FRINGE PAIRS IN GENERALIZED MSTD SETS

MEGUMI ASADA, SARAH MANSKI, STEVEN J. MILLER, AND HONG SUH

Abstract. A *More Sums Than Differences* (MSTD) set is a set A for which $|A+A| > |A-A|$. Martin and O'Bryant proved that the proportion of MSTD sets in $\{0, 1, \dots, n\}$ is bounded below by a positive number as n goes to infinity. Iyer, Lazarev, Miller and Zhang introduced the notion of a *generalized MSTD set*, a set A for which $|sA - dA| > |\sigma A - \delta A|$ for a prescribed $s + d = \sigma + \delta$. We offer efficient constructions of k-generational MSTD sets, sets A where $A, A + A, \dots, kA$ are all MSTD. We also offer an alternative proof that the proportion of sets A for which $|sA-dA|-|\sigma A-\delta A| = x$ is positive, for any $x \in \mathbb{Z}$. We prove that for any $\epsilon > 0$, $\Pr(1-\epsilon < \log|sA-dA|/\log|\sigma A-\delta A| < 1+\epsilon)$ goes to 1 as the size of A goes to infinity and we give a set A which has the current highest value of $\log|A+A|/\log|A-A|$. We also study decompositions of intervals $\{0, 1, \dots, n\}$ into MSTD sets and prove that a positive proportion of decompositions into two sets have the property that both sets are MSTD.

Contents

1. Introduction 2
2. Background 4
2.1. Definitions 4
2.2. Important Characteristics of Rich Sets 5
2.3. Minimal Fringe Pairs 6
3. Efficient constructions of k-generational sets and their generalizations 7
3.1. Previous Constructions 7
3.2. Proof of Theorem 1.1 7
3.3. Super k-generational MSTD Sets 9
4. Arbitrary differences 10
4.1. Preliminaries 10
4.2. Proof of Theorem 1.2 11
5. Another MSTD construction and $\frac{\log|sA-dA|}{\log|\sigma A-\delta A|}$ 13
5.1. New Construction and Large Ratio 13
5.2. Proof of Theorem 1.3 15
6. Bi-MSTD Sets 17
6.1. Definitions 17

Date: October 17, 2016.

2010 *Mathematics Subject Classification.* 11P99 (primary), 11K99 (secondary).

Key words and phrases. Sum-dominant sets, generalized MSTD sets, k-generational sets, arbitrary differences, fringe pairs.

Our research was conducted as part of the SMALL 2015 REU program. The first, second, and fourth named authors were supported by Williams College and NSF grant DMS-1347804. The third named author was supported by NSF grant DMS-1265673. The first named author was also supported by a Clare Boothe Luce scholarship. We thank Gal Gross and Kevin O'Bryant for many enlightening conversations and the referee for constructive comments.

6.2. Proof of Theorem 1.4 18
7. Future work 18
References 19

1. Introduction

A *More Sums Than Differences* (MSTD) set is a set A where the sumset of A, denoted $A+A = \{a_1+a_2 : a_1, a_2 \in A\}$, has a greater cardinality than the difference set, $A-A = \{a_1-a_2 : a_1, a_2 \in A\}$. Since addition is commutative while subtraction is not, it is reasonable to expect the difference set to be larger than the sumset in most cases; such sets are said to be difference-dominated, while sets where $|A+A| = |A-A|$ are called sum-difference balanced. Interestingly, while most sets of $\{0, 1, \dots, n\}$ are difference dominated as $n \to \infty$, Martin and O'Bryant [MO] proved that a positive percentage *are* sum-dominated. They predict an expected limiting density of MSTD sets of about 4.5×10^{-4}; the best lower bound at the time of writing is due to Zhao [Zh1]: 4.28×10^{-4}.

Conway found the first example of an MSTD set

$$\{0, 2, 3, 4, 7, 11, 12, 14\}$$

in the 1960's. Hegarty later proved that there were no MSTD sets of smaller cardinality [He]. While some work has been generalized to MSTD sets in abelian groups [MV, Na4, Zh3], we concern ourselves with MSTD sets A with $A \subset \mathbb{Z}$. Some progress has been made in constructing infinite families of such sets [Na1, Na2, Na3, Na4, Zh1, He, MOS, MS, Ru1, Ru2], as well as MSTD sets with additional properties. One example is due to Iyer, Lazarev, Miller and Zhang [ILMZ], who developed constructions for *generalized MSTD sets*, sets A that satisfy $|sA - dA| > |\sigma A - \delta A|$ for a given $s + d = \sigma + \delta$. This provided them with the framework to construct k-generational sets, a rarer class of MSTD sets A for which $A, A+A, \dots, kA$ are all sum-dominant.

