + \mathbf{d}\{X_2^*, X_2 + \overline{X}_2\} - \mathbf{d}\{X_2^*, X_2\}\big)\\ &\ge k - \frac{\eta}{2}\,\mathbf{d}\{X_1, X_1\} - \frac{\eta}{2}\,\mathbf{d}\{X_2, X_2\}\end{aligned} \tag{16}$$

and

$$\begin{aligned}\mathbf{d}\{X_1 \mid X_1 + \overline{X}_1; X_2 \mid X_2 + \overline{X}_2\} &\ge k - \eta\big(\mathbf{d}\{X_1^*; X_1 \mid X_1 + \overline{X}_1\} - \mathbf{d}\{X_1^*, X_1\}\\ &\qquad + \mathbf{d}\{X_2^*, X_2 \mid X_2 + \overline{X}_2\} - \mathbf{d}\{X_2^*, X_2\}\big)\\ &\ge k - \frac{\eta}{2}\,\mathbf{d}\{X_1, X_1\} - \frac{\eta}{2}\,\mathbf{d}\{X_2, X_2\}.\end{aligned} \tag{17}$$

Corollary 7.2, with (Y_1, Y_2, Y_3, Y_4) set to $(X_2, X_1, \overline{X}_2, \overline{X}_1)$ this time, yields

$$\begin{aligned}2k &= \mathbf{d}\{X_2 + \overline{X}_2;\ X_1 + \overline{X}_1\} + \mathbf{d}\{X_2 \mid X_2 + \overline{X}_2;\ X_1 \mid X_1 + \overline{X}_1\}\\ &\qquad + \mathbf{I}\{X_1 + X_2 : X_1 + \overline{X}_1 \mid X_1 + X_2 + \overline{X}_1 + \overline{X}_2\}\\ &= \mathbf{d}\{X_1 + \overline{X}_1;\ X_2 + \overline{X}_2\} + \mathbf{d}\{X_1 \mid X_1 + \overline{X}_1;\ X_2 \mid X_2 + \overline{X}_2\} + I_2,\end{aligned} \tag{18}$$

into which we substitute (16) and (17) to obtain

$$I_2 \le \eta\big(\mathbf{d}\{X_1, X_1\} + \mathbf{d}\{X_2, X_2\}\big).$$

It remains to bound $\mathbf{d}\{X_1, X_1\} + \mathbf{d}\{X_2, X_2\}$.

Since $\mathbf{H}\{X_1 + \overline{X}_1\} = \mathbf{H}\{X_1\} + \mathbf{d}\{X_1, X_1\}$ and similarly for X_2, we may expand

$$\begin{aligned}\mathbf{d}\{X_1 + \overline{X}_1; X_2 + \overline{X}_2\} &= \mathbf{H}\{X_1 + \overline{X}_1 + X_2 + \overline{X}_2\}\\ &\qquad - \frac{1}{2}\,\mathbf{H}\{X_1 + \overline{X}_1\} - \frac{1}{2}\,\mathbf{H}\{X_2 + \overline{X}_2\}\\ &= \mathbf{H}\{X_1 + \overline{X}_1 + X_2 + \overline{X}_2\} - \frac{1}{2}\,\mathbf{H}\{X_1\} - \frac{1}{2}\,\mathbf{H}\{X_2\}\\ &\qquad - \frac{1}{2}\,\mathbf{d}\{X_1, X_1\} - \frac{1}{2}\,\mathbf{H}\{X_2, X_2\}.\end{aligned}$$

But recall that the second conclusion of Lemma 8.2 was

$$\mathbf{H}\{X_1 + X_2 + \overline{X}_1 + \overline{X}_2\} \le \frac{1}{2}\,\mathbf{H}\{X_1\} + \frac{1}{2}\,\mathbf{H}\{X_2\} + (2 + \eta)k - I_1,$$

meaning that

$$\mathbf{d}\{X_1 + \overline{X}_1; X_2 + \overline{X}_2\} \le (2 + \eta)k - \frac{1}{2}\,\mathbf{d}\{X_1, X_1\} - \frac{1}{2}\,\mathbf{d}\{X_2, X_2\} - I_1.$$

