A Note on Sumsets using Entropy

Adam Marcus* Prasad Tetali†

revised: April 2008

Gyarmati, Matolcsi, and Ruzsa recently noted [2] that Han-type inequalities can be applied to sumsets in much the same way that they can be applied to characteristic functions of sets of random variables (the usual situation). While it is not true (in general) that sumsets satisfy a log-submodular relation in an obvious way, it is natural to ask whether they permit a weaker property, by way of fractional subadditivity. It is classical (and recently reviewed in [4]) that fractional subadditivity is weaker than log-submodularity and more general than Han's inequalities. Here, we extend an argument of [2] (more precisely, an idea in the proof of Theorem 1.2 in their paper), make further use of entropy, and show general fractional subadditivity properties for sumsets that imply some of the results and conjectures in[2] as easy corollaries.

It should be noted that some of the ideas in this paper were discovered independently by Balister and Bollobás [1] who, following the works of [2] and [4], also developed a hierarchy of entropy inequalities. This paper, however, constains some new ideas that can be used to extend the results in [1] beyond sumsets in ways that were not considered in that paper.

1 Definitions

The important property of sumsets that we wish to exploit is that, for a fixed element a, the sum $a+b$ depends only on b (no further knowledge about how a and b relate is needed). This idea leads to a more general class of functions.

Definition 1. Let $X_1, X_2, \ldots, X_k$ be finite sets. Any subset $S \subset [k]$ corresponds to a different product space $X_S = \prod_{i \in S} X_i$. For sets $S \subseteq T \subseteq [k]$, we define the projection function $\pi_S : X_T \to X_S$ in the natural way: for $x \in T$, let $\pi_S(x) = (x_{i_1}, \ldots, x_{i_{|S|}})$ where $i_j \in S$. When the meaning is clear, we will write $\pi_i(x)$ for $\pi_{\{i\}}(x)$.

We will denote $Q(X_1, X_2, \ldots, X_k)$ to be the space that is a disjoint union of each of the spaces X_S, for $S \subseteq [k]$. Formally,

$$Q(X_1, X_2, \ldots, X_k) = \bigcup_{S \in [k]} \left\{ (x_{i_1}, \ldots, x_{i_{|S|}}) : x_i \in X_i, S = \{i_1, \ldots, i_{|S|}\} \right\}$$

Let Y be any space and $f : Q(X_1, \ldots, X_k) \to Y$ be any function. Then, for a set $S \subset [k]$, we define $f_S : X_S \to Y$ to be the restriction of f to only those inputs that came from X_S. We will abuse notation by writing, for $S \subseteq T$ and $x \in T$, $f_S(x)$ to mean $f_S(\pi_S(x))$.

*School of Mathematics, Georgia Tech, Atlanta, GA 30332-0160

†School of Mathematics and School of Computer Science, Georgia Tech, Atlanta, GA 30332-0160; research supported in part by NSF grant DMS-0701043.

Let f be such a function and let $\mathcal{G}$ be a collection of subsets of $[k]$. We will say that f is *deterministic with respect to* $\mathcal{G}$ if for all $S \in \mathcal{G}$ and for all $x, y \in X_{[k]}$ we have that $f(x) = f(y)$ whenever both $f_S(x) = f_S(y)$ and $f_{\overline{S}}(x) = f_{\overline{S}}(y)$ (here, $\overline{S} = [k] \setminus S$).

In essence the definition above is designed to capture the property of sumsets that was mentioned earlier. For a function f to be deterministic with respect to a single set $S \subset [k]$, it must be that $f_S(x)$ and $f_{\overline{S}}(x)$ uniquely determine the value of $f(x)$. Then, being deterministic with respect to a collection $\mathcal{G}$, is nothing more than being deterministic with respect to all $G \in \mathcal{G}$. The following examples show that both Cartesian products of sets and linear combinations of sets (and so, in particular, sumsets) are deterministic with respect to $\mathcal{G}$ for any $\mathcal{G}$.

Example 1. *Let V be a vector space over the reals with basis vectors $\{v_1, \ldots, v_k\}$. Let $X_1, \ldots, X_k \subseteq \mathbb{R}$ and define $f : Q(X_1, \ldots, X_k) \to V$ such that $f_S(x) = \sum_{i \in S} \pi_i(x) v_i$. Then f is deterministic with respect to $\mathcal{G}$ for all $\mathcal{G} \subseteq [k]$.*

Proof. Let $x \in X_T$ for some $T \subseteq [k]$ and let $G \in \mathcal{G}$. Then

$$f(x) = \sum_{i \in T} \pi_i(x) v_i = \sum_{i \in (G \cap T)} \pi_i(x) v_i + \sum_{i \in (\overline{G} \cap T)} \pi_i(x) v_i = f_G(x) + f_{\overline{G}}(x).$$