A unifying strategy used throughout this paper is the manipulation of *fringe pairs*, a concept introduced by Zhao [Zh2] (see Definition 2.1). Our first set of results concern fringe constructions for generalized MSTD sets that are much more efficient than the ones in [ILMZ]. Before stating these we first set some notation:

- $|A|$ is the cardinality of A,
- $[a, b] = \{a, a+1, \dots, b\}$,
- $mA = \{\sum_{i=1}^{m} a_i : a_i \in A\}$ for $m \geq 1$ (note this is *not* m times each element of A),
- $m \cdot A = \{ma : a \in A\}$,
- $-A = \{-a : a \in A\}$, $-mA = -(mA)$,
- $A^c = [\min A, \max A] \setminus A$.

Our first result significantly improves upon the constructions of k-generational sets given by Iyer et al. [ILMZ], which used base expansion and therefore yielded sets A which grow astronomically with k.

Theorem 1.1. *For any $q > 2$, there exists a set A with $|A| = O(q^2)$ such that for all $s + d = \sigma + \delta$ with $s > \sigma$,*

$$|sA - dA| > |\sigma A - \delta A|. \tag{1.1}$$

Further, given $k > 0$ there exists a k-generational set A with $|A| = O(k)$.

We prove Theorem 1.1 in §3.2. Our constructions generate explicit k-generational sets that are much smaller than those previously constructed, making them far easier to manipulate and verify. We then apply the fringe pairs framework to give a new proof of the following result in §4.2.

Theorem 1.2. *For any $x \in \mathbb{Z}$ and $0 \le d < \delta \le \sigma < s$ with $s + d = \sigma + \delta$, the proportion of $A \subseteq [0, n]$ such that $|sA - dA| - |\sigma A - \delta A| = x$ is bounded below by a positive number as n goes to infinity.*

We also study the values of $\log|sA - dA|/\log|\sigma A - \delta A|$. These are natural objects to study as they normalize the excess of one generalized sum or difference set relative to the number of elements of that set. In other words, if we want to construct a set with many more sums than differences we don't want to do so through base expansion and taking exponentially large sets; we want to construct small sets with the desired excess. The quantity $\log|A+A|/\log|A-A|$ was studied by Ruzsa in [Ru2], where he proved that there are a "multitude" of sets with $\log|A + A|/\log|A - A| > 1 + c$ with $c > 0$. Precisely, he proved that there exists $c > 0$ such that for every sufficiently large n there exists a set such that

$$|A| \;=\; n, \quad |A - A| \;\le\; n^{2-c}, \quad |A + A| \;\ge\; n^2/2 - n^{2-c},$$

or in other words, $\log|A+A|/\log|A-A| > 1+c_1$ where $c_1 > 0$. We show that for any $\epsilon > 0$ there are not enough sets $A \subseteq [0, n]$ with $\log|sA - dA|/\log|\sigma A - \delta A| > 1 + \epsilon$ to constitute a positive proportion. In fact, we prove a stronger statement about $|sA - dA|/|\sigma A - \delta A|$ in §5.2.

Theorem 1.3. *Choosing subsets $A \subset [0, n]$ uniformly, for every $\epsilon > 0$ we have*

$$\lim_{n\to\infty} \Pr\left(1 - \epsilon < \frac{\log|sA - dA|}{\log|\sigma A - \delta A|} < 1 + \epsilon\right) = 1. \tag{1.2}$$

In fact,

$$\lim_{n\to\infty} \Pr\left(\frac{s + d - 1}{s + d} < \frac{|sA - dA|}{|\sigma A - \delta A|} < \frac{s + d}{s + d - 1}\right) \;=\; 1. \tag{1.3}$$

We also construct a set A for which

$$\frac{\log|A + A|}{\log|A - A|} \;=\; \frac{\log 892}{\log 765} \;=\; 1.02313, \tag{1.4}$$

which is larger than the previous largest value of

$$\frac{\log|A + A|}{\log|A - A|} \;=\; \frac{\log 91}{\log 83} \;=\; 1.0208 \tag{1.5}$$

found in [He, MO].

Finally, we investigate decompositions of the interval $[0, n]$ into two disjoint MSTD sets in §6. It turns out that a positive proportion of decompositions of $[0, n]$ into two disjoint sets have the property that both components are MSTD. If both $A \subseteq [0, n]$ and $[0, n] \setminus A$ are MSTD, we say that A is *bi-MSTD*.

Theorem 1.4. *Let A be a uniform random subset of $[0, n]$ with $n \ge 19$. Then the probability that A is bi-MSTD is bounded below by a positive number (and there are no bi-MSTD sets for $n < 19$).*

The proof of Theorem 1.4 is given in §6.2.

The paper is organized as follows. We define fringe pairs and prove some needed properties in §2. We then prove our results on efficient k-generational constructions in §3, on arbitrary differences in §4, on the ratio of the logarithms in §5 and bi-MSTD sets in §6. We conclude in §7 with topics for future investigations.