Chaining this with (16) yields

$$k - \frac{\eta}{2}\,\mathbf{d}\{X_1, X_1\} - \frac{\eta}{2}\,\mathbf{d}\{X_2, X_2\} \le (2+\eta)k - \frac{1}{2}\,\mathbf{d}\{X_1, X_1\} - \frac{1}{2}\,\mathbf{d}\{X_2, X_2\} - I_1,$$

which simplifies to

$$\mathbf{d}\{X_1, X_1\} + \mathbf{d}\{X_2, X_2\} \le \frac{2k + 2\eta k - 2I_1}{1-\eta} = 2k + \frac{2(2\eta k - I_1)}{1-\eta}$$

and hence

$$I_2 \le 2\eta k + \frac{2\eta(2\eta k - I_1)}{1-\eta}. \quad \blacksquare$$

9. Endgame

We now approach the phase of the proof which the authors of [2] theatrically call the 'endgame.' There is just one more lemma we need before we can give a full proof of Proposition 6.4.

Lemma 9.1. *Let X_1 and Y_1 be any $\mathbf{F}_2^n$-valued random variables, and let T_1, T_2, and T_3 be $\mathbf{F}_2^n$-valued random variables such that $T_1 + T_2 + T_3 = 0$ holds identically. Putting*

$$\delta = \mathbf{I}\{T_1 : T_2\} + \mathbf{I}\{T_1 : T_3\} + \mathbf{I}\{T_2 : T_3\}$$

and letting ψ be the functional given by

$$\psi\{T_1', T_2'\} = \mathbf{d}\{T_1', T_2'\} + \eta\big(\mathbf{d}\{X_1^*, T_1'\} - \mathbf{d}\{X_1^*, X_1\} + \mathbf{d}\{X_2^*, T_2'\} - \mathbf{d}\{X_2^*, X_2\}\big),$$

there exist random variables T_1' and T_2' such that

$$\psi\{T_1', T_2'\} \le \left(1 + \frac{\eta}{3}\right)\delta + \frac{\eta}{3}\sum_{i=1}^{2}\sum_{j=1}^{3}\big(\mathbf{d}\{X_i^*, T_j\} - \mathbf{d}\{X_i^*, X_i\}\big)$$

Proof. From the fact that any two of the T_j determine the full triple (T_1, T_2, T_3), we have

$$\mathbf{H}\{T_1, T_2\} = \mathbf{H}\{T_1, T_3\} = \mathbf{H}\{T_2, T_3\} = \mathbf{H}\{T_1, T_2, T_3\}.$$

Then since $T_1 + T_2 = T_3$ we may apply Lemma 4.6 with $(X, Y, Z) = (T_1, T_2, T_3)$ to bound

$$\begin{aligned}
\sum_{t_3 \in \mathbf{F}_2^n} &\mathbf{P}\{T_3 = t_3\}\,\mathbf{d}\big\{(T_1 \mid T_3 = t_3); (T_2 \mid T_3 = t_3)\big\} \\
&\le 2\,\mathbf{I}\{T_1 : T_2\} + 2\,\mathbf{H}\{T_3\} - \mathbf{H}\{T_1, T_2\} \\
&= 2\,\mathbf{H}\{T_1\} + 2\,\mathbf{H}\{T_2\} + 2\,\mathbf{H}\{T_3\} - 3\,\mathbf{H}\{T_1, T_2\} \\
&= \mathbf{I}\{T_1 : T_2\} + \mathbf{I}\{T_1 : T_3\} + \mathbf{I}\{T_2 : T_3\} \\
&= \delta.
\end{aligned}$$