Thus knowing $f_G(x)$ and $f_{\overline{G}}(x)$ uniquely determines $f(x)$. Since this is true for any $G \in \mathcal{G}$, f is deterministic with respect to $\mathcal{G}$. □

Example 2. *Let $\mathcal{A}$ be an Abelian group and $X_1, \ldots, X_k \subseteq \mathcal{A}$ and let $c_1, \ldots, c_k \in \mathbb{Z}$. Define $f : Q(X_1, \ldots, X_k) \to \mathcal{A}$ such that $f_S(x) = \sum_{i \in S} c_i \pi_i(x)$. Then f is deterministic with respect to $\mathcal{G}$ for all $\mathcal{G} \subseteq [k]$.*

Proof. The proof is identical to Example 1, only replacing c_i with v_i. □

2 Proof

Our goal is to prove Theorem 1, stated below. Note that Example 1, which shows that the characteristic function is deterministic, shows that Theorem 1 is, in fact, a generalization of normal subadditivity.

Theorem 1. *Let $X_1, X_2, \ldots, X_k$ be finite sets, $\mathcal{G} = \{(\alpha_G, G)\}$ be a fractional covering of $[k]$, and f be a function on $Q(X_1, \ldots, X_k)$ that is deterministic with respect to $\mathcal{G}$. Then for any set $R \subseteq f(X_{[k]})$,*

$$|R| \leq \prod_{G \in \mathcal{G}} \left| f_G\left(f^{-1}_{[k]}(R)\right)\right|^{\alpha_G}$$

Proof. We first choose an arbitrary linear order on $\bigcup X_i$. Now let R be given, and for each $r \in R$, let x_r be the first element of $f^{-1}_{[k]}(R) \subseteq X_{[k]}$ in lexicographical order, and let $X^R = \{x_r : r \in R\}$. Let Z be a random variable chosen uniformly from X^R, and for $i \in [k]$, let $Z_i = \pi_i(Z)$. Then by fractional additivity,

$$\log(|R|) = H(Z) = H(Z_1, \ldots, Z_k) \leq \sum_{G \in \mathcal{G}} \alpha_G H(Z_G) \tag{1}$$

where $Z_G = \{Z_{i_1}, \ldots, Z_{i_{|G|}}\}$, where $G = \{i_j\} \subseteq [k]$. Hence $Z_G = \pi_G(Z)$ and by the chain rule of entropy, for each $G \in \mathcal{G}$, we have that:

$$H(\pi_G(Z)|f_G(Z)) + H(f_G(Z)) = H\left(Z_G, f(Z_G)\right) = H(f_G(Z)|\pi_G(Z)) + H(\pi_G(Z)). \tag{2}$$

Here, $H(f_G(Z)|\pi_G(Z)) = 0$ since f_G is a deterministic function, and so plugging in to the above equation gives:

$$\log\left(|R|\right) \leq \sum_{G\in\mathcal{G}} \alpha_G H\left(Z_G\right) = \sum_{G\in\mathcal{G}} \alpha_G \Big(H(f_G(Z)) - H(\pi_G(Z)|f_G(Z)) \Big).$$

Now the key point is the following somewhat surprising claim, whose proof is more or less obvious; this is also the essence (in addition to Han's inequality) of the proof of Theorem 1.2 in [2].
Claim: $H(\pi_G(Z)|f_G(Z)) = 0$ for all $G \in \mathcal{G}$.

Proof. It suffices to show that, for every G, f_G is a one-to-one function when the domain is restricted to $\pi_G(Z)$. Assume not. Then there are two elements $a \neq b \in X_G$ such that $f_G(a) = f_G(b)$ and both $Pr(Z_G = a)$ and $Pr(Z_G = b)$ are non-zero. Thus there must be "preimages" $A, B \in X^R$ such that $\pi_G(A) = a$ and $\pi_G(B) = b$ and $A \neq B$ (since otherwise $a = b$).

Without loss of generality, let $A < B$ in lexicographical order on X_G, and consider $b' = \pi_{\overline{G}}(B) \in X_{\overline{G}}$. Let $A' \in X_{[k]}$ be the vector

$$A'(i) = \begin{cases} a_i & \text{for } i \in G \\ b'(i) & \text{for } i \notin G \end{cases}$$

Clearly $A' < B$ in lexicographical order, and since

$$f_G(A') = f_G(a) = f_G(b) = f_G(B) \quad \text{and} \quad f_{\overline{G}}(A') = f_{\overline{G}}(b') = f_{\overline{G}}(B)$$

we have that $f(A') = f(B)$. This is a contradiction, however, since we assumed B to be the smallest such element in $X_{[k]}$ in lexicographical order. □