2. Background

2.1. **Definitions.** Martin and O'Bryant used the idea that the fringes of A essentially determine whether or not A is MSTD, as the middle does not contribute as much to $|A+A|-|A-A|$ since almost all middle elements of the sum and difference set are attained [MO]. This is because the number of elements of a set A chosen uniformly among subsets of $[0,n]$ is tightly concentrated around $(n+1)/2$, and there are so many ways of writing middle numbers as a sum or difference that with high probability all are realized. Zhao built on these ideas to prove that the proportion of MSTD sets in $[0,n]$ converges as n goes to infinity. We extend many of Zhao's definitions and results in [Zh2] to the setting of generalized MSTD sets. Of particular interest are Definition 2.1 and Definition 2.3, which define *fringe pairs* and *rich sets* for generalized MSTD sets. The results in this section will prove valuable in the proofs of our main results in Sections 3.2, 4.2, and 6.2. In the following, we fix $s+d=\sigma+\delta$.

Definition 2.1. *Let s,d,σ,δ be non-negative integers such that $0\le d<\delta\le\sigma<s$ and $s+d=\sigma+\delta$. A generalized MSTD fringe pair of order k for $(s,d),(\sigma,\delta)$ is a pair $(L,R;k)$ such that $L,R\subseteq[0,k]$ with $0\in L,R$ and*

$$\begin{aligned}&|(sL+dR)\cap[0,k]|+|(sR+dL)\cap[0,k]|\\ &\quad>\ |(\sigma L+\delta R)\cap[0,k]|+|(\sigma R+\delta L)\cap[0,k]|.\end{aligned}\tag{2.1}$$

We impose an order on the set of generalized MSTD fringe pairs by having $(L,R;k)\le(L',R';k')$ if $k\le k'$ and

$$L\ =\ L'\cap[0,k],\qquad R=R'\cap[0,k],\qquad [k+1,k']\subseteq sL'+dR',sR'+dL'.\tag{2.2}$$

$(L,R;k)<(L',R';k')$ is defined as expected with $k<k'$ as well as the stipulations listed in (2.1).

Definition 2.2. *A minimal generalized MSTD fringe pair is a pair $(L,R;k)$ such that $(L,R;k)\le(L',R';k')$ for all generalized MSTD fringe pairs $(L',R';k')$ with $k'\ge k$.*

We later use the partial order of generalized MSTD fringe pairs to reduce the study of all generalized MSTD fringe pairs to simply the minimal ones.

Definition 2.3. *A set $A\subseteq[0,n]$ is a* k-rich set *with MSTD fringe pair $(L,R;k)$ if*

$$\begin{aligned}2k\ <\ n,\qquad A\cap[0,k]\ &=\ L,\qquad A\cap[n-k,n]\ =\ n-R,\\ [k+1,2n-k-1]\ &\subseteq\ A+A.\end{aligned}\tag{2.3}$$

When we do not specify a fringe pair and say A is k-rich, we often simply mean $[k+1,2n-k-1]\subseteq A+A$. Sometimes we simply say that A is rich if k is clear from the context.

2.2. **Important Characteristics of Rich Sets.** It turns out that a rich set with generalized MSTD fringe pair $(L, R; k)$ is a generalized MSTD with respect to (s, d), (σ, δ) (Lemma 2.5). We first prove a simple lemma.

Lemma 2.4. *A k-rich set $A \subseteq [0, n]$ has the property*

$$[-dn + k + 1, sn - k - 1] \subseteq sA - dA \tag{2.4}$$

for any $s + d > 2$.

Proof. Suppose without loss of generality that $s \geq d$. For $d \geq s$, simply switch the roles of s and d. Since $s + d > 2$ and $s \geq d$, we have $s \geq 2$. Thus $sA - dA = 2A + (s-2)A - dA$. Then

$$\begin{aligned} 2A + (s-2)A - dA &\supseteq [k+1, 2n-k-1] + (s-2)A - dA \\ &\supseteq [k+1, 2n-k-1] + (s-2)\{0, n\} - d\{0, n\} \\ &= [-dn + k + 1, sn - k - 1] \end{aligned} \tag{2.5}$$

since $[k+1, 2n-k-1] \cup [n+k+1, 3n-k-1] \cup \{2n\} = [k+1, 3n-k-1]$ and so on. □

Lemma 2.5. *Let $0 \leq d < \delta \leq \sigma < s$ and $s + d = \sigma + \delta$. A rich set $A \subseteq [0, n]$ with generalized MSTD fringe pair $(L, R; k)$ satisfies $|sA - dA| > |\sigma A - \delta A|$.*