Letting Z be a random variable independent of both X_1^* and X_2^*, we apply Lemma 2.5 with $(X, Z, Y, W) = (X_1^*, Z, T_1, T_3)$ to obtain

$$\begin{aligned}\sum_{t_3 \in \mathbf{F}_2^n} \mathbf{P}\{T_3 = t_3\}\big(\mathbf{d}\big\{X_1^*; (T_1 \mid T_3 = t_3)\big\} - \mathbf{d}\{X_1^*, X_1\}\big)& \\ = \mathbf{d}\{X_1^*, T_1 \mid T_3\} - \mathbf{d}\{X_1^*, X_1\}& \\ \le \mathbf{d}\{X_1^*, T_1\} + \frac{1}{2}\mathbf{I}\{T_1 : T_3\} - \mathbf{d}\{X_1^*, X_1\},&\end{aligned}$$

and applying the same lemma with $(X, Z, Y, W) = (X_2^*, Z, T_2, T_3)$ yields

$$\begin{aligned}\sum_{t_3 \in \mathbf{F}_2^n} \mathbf{P}\{T_3 = t_3\}\big(\mathbf{d}\big\{X_2^*; (T_2 \mid T_3 = t_3)\big\} - \mathbf{d}\{X_2^*, X_2\}\big)& \\ \le \mathbf{d}\{X_2^*, T_2\} + \frac{1}{2}\mathbf{I}\{T_2 : T_3\} - \mathbf{d}\{X_2^*, X_2\}.&\end{aligned}$$

Putting the three observations together, we have

$$\begin{aligned}&\sum_{t_3 \in \mathbf{F}_2^n} \mathbf{P}\{T_3 = t_3\}\psi\big\{(T_1 \mid T_3 = t_3); (T_2 \mid T_3 = t_3)\big\} \\ &\quad = \sum_{t_3 \in \mathbf{F}_2^n} \mathbf{P}\{T_3 = t_3\}\Big(\mathbf{d}\big\{(T_1 \mid T_3 = t_3); (T_2 \mid T_3 = t_3)\big\} \\ &\qquad\qquad + \eta\,\mathbf{d}\big\{X_1^*; (T_1 \mid T_3 = t_3)\big\} + \eta\,\mathbf{d}\big\{X_2^*; (T_2 \mid T_3 = t_3)\big\} \\ &\qquad\qquad - \eta\,\mathbf{d}\{X_1^*, X_1\} - \eta\,\mathbf{d}\{X_2^*, X_2\}\Big) \\ &\quad \le \delta + \eta\big(\mathbf{d}\{X_1^*, T_1\} - \mathbf{d}\{X_1^*, X_1\} + \mathbf{d}\{X_2^*, T_2\} - \mathbf{d}\{X_2^*, X_2\} \\ &\qquad\qquad + \frac{1}{2}\mathbf{I}\{T_1 : T_3\} + \frac{1}{2}\mathbf{I}\{T_2 : T_3\}\big),\end{aligned}$$

Choosing some t_3 in the support of T_3 such that the value of ψ inside the sum is minimal and then setting $T'_{1,3} = (T_1 \mid T_3 = t_3)$ and $T'_{2,3} = (T_2 \mid T_3 = t_3)$, we have

$$\begin{aligned}\psi\big\{T'_{1,3}, T'_{2,3}\big\} \le \delta + \eta\big(\mathbf{d}\{X_1^*, T_1\} - \mathbf{d}\{X_1^*, X_1\} + \mathbf{d}\{X_2^*, T_2\} - \mathbf{d}\{X_2^*, X_2\}& \\ + \frac{1}{2}\mathbf{I}\{T_1 : T_3\} + \frac{1}{2}\mathbf{I}\{T_2 : T_3\}\big)&\end{aligned}$$

We can now repeat this for all six permutations (α, β, γ) of $(1, 2, 3)$ and average the corresponding bounds for $(T_\alpha, T_\beta, T_\gamma)$ to obtain

$$\begin{aligned}\frac{1}{6}\sum_{(\alpha,\beta,\gamma) \in \mathfrak{S}_3} \psi\big\{T'_{\alpha,\gamma}, T'_{\beta,\gamma}\big\} \le \delta - \eta\,\mathbf{d}\{X_1^*, X_1\} - \eta\,\mathbf{d}\{X_2^*, X_2\}& \\ + \frac{\eta}{6}\sum_{(\alpha,\beta,\gamma) \in \mathfrak{S}_3} \big(\mathbf{d}\{X_1^*, T_\alpha\} + \mathbf{d}\{X_2^*, T_\beta\}& \\ + \frac{1}{2}\mathbf{I}\{T_\alpha : T_\gamma\} + \frac{1}{2}\mathbf{I}\{T_\beta : T_\gamma\}\big).&\end{aligned}$$