So Equation (2) reduces to $H(\pi_G(Z)) = H(f_G(Z))$. Plugging this into Equation (1) yields:

$$\log\left(|R|\right) \leq \sum_{G\in\mathcal{G}} \alpha_G H\left(f_G(Z)\right) \leq \sum_{G\in\mathcal{G}} \alpha_G \log\left(\left|f_G\left(X^R\right)\right|\right) \leq \sum_{G\in\mathcal{G}} \alpha_G \log\left(\left|f_G(f_{[k]}^{-1}(R))\right|\right)$$

where the last inequality is due to the fact that $X^R \subseteq f_{[k]}^{-1}(R))$, and so our claimed result is true. □

3 Corollaries

The following corollaries are a slight variant on Theorem 1 in that they do not directly define R; rather, they pick a subset S of the image of one of the subspaces, then lift the preimage of S up to the top space.

Corollary 1. *Let $A, B_1, B_2, \ldots, B_k \subseteq \mathbb{R}$ and define $\overline{B_i} = B_1 + \ldots + B_{i-1} + B_{i+1} + \ldots + B_k$ for $i = 1, \ldots, k$ and $B = B_1 + \ldots + B_k$. Then for any $S \subseteq B$, we have that*

$$|A + S|^k \leq |S| \prod_{i=1}^{k} |A + \overline{B_i}|$$

Proof. Set $X_1 = A$ and $X_i = B_{i-1}$ for $i = 2, \dots, k+1$, and $X = X_1, \dots, X_{k+1}$. By Example 2, the collection of functions $f_S(x) = \sum_{i \in S} x_i$ is deterministic with respect to any $\mathcal{G}$. Note that $S \subseteq f_{\overline{\{1\}}}(X)$, so let $Q = f^{-1}_{\overline{\{1\}}}(S)$, and set $R = \{f(a, b_1, \dots, b_k)\}$ for all $a \in A$ and all $(b_1, \dots, b_k) \in Q$. Now choose

$$\mathcal{G} = \left\{ \left(\frac{1}{k}, \overline{\{i\}} \right) \right\}_{i \in [k+1]}.$$

By Theorem 1, we have that

$$|R| \leq \prod_{G \in \mathcal{G}} \left| f_G \left(f^{-1}_{[k]}(R) \right) \right|^{\alpha_G}$$

and we use that

$$f_{\overline{\{1\}}} \left(f^{-1}_{[k]}(R) \right) = S \qquad \text{and} \qquad f_{\overline{\{i\}}} \left(f^{-1}_{[k]}(R) \right) \subseteq A + \overline{B_{i-1}}$$

for $i = 2, \dots, k+1$. The inequality follows. □

Corollary 2. *Let $A, B_1, B_2, \dots, B_k \subseteq \mathbb{R}$. Then for any $S \subseteq B$, we have that*

$$|A + S|^k \leq |S|^{k-1} \prod_{i=1}^{k} |A + B_i|$$

Proof. The proof is the same as Corollary 2, just with a different covering:

$$\mathcal{G} = \left\{ \left(\frac{1}{k}, \{1, i\} \right) \right\}_{i \in [k+1]}.$$

□

Remark. Clearly, various other covering families yield similar corollaries. We mention these only because they offer direct generalizations of Theorem 1.5 in [2] and are independent of the results in [1].

4 Further Research and Acknowledgments

While the results in this chapter provide an entire collection of inequalities for sumsets, a number of known inequalities are not known to be implied by this method (see [5]). It would be interesting to see if these inequalities can be deduced by the fractional subadditivity (or other properties) of entropy. In particular, recent work of Madiman [3] derived sumset inequalities from the *entropy power inequality*. Another possible direction of research would be to consider similar results in nonabelian groups. Our results do not immediately extend to nonabelian groups, but perhaps a more thorough analysis would give similar results.

We thank Imre Ruzsa for sharing the preprint [2], for helpful discussions, and finally for informing us of the independent and recent work of Balister–Bollobás. We also thank Béla Bollobás for promptly sending us the preprint [1], which contains results of independent interest.

References

[1] P. Balister, B. Bollobás, Projections, Entropy, and Sumsets, preprint (October 2007).

[2] K. Gyarmati, M. Matolcsi, and I. Ruzsa, A Superadditivity and Submultiplicativity Property for Cardinalities of Sumsets, preprint (June 2007).

[3] M. Madiman, On the Entropy of Sums, in *Proc. 2008 IEEE Information Theory Workshop*, (2008).

[4] M. Madiman, P. Tetali, Information Inequalities for Joint Distribution of Random Variables: Interpretations and Applications, submitted (May 2007).

[5] M. Nathanson, *Additive Number Theory: Inverse Problems and Geometry of Sumsets*, No. 165 in Graduate Texts in Mathematics, Springer, 1996.