Proof. We split $sA - dA$ into two portions, the fringe and the middle. We assert the following two statements, which are sufficient to prove the lemma:

$$\begin{aligned} &|(sA - dA) \cap ([-dn, -dn + k] \cup [sn - k, sn])| \\ &\qquad > |(\sigma A - \delta A) \cap ([-\delta n, -\delta n + k] \cup [\sigma n - k, \sigma n])| \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} &|(sA - dA) \cap [-dn + k + 1, sn - k - 1]| \\ &\qquad \geq |(\sigma A - \delta A) \cap [-\delta n + k + 1, \sigma n - k - 1]|. \end{aligned} \tag{2.7}$$

Since A is a rich set, $(sA - dA) \supseteq [-dn + k + 1, sn - k - 1]$. Thus (2.7) follows. To show (2.6), we notice that

$$\begin{aligned} &|(sA - dA) \cap ([-dn, -dn + k] \cup [sn - k, sn])| \\ &\qquad = |(sA - dA) \cap [-dn, -dn + k]| + |(sA - dA) \cap [sn - k, sn]|. \end{aligned} \tag{2.8}$$

Since $A \subseteq [0, n]$ has generalized MSTD fringe pair $(L, R; k)$, we have $L = A \cap [0, k]$ and $R = n - (A \cap [n - k, n])$. Using $\cong$ to denote equivalence under translation and dilation, we find

$$\begin{aligned} (sL + dR) \cap [0, k] &= (s(A \cap [0, k]) + dn - d(A \cap [n - k, n])) \cap [0, k] \\ &\cong (s(A \cap [0, k]) - d(A \cap [n - k, n])) \cap [-dn, -dn + k] \\ &= (sA - dA) \cap [-dn, -dn + k]. \end{aligned} \tag{2.9}$$

Similarly, we have

$$\begin{aligned} (sR + dL) \cap [0, k] &= (sn - s(A \cap [n - k, n]) + d(A \cap [0, k])) \cap [0, k] \\ &\cong (s(A \cap [n - k, n]) - d(A \cap [0, k])) \cap [sn - k, sn] \\ &= (sA - dA) \cap [sn - k, sn]. \end{aligned} \tag{2.10}$$

Now (2.6) follows from the definition of a generalized MSTD fringe pair. □

2.3. **Minimal Fringe Pairs.** Now we show that the order we imposed on fringe pairs earlier is completely determined by k, k' for fringe pairs $(L, R; k)$ and $(L', R'; k')$ corresponding to a rich set $A \subseteq [0, n]$.

Lemma 2.6. *Let $A \subseteq [0, n]$ be a rich set. Let $(L, R; k)$ and $(L', R'; k')$ be two generalized MSTD fringe pairs corresponding to A. If $k = k'$, then $(L, R; k) = (L', R'; k')$. If $k < k'$, then $(L, R; k) < (L', R'; k')$.*

Proof. The first statement follows trivially. If $k < k'$, then we need to show

$$L = L' \cap [0, k], \quad R = R' \cap [0, k], \quad [k + 1, k'] \subseteq sL' + dR', sR' + dL'.$$

We have $A \cap [0, k] = L, A \cap [0, k'] = L'$, and the analogous statements for R and R'. Since $k < k'$, $L = A \cap [0, k] = (A \cap [0, k']) \cap [0, k] = L' \cap [0, k]$. The same holds for R and R'. Finally, we have

$$[-dn + k' + 1, sn - k' - 1] \subseteq [-dn + k + 1, sn - k - 1] \subseteq sA - dA.$$

Thus

$$[-dn + k + 1, -dn + k'] \subseteq sA - dA, \quad [sn - k - 1, sn - k'] \subseteq sA - dA.$$

An argument from the proof of Lemma 2.5 shows that $[-dn + k + 1, -dn + k'] \subseteq sA - dA \iff [k + 1, k'] \subseteq sL + dR$ and $[sn - k - 1, sn - k'] \subseteq sA - dA \iff [k + 1, k'] \subseteq sR + dL$. Therefore $(L, R; k) < (L', R'; k')$. □

Lemma 2.7. *Let $(L, R; k)$ be the minimal generalized MSTD fringe pair of a rich set $A \subseteq [0, n]$. Then $(L, R; k)$ is minimal in the partial ordering of all generalized MSTD fringe pairs. Also, for every $k < k' < n/2$, $(L', R'; k')$ is also a generalized MSTD fringe pair of A, where $L' = A \cap [0, k'], R' = (n - A) \cap [0, k']$, and every generalized MSTD fringe pair of A has this form.*

Proof. Suppose $(L, R; k)$ is not minimal. Then there is a generalized MSTD fringe pair $(L', R'; k') < (L, R; k)$. So $k' < k$ and

$$L' = L \cap [0, k'], \quad R' = R \cap [0, k'], \quad [k + 1, k'] \subseteq sL' + dR', sR' + dL'.$$

This implies that $[k+1, 2n-k-1] \subseteq A+A$ and therefore A is rich with generalized MSTD fringe pair $(L', R'; k')$ as well, which contradicts the minimality of $(L, R; k)$ when attached to A.