Since each of $\{1,2,3\}$ appears twice as α and twice as β, we have

$$\sum_{(\alpha,\beta,\gamma)\in\mathfrak{S}_3}\big(\mathbf{d}\{X_1^*,T_\alpha\}+\mathbf{d}\{X_2^*,T_\beta\}\big)=2\sum_{i=1}^{2}\sum_{j=1}^{3}\mathbf{d}\{X_i^*,T_j\}$$

and

$$\sum_{(\alpha,\beta,\gamma)\in\mathfrak{S}_3}\left(\frac{1}{2}\,\mathbf{I}\{T_\alpha:T_\gamma\}+\frac{1}{2}\,\mathbf{I}\{T_\beta:T_\gamma\}\right)=2\delta.$$

Consequently, the average of $\psi\big\{T'_{\alpha,\gamma},T'_{\beta,\gamma}\big\}$ over all permutations (α,β,γ) is bounded above by

$$\left(1+\frac{\eta}{3}\right)\delta-\eta\,\mathbf{d}\{X_1^*,X_1\}-\eta\,\mathbf{d}\{X_2^*,X_2\}+\frac{\eta}{3}\sum_{i=1}^{2}\sum_{j=1}^{3}\mathbf{d}\{X_i^*,T_j\},$$

so the result follows by letting (T'_1,T'_2) equal the pair $(T'_{\alpha,\gamma},T'_{\beta,\gamma})$ for the choice of (α,β,γ) that minimises $\psi\big\{T'_{\alpha,\gamma},T'_{\beta,\gamma}\big\}$. ∎

We are now able to prove Proposition 6.4, which we shall restate in the contrapositive. Just as in previous sections, the functional τ depends on the fixed random variables X_1^* and X_2^* as well as the choice of constant $\eta=1/9$.

Proposition 6.4′. *Let X_1 and X_2 be $\mathbf{F}_2^n$-valued random variables with the property that $\tau\{X'_1,X'_2\}\geq\tau\{X_1,X_2\}$ for all random variables X'_1 and X'_2 on $\mathbf{F}_2^n$. Then $\mathbf{d}\{X_1,X_2\}=0$.*

Proof. Let $k=\mathbf{d}\{X_1,X_2\}$. Let X_1 and $\overline{X}_1$ be copies of X_1 and X_2 and $\overline{X}_2$ copies of X_2 such that X_1, $\overline{X}_1$, X_2, and $\overline{X}_2$ are all independent. Let I_1 and I_2 be as in the previous section, and let

$$I_3=\mathbf{I}\{\overline{X}_1+X_2:X_1+\overline{X}_1\mid X_1+X_2+\overline{X}_1+\overline{X}_2\},$$

so that $I_3=I_2$ (which follows from interchanging the rôles of X_1 and $\overline{X}_1$). For brevity of notation, let

$$U=X_1+X_2,\qquad V=\overline{X}_1+X_2,\qquad W=X_1+\overline{X}_1,$$

and

$$S=X_1+X_2+\overline{X}_1+\overline{X}_2.$$

Then

$$I_1=\mathbf{I}\{U:V\mid S\},\qquad I_2=\mathbf{I}\{W:U\mid S\},\qquad\text{and}\qquad I_3=\mathbf{I}\{V:W\mid S\}.$$

Lemma 8.2 gave us

$$I_3=I_2\leq 2\eta k+\frac{2\eta(2\eta k-I_1)}{1-\eta},$$

so setting

$$\overline{\delta} = \mathbf{I}\{U : V \mid S\} + \mathbf{I}\{W : U \mid S\} + \mathbf{I}\{V : W \mid S\},$$

we have

$$\begin{aligned}
\overline{\delta} &\le I_1 + 4\eta k + \frac{4\eta(2\eta k - I_1)}{1-\eta} \\
&= \frac{I_1 - \eta I_n + 4\eta k - 4\eta^2 k + 8\eta^2 k - 4\eta I_1}{1-\eta} \\
&= 6\eta k + \frac{(1-5\eta)I_1 + 4\eta k + 4\eta^2 k - 6\eta k + 6\eta^2 k}{1-\eta} \qquad (19) \\
&= 6\eta k + \frac{(1-5\eta)I_1 - 2\eta k + 10\eta^2 k}{1-\eta} \\
&= 6\eta k - \frac{1-5\eta}{1-\eta}(2\eta k - I_1)
\end{aligned}$$