So $(L, R; k)$ is the minimal generalized MSTD fringe pair of A. Take $k < k'$. Then the only possible generalized MSTD fringe pair is $(L', R'; k')$ where

$$L' = A \cap [0, k'], \quad R' = A \cap [0, k'].$$

By the previous lemma, $(L', R'; k') > (L, R; k)$. A simple computation confirms that $(L', R'; k')$ is also a generalized MSTD fringe pair. □

The above results show that we may count all generalized MSTD sets by counting all minimal generalized MSTD fringe pairs.

Remark 2.8. *Iyer et al. [ILMZ] proved that the proportion of generalized MSTD sets is bounded below by a positive number as n goes to infinity. Zhao's methods should easily generalize to show that this proportion converges. Our definition of a rich set is identical to Zhao's. The only difference is the nature of the fringe pairs, but the fringe pairs are not involved in the proof of convergence.*

*Zhao introduced k-*affluent sets*, sets A for which $[k + 1, 2n - k - 1] \subseteq A + A$ and $[-n + k + 1, n - k - 1] \subseteq A - A$ (see Definition 4.1), in order to show that for*

any $m \in \mathbb{Z}$ the proportion of A which satisfy $|A+A|-|A-A| = m$ converges as $n \to \infty$. The analogous result for generalized MSTD sets does not require affluent sets (we use affluent sets in this paper, albeit for a different purpose). Since we may assume that $s+d = \sigma+\delta \geq 3$, both $sA-dA$ and $\sigma A - \delta A$ contain a copy of $A+A$, up to a minus sign. Thus all the convergence results in [Zh2] should be applicable to the generalized MSTD case. We do not pursue this path as it is too repetitive to merit detailing.

3. Efficient constructions of k-generational sets and their generalizations

3.1. **Previous Constructions.** In [ILMZ], Iyer et al. gave a construction of *k-generational sets*, which are sets A such that $A, A+A, \ldots, kA$ are all MSTD. The primary tool was a technique called base expansion, summarized in the following proposition.

Proposition 3.1 (Iyer et al. [ILMZ])**.** *Fix a positive integer k. Say $A_1, \ldots, A_k \subseteq \mathbb{Z}^+$. Choose some $m > k \cdot \max(\cup_{i=1}^k A_k)$. Let $C = A_1 + m \cdot A_2 + \cdots + m^{k-1} \cdot A_k$, where $\cdot$ denotes scalar multiplication. Then*

$$|sC - dC| \;=\; \prod_{j=1}^{k} |sA_j - dA_j|$$

for all $s+d \leq k$.

Using base expansion, we may choose A_j such that $|jA_j + jA_j| > |jA_j - jA_j|$ (and $|sA_j - dA_j| = |\sigma A_j - \delta A_j|$ for $s+d = \sigma+\delta \neq 2j$) and create the appropriate C prescribed above, which is k-generational. However, a major drawback of base expansion is that the set C grows large very quickly (we explore this issue in greater detail in §5, where we investigate the ratio of the logarithms of the cardinalities). According to the construction in [ILMZ], the middle of A_j has at least $2(2jr-4j+1)$ elements where $r = 4j+2$. Thus $|A_j| = \Omega(j^2)$, which means there is a constant c such that $cj^2 \leq |A_j|$. By the proof of Lemma 4.3 in [ILMZ], we have $|C| = \prod_{j=1}^k |A_j|$. Therefore $|C| = \Omega(k!^2)$.

This is a huge growth rate that leads to sets which are computationally impractical to work with even when $k = 2$. For example, an optimization of the construction with base expansion yields a 5-generational set of $2,685,375$ elements. Our construction reduces this size to 35 (see Proposition 3.4). We are able to create much more reasonably sized sets by choosing appropriate fringe pairs and filling out the middle rather than using base expansion, which has no regard for the size of the set.

3.2. **Proof of Theorem 1.1.** We first prove the existence of a k-generational fringe pair in Proposition 3.3. In Proposition 3.4, we construct a rich set with the k-generational fringe pair from Proposition 3.3, giving us a k-generational set. We note that our rich set varies linearly with k to complete the proof.

Lemma 3.2. *Let $R = [0,m] \cup \{q\}$. Then*

$$kR \;=\; [0,km] \cup [q, q+(k-1)m] \cup \cdots \cup [(k-1)q, (k-1)q+m] \cup \{kq\}.$$

Further, if $q > km$, then

$$|kR| \;=\; \frac{(mk+2)(k+1)}{2}.$$

Proof. Observe that

$$k(A \cup B) \;=\; kA \cup ((k-1)A + B) \cup \cdots \cup (A + (k-1)B) \cup kB.$$

Set $A = [0, m], B = \{q\}$ to obtain the result. Note then if $km < q$, $|kR| = \frac{(k)(k+1)m}{2} + k + 1 = \frac{(mk+2)(k+1)}{2}$. □

Proposition 3.3. *Let $L = \{0\}, R = \{0, 1, 3\}$. Then the fringe pair $(L, R, 6k)$ is k-generational in the sense that for all $1 \le j \le k$,*

$$|(2j)L \cap [0, 6k]| + |(2j)R \cap [0, 6k]| \;>\; 2|(jL + jR) \cap [0, 6k]|.$$

Proof. Fix $1 \le j \le k$. In Lemma 3.2 set $m = 1, q = 3, k = 2j$ to get $(2j)R = [0, 6j] \setminus \{6j - 1\}$. Also, $jL + jR = [0, 3j] \setminus \{3j - 1\}$. Therefore

$$|(2j)L \cap [0, 6k]| + |(2j)R \cap [0, 6k]| \;=\; 1 + 6j, \quad 2|(jL + jR) \cap [0, 6k]| = 2 \cdot 3j. \;\square$$

By Lemma 2.5, we may pick any rich set A with fringe pair $(\{0\}, \{0, 1, 3\}; 6k)$ and A is k-generational. In Proposition 3.4, we explicitly construct such a set.

Proposition 3.4. *Let $A = \{0\} \cup [6k + 1, 12k + 1] \cup (18k + 2 - \{0, 1, 3\})$. Then A is a rich set with k-generational fringe pair $(L, R, 6k)$ and thus k-generational.*

Proof. It suffices to show that $[6k+1, 30k+3] \subseteq A+A$. Note that $|[6k+1, 12k+1]| = 6k + 1$. Compute

$$\begin{aligned} A + A \;&\supseteq\; [6k + 1, 12k + 1] \cup 2[6k + 1, 12k + 1] \cup ([6k + 1, 12k + 1] + \{18k + 2\}) \\ &=\; [6k + 1, 30k + 3]. \end{aligned}$$

□

Since A has $6k+5$ elements, the growth of the size of these k-generational sets A is $|A| = O(k)$. This is a significant improvement over the previous best construction of k-generational sets, which have size $|A| = \Omega(k!^2)$.

Remark 3.5. *This construction provides us with a nice 2-generational MSTD set with 17 elements,*

$$\{0, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 35, 37, 38\},$$

as well as a 3-generational set of size 23,

$$\{0, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 53, 55, 56\}.$$

Remark 3.6. *One advantage of the base expansion method is that for any given sequence $\{(s_i, d_i, \sigma_i, \delta_i)\}_{i=2}^{q}$ with $s_i + d_i = \sigma_i + \delta_i = i$ and $s_i \ne \sigma_i$, $s_i \ne \delta_i$, the base expansion method can construct a set A which satisfies $|s_iA - d_iA| > |\sigma_iA - \delta_iA|$. Our construction is specific to k-generational sets. Though we do not reach the full level of generality achieved by Iyer et al., for any given sequence $\{(s_i, d_i, \sigma_i, \delta_i)\}_{i=2}^{q}$ with $s_i + d_i = \sigma_i + \delta_i = i$ and $0 \le d_i < \delta_i \le \sigma_i < s_i$, we describe an efficient construction of a set A which satisfies $|s_iA - d_iA| > |\sigma_iA - \delta_iA|$ in §3.3.*

3.3. **Super k-generational MSTD Sets.** We prove the existence of a stronger form of k-generational sets, *super k-generational sets*, in Proposition 3.9.

Definition 3.7. *A* super k-generational MSTD sets *is a set A in which for all $s+d=\sigma+\delta\le k$ with $0\le d<\delta\le\sigma<s$, $|sA-dA|>|\sigma A-\delta A|$.*

Lemma 3.8. *Consider the set $\{0,1,q\}$ with $q>2$. We have $k\{0,1,q\}=[0,k]\cup[q,q+k-1]\cup[2q,2q+k-2]\cup\cdots\cup\{kq\}$.*

Proof. This is a consequence of Lemma 3.2. □

Note that if $k<q$, then $|k\{0,1,q\}|=\frac{(k+1)(k+2)}{2}$. Similarly if $k=q$ then $|k\{0,1,q\}|=\frac{(k+1)(k+2)}{2}-1$.

Proposition 3.9. *Let $L=\{0\},R=\{0,1,q\},k=q^2$ where $q>2$. Then for any $s+d=\sigma+\delta\le q$ with $0\le d<\delta\le\sigma<s$, $(L,R;k)$ is a generalized MSTD fringe pair.*

Proof.
Case 1: Let $d=0$. Observe that $sL=\{0\},sR=[0,s]\cup[q,q+s-1]\cup\cdots\cup\{sq\}$. Also $\sigma L+\delta R=[0,\delta]\cup[q,q+\delta-1]\cup\cdots\cup\{\delta q\}$ and $\delta L+\sigma R=[0,\sigma]\cup[q,q+\sigma-1]\cup\cdots\cup\{\sigma q\}$. Therefore

$$|(sL+dR)\cap[0,k]|+|(sR+dL)\cap[0,k]|\ \ge\ 1+\frac{(s+1)(s+2)}{2}-1$$

and

$$|(\sigma L+\delta R)\cap[0,k]|+|(\sigma R+\delta L)\cap[0,k]|\ \le\ \frac{(\sigma+1)(\sigma+2)}{2}+\frac{(\delta+1)(\delta+2)}{2}.$$