Now we perform invocations of Lemma 3.6 to obtain bounds on various Ruzsa distances. Setting $(X, Y, Z, W) = (X_1^*, X_1, X_2, \overline{X}_1 + \overline{X}_2)$, gives

$$\begin{aligned}
\mathbf{d}\{X_1^*; U \mid S\} - \mathbf{d}\{X_1^*, X_1\} &\le \frac{1}{2}\big(\mathbf{H}\{S\} + \mathbf{H}\{U\} - \mathbf{H}\{X_1\} - \mathbf{H}\{\overline{X}_1 + \overline{X}_2\}\big) \\
&= \frac{1}{2}\big(\mathbf{H}\{S\} - \mathbf{H}\{X_1\}\big),
\end{aligned}$$

where in the second line we used the fact that

$$\mathbf{H}\{U\} = \mathbf{H}\{X_1 + X_2\} = \mathbf{H}\{\overline{X}_1 + \overline{X}_2\}.$$

Similarly, we have

$$\begin{aligned}
\mathbf{d}\{X_2^*; U \mid S\} - \mathbf{d}\{X_2^*, X_2\} &\le \frac{1}{2}\big(\mathbf{H}\{S\} - \mathbf{H}\{X_2\}\big), \\
\mathbf{d}\{X_1^*; V \mid S\} - \mathbf{d}\{X_1^*, X_1\} &\le \frac{1}{2}\big(\mathbf{H}\{S\} - \mathbf{H}\{X_1\}\big), \\
\mathbf{d}\{X_2^*; V \mid S\} - \mathbf{d}\{X_2^*, X_2\} &\le \frac{1}{2}\big(\mathbf{H}\{S\} - \mathbf{H}\{X_2\}\big),
\end{aligned}$$

and

$$\mathbf{d}\{X_1^*, W \mid S\} - \mathbf{d}\{X_1^*, X_1\} \le \frac{1}{2}\big(\mathbf{H}\{S\} + \mathbf{H}\{W\} - \mathbf{H}\{X_1\} - \mathbf{H}\{W'\}\big),$$

where $W' = X_2 + \overline{X}_2$. To address the asymmetry in the last bound, we note that for any fixed value s taken by S, we have $W' = W + s$, so $\mathbf{d}\{X_2^*; W \mid S\} = \mathbf{d}\{X_2^*; W' \mid S\}$. Now applying Lemma 3.6 to W' yields

$$\mathbf{d}\{X_2^*; W \mid S\} - \mathbf{d}\{X_2^*, X_2\} \le \frac{1}{2}\big(\mathbf{H}\{S\} + \mathbf{H}\{W\} - \mathbf{H}\{X_2\} - \mathbf{H}\{W'\}\big).$$

The sum of these six inequalities is

$$\sum_{i=1}^{2} \sum_{A \in \{U,V,W\}} \big(\mathbf{d}\{X_i^*; A|S\} - \mathbf{d}\{X_i^*, X_i\} \big) \le 3\,\mathbf{H}\{S\} - \frac{3}{2}\,\mathbf{H}\{X_1\} - \frac{3}{2}\,\mathbf{H}\{X_2\}. \quad (20)$$

The second part of Lemma 8.2 states that

$$\mathbf{H}\{S\} \le \frac{1}{2}\,\mathbf{H}\{X_1\} + \frac{1}{2}\,\mathbf{H}\{X_2\} + (2+\eta)k - I_1,$$

which can be plugged into (20) to give

$$\begin{aligned} \sum_{i=1}^{2} \sum_{A \in \{U,V,W\}} \big(\mathbf{d}\{X_i^*; A \mid S\} - \mathbf{d}\{X_i^*, X_i\} \big) &\le 3(2+\eta)k - 3I_1 \\ &\le 6k - 3\eta k + 6\eta k - 3I_1 \\ &= (6-3\eta)k + 3(2\eta k - I_1). \end{aligned} \quad (21)$$