Since $s>\sigma$ and σ,δ are nonzero, it is clear that $s^2>\sigma^2+\delta^2$ (remember $d=0$ in this case, so $s=\sigma+\delta$). Thus

$$\begin{aligned}|(sL+dR)\cap[0,k]|+|(sR+dL)\cap[0,k]|&\\ >\ |(\sigma L+\delta R)\cap[0,k]|+|(\sigma R+\delta L)\cap[0,k]|,&\end{aligned}$$

and $(L,R;k)$ is a generalized MSTD fringe pair.

Case 2: Let $d\neq 0$. Observe that

$$s_iL+d_iR\ =\ [0,d_i]\cup[q,q+d_i-1]\cup\cdots\cup\{d_iq\}$$

and

$$d_iL+s_iR\ =\ [0,s_i]\cup[q,q+s_i-1]\cup\cdots\cup\{s_iq\}.$$

Then

$$|(s_iL+d_iR)\cap[0,k]|+|(s_iR+d_iL)\cap[0,k]|\ =\ \frac{(s_i+1)(s_i+2)}{2}+\frac{(d_i+1)(d_i+2)}{2}.$$

We have $s^2+d^2>\sigma^2+\delta^2$; this follows from $0\le d<\delta\le\sigma<s$ (thus $2\sigma\delta>2sd$, and the claim follows from combining that inequality with $(s+d)^2=(\sigma+\delta)^2$). Thus

$$\begin{aligned}|(sL+dR)\cap[0,k]|+|(sR+dL)\cap[0,k]|&\\ >\ |(\sigma L+\delta R)\cap[0,k]|+|(\sigma R+\delta L)\cap[0,k]|.&\end{aligned}$$

and $(L,R;k)$ is a generalized MSTD fringe pair. □

The existence of such a fringe pair gives an infinite family of super k-generational MSTD sets and proves their positive density as n approaches infinity. For the sake of completeness, we give one possible construction of a super k-generational set from a super k-generational fringe pair.

Proposition 3.10. *Let $A = \{0\} \cup [k+1, 2k+2] \cup (3k+3 - \{0, 1, q\})$ with $q > 2$ and $k = q^2$. Then A is a rich set with fringe pair (L, R, k) for all $s + d = \sigma + \delta \leq q$ and $|A| = q^2 + 6$.*

Proof. It suffices to show $[k+1, 5k+5] \subseteq A + A$, which follows from

$$\begin{aligned} A + A &\supseteq [k+1, 2k+2] \cup 2[k+1, 2k+2] \cup ([k+1, 2k+2] + \{3k+3\}) \\ &= [k+1, 5k+5]. \end{aligned}$$

□

Remark 3.11. *This construction gives a super 4-generational set with 22 elements:*

$$\{0, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 47, 50, 51\}$$

4. Arbitrary differences

We now turn our attention to a simple construction for attaining specific differences between $|A + A|$ and $|A - A|$, and more generally $|sA - dA|$ and $|\sigma A - \delta A|$. Though Martin and O'Bryant and Iyer et al. have already proved that for any integer m, a positive percentage of sets have the property that $|A+A| - |A-A| = m$ and a positive percentage of sets have the property that $|sA - dA| - |\sigma A - \delta A| = m$ respectively, we include the proof to advocate for the loose notion that fringe pairs are a clean perspective with which to think about arbitrary differences. We first give necessary definitions for our construction before proving Theorem 1.2 in §4.2.

4.1. Preliminaries.

Definition 4.1. *Let k and n be positive integers with $2k < n$. Let $A \subseteq [0, n]$. We say A is* k-affluent *with generalized MSTD fringe pair $(L, R; k)$ if $[k+1, 2n-k-1] \subseteq A + A$ and $[-n+k+1, n-k-1] \subseteq A - A$.*

Note that an affluent set has all middle sums and differences present; therefore, discrepancies in numbers of sums and differences are completely determined by the fringes.