Let ψ be defined as in Lemma 9.1. By Lemma 8.1, $k \le \psi\{T_1', T_2'\}$ for all random variables T_1' and T_2', so for any random variables (T_1, T_2, T_3), the pair (T_1', T_2') furnished by Lemma 9.1 satisfies

$$k \le \psi\{T_1', T_2'\} \le \left(1 + \frac{\eta}{3}\right)\delta + \frac{\eta}{3} \sum_{i=1}^{2} \sum_{j=1}^{3} \big(\mathbf{d}\{X_i^*, T_j\} - \mathbf{d}\{X_i^*, X_i\} \big).$$

In particular, since $(U + V + W \mid S = s)$ is identically zero for all possible values s of S, we can set

$$T_1 = (U \mid S = s), \qquad T_2 = (V \mid S = s), \qquad \text{and} \qquad T_3 = (W \mid S = s)$$

and then average the above inequality over all $s \in \mathbf{F}_2^n$, weighted by $\mathbf{P}\{S = s\}$ to obtain

$$k \le \left(1 + \frac{\eta}{3}\right)\overline{\delta} + \frac{\eta}{3} \sum_{i=1}^{2} \sum_{A \in \{U,V,W\}} \big(\mathbf{d}\{X_i^*, A \mid S\} - \mathbf{d}\{X_i^*, X_i\} \big).$$

We now substitute the upper bounds (19) and (21) to get

$$\begin{aligned} k &\le \left(1 + \frac{\eta}{3}\right)\left(6\eta k - \frac{1-5\eta}{1-\eta}(2\eta k - I_1)\right) + \frac{\eta}{3}\Big((6-3\eta)k + 3(2\eta k - I_1)\Big) \\ &= (8\eta - \eta^2)k + \left(\eta - \left(1 + \frac{\eta}{3}\right)\frac{1-5\eta}{1-\eta}\right)(2\eta k - I_1). \end{aligned}$$

With the choice $\eta = 1/9$,

$$\eta - \Big(1 + \frac{\eta}{3}\Big)\frac{1-5\eta}{1-\eta} = -\frac{11}{27} \le 0$$

and $2\eta k - I_1 \ge 0$ by Lemma 8.2, hence one concludes that

$$k \le (8\eta + \eta^2)k = \frac{73}{81}k,$$

which is nonsense unless $k = 0$. ∎

10. Sum-product phenomena

We prove the following lemma.

Lemma 10.1. *Let X and Y be random variables taking values in $\mathbf{F}_p^*$ for some prime p. There exists $u \in \mathbf{F}_p^*$ such that*

$$\mathbf{H}\{X, uY\} \ge \frac{\mathbf{H}\{X,Y\}}{2^{\mathbf{H}\{X,Y\}}} - 1,$$

Proof. Let U be a random variable uniform on $\mathbf{F}_p^*$ and independent of both X and Y. Independence gives us the identity $\mathbf{H}\{X + UY\} = \mathbf{H}\{U\} = \log(p-1)$. We let (X_1, Y_1), (X_2, Y_2), and U be conditionally independent trials of (X, Y), (X, Y), and U respectively, relative to $X + UY$. Put

$$H = \mathbf{H}\{X_1, Y_1, X_2, Y_2, U\}.$$

Let $W = 1$ if $X_1 = X_2$ and $W = 0$ otherwise. Then since $X_1 + UY_1 = X + UY$, we have

$$\begin{aligned} H &= \mathbf{H}\{X_1, Y_1, X_2, Y_2, W, X + UY\} \\ &= \mathbf{H}\{U\} + \mathbf{H}\{W\} + \mathbf{H}\{X_1, Y_1, X_2, Y_2 \mid W, X + UY\}. \end{aligned}$$

If $X_1 = X_2$, then $Y_1 = Y_2$, so by conditional independence,

$$\begin{aligned} H &= \mathbf{H}\{U\} + \mathbf{H}\{W\} + \mathbf{P}\{X_1 = X_2 \mid X + UY\}\,\mathbf{H}\{X_1, Y_1 \mid X + UY\} \\ &\qquad + \mathbf{P}\{X_1 \ne X_2 \mid X + UY\}\,\mathbf{H}\{X_1, Y_1, X_2, Y_2 \mid X + UY\}. \\ &= \mathbf{H}\{U\} + \mathbf{H}\{W\} + \mathbf{P}\{X_1 = X_2\}\big(\mathbf{H}\{X, Y \mid X + UY\}\big) \\ &\qquad + \mathbf{P}\{X_1 \ne X_2 \mid X + UY\}\big(2\,\mathbf{H}\{X, Y \mid X + UY\}\big) \\ &= \mathbf{H}\{U\} + \mathbf{H}\{W\} + \mathbf{H}\{X, Y \mid X + UY\} \\ &\qquad + \mathbf{P}\{X_1 \ne X_2 \mid X + UY\}\,\mathbf{H}\{X, Y \mid X + UY\} \end{aligned}$$

But

$$\mathbf{H}\{X, Y \mid X + UY\} = \mathbf{H}\{X, Y, X + UY\} - \mathbf{H}\{X + UY\} = \mathbf{H}\{X, Y\}$$

by independence of (X, Y) and U, so

$$H = \big(1 + \mathbf{P}\{X_1 \neq X_2 \mid X + UY\}\big)\, \mathbf{H}\{X, Y\} + \mathbf{H}\{U\} + \mathbf{H}\{W\}.$$

On the other hand, by invertibility of each $u \in \mathbf{F}_p^*$, we have

$$\begin{aligned} H - \mathbf{H}\{U\} &= \mathbf{H}\{X_1, Y_1, X_2, Y_2 \mid U\} \\ &= \frac{1}{p-1} \sum_{u \in \mathbf{F}_p^*} \mathbf{H}\{X_1, Y_1, X_2, Y_2 \mid U = u\} \\ &= \frac{1}{p-1} \sum_{u \in \mathbf{F}_p^*} \mathbf{H}\{X_1, uY_1, X_2, uY_2\} \\ &= \frac{1}{p-1} \sum_{u \in \mathbf{F}_p^*} \mathbf{A}\{X, uY\}. \end{aligned}$$

Consequently there exists $u \in \mathbf{F}_p^*$ such that

$$\begin{aligned} \mathbf{A}\{X, uY\} &\leq H - \mathbf{H}\{U\} \\ &= \big(1 + \mathbf{P}\{X_1 \neq X_2 \mid X + UY\}\big)\, \mathbf{H}\{X, Y\} + \mathbf{H}\{W\} \\ &\leq \big(1 + \mathbf{P}\{X_1 \neq X_2 \mid X + UY\}\big)\, \mathbf{H}\{X, Y\} + 1, \end{aligned}$$

and for this choice of u we have

$$\mathbf{H}\{X + uY\} \geq \mathbf{P}\{X_1 = X_2 \mid X + UY\}\, \mathbf{H}\{X, Y\} - 1.$$

But

$$\mathbf{P}\{X_1 = X_2 \mid X + UY\} = \mathbf{P}\big\{(X_1, Y_1) = (X_2, Y_2) \mid X + UY\big\},$$

and by [TODO: add lemma], we have

$$\begin{aligned} \mathbf{P}\{(X_1, Y_1) = (X_2, Y_2) \mid X + UY\} &\geq 2^{-\mathbf{H}\{X, Y \mid X + UY\}} \\ &= 2^{-\mathbf{H}\{X, Y\}}, \end{aligned}$$

which immediately yields the desired inequality. ∎

References

[1] Imre Zoltán Ruzsa, "Sums of finite sets," *Number Theory: New York Seminar 1991–1995* (1996).

[2] William Timothy Gowers, Ben Green, Freddie Manners, and Terence Tao, "On a conjecture of Marton," *arXiv:2311.05762* (2023), 33 pp.

[3] Ben Green, Freddie Manners, and Terence Tao, "Sumsets and entropy revisited," *arXiv:2306.13403* (2023), 37 pp.

[4] Imre Zoltán Ruzsa, "Sumsets and entropy," *Random Structures and Algorithms* **34** (2008), 1–10.

[5] Terence Tao, "Sumset and inverse sumset theory for Shannon entropy," *Combinatorics, Probability, and Computing* **19** (2010), 603–639.