Proposition 4.2. *Given $m > 0$, $L = \{0\}$, $R = [0, m] \cup \{q\}$, and $k = 2q$ with $q > 2m$, if A is k-affluent with fringe pair $(L, R; k)$, then $|A + A| - |A - A| = m$.*

Proof. Observe that $|2L \cap [0, k]| = 1$ and $|2R \cap [0, k]| = 3m + 3$ by Lemma 3.2. Furthermore, $|(L+R) \cap [0, k]| = m+2$. It follows that $|2L \cap [0, k]| + |2R \cap [0, k]| - 2|(L + R) \cap [0, k]| = m$. Since A is affluent, the difference between sums and differences is completely determined by the fringes. Thus $|A + A| - |A - A| = m$. □

Proposition 4.3. *Given $m > 0$, $L = [0, m]$, $R = [0, m] \cup \{q\}$ and $k = 2q$ with $q > 2m$, if A is k-affluent with fringe pair $(L, R; k)$, then $|A + A| - |A - A| = -m$.*

Proof. Observe that $|2L \cap [0, k]| = 2m + 1$ and $|2R \cap [0, k]| = 3m + 3$ by Lemma 3.2. Furthermore, $(L + R) \cap [0, k] = [0, 2m] \cup [q, q + m]$ and has cardinality $3m + 2$. It follows that $2|(L + R) \cap [0, k]| - (|2L \cap [0, k]| + |2R \cap [0, k]|) = m$. Since A is affluent, the difference between sums and differences is completely determined by the fringes. Thus $|A + A| - |A - A| = -m$. □

Another advantage of this simple proof is that it is easily generalizable to the case of $s+d=\sigma+\delta$. This construction can be extended to attain $|sA-dA|-|\sigma A-\delta A| = (\sigma\delta - sd)m$ by again using Lemma 3.2. Then some further work shows that any arbitrary difference can be attained with a positive proportion as n goes to infinity. The idea of the proof is quite simple: for positive differences $|sA-dA|-|\sigma A-\delta A| = x \geq 0$, we take the same fringe as in Proposition 4.2, $L=\{0\}, R=[0,m]\cup\{q\}$ with $m \geq x$. Then from $k=(s+d)q$, we simply decrease k until $(L,R;k)$ has the property

$$\begin{aligned}|(sL+dR)\cap[0,k]| \ &+ \ |(sR+dL)\cap[0,k]| \\ &- \ |(\sigma L+\delta R)\cap[0,k]| \ - \ |(\sigma R+\delta L)\cap[0,k]| \ = \ x.\end{aligned}$$

This quantity is ultimately reflected in $|sA-dA|-|\sigma A-\delta A|$ if A is k-affluent. The negative differences case is similar but for the left fringe, which we take from Proposition 4.3 to be $L=[0,m]$.

4.2. **Proof of Theorem 1.2.** We divide our proof into two theorems proving the positive and negative cases.

Theorem 4.4. *Let $x \geq 0$ be a nonnegative integer and $s+d=\sigma+\delta$ with $d<\delta\leq\sigma<s$. Then the proportion of $A\subseteq[0,n]$ satisfying $|sA-dA|-|\sigma A-\delta A|=x$ is bounded below by a positive number as n goes to infinity.*

Proof. Let $L=\{0\}$ and $R=[0,m]\cup\{q\}$ with $q>(s+d)m$. Then by Lemma 3.2,

$$cR \ = \ [0,cm]\cup[q,q+(c-1)m]\cup\cdots\cup[(c-1)q,(c-1)q+m]\cup\{cq\},$$

where we define $0R$ to be $\{0\}$. Notice that if $c<\gamma$, then $cR\subseteq\gamma R$. Now since $d<\delta\leq\sigma<s$, we have the relation

$$dR \ \subseteq \ \delta R \ \subseteq \ \sigma R \ \subseteq \ sR. \tag{4.1}$$

Now define the quantity

$$f(k) \ = \ |dR\cap[0,k]|+|sR\cap[0,k]|-|\delta R\cap[0,k]|-|\sigma R\cap[0,k]|.$$

Since $L=\{0\}$, the quantity $f(k)$ is also equal to

$$\begin{aligned}|(sL+dR)\cap[0,k]| + \ |(sR+dL)\cap[0,k]| \\ - |(\sigma L+\delta R)\cap[0,k]| \ - \ |(\sigma R+\delta L)\cap[0,k]|.\end{aligned}$$

We have the following implications:

(1) $k\in dR \implies f(k)-f(k-1) \ = \ 0,$
(2) $k\in\delta R, \notin dR \implies f(k)-f(k-1) \ = \ -1,$
(3) $k\in\sigma R, \notin \delta R \implies f(k)-f(k-1) \ = \ 0,$
(4) $k\in sR, \notin \sigma R \implies f(k)-f(k-1) \ = \ 1,$
(5) $k\notin sR \implies f(k)-f(k-1) \ = \ 0.$

Therefore, $f(k)$ and $f(k-1)$ differ at most by one. Since $|cR|=(mc+2)(c+1)/2$,

$$\begin{aligned}f(sq) \ &= \ \frac{(md+2)(d+1)}{2}+\frac{(ms+2)(s+1)}{2}-\frac{(m\delta+2)(\delta+1)}{2}-\frac{(m\sigma+2)(\sigma+1)}{2} \\ &= \ \frac{1}{2}m(d(d+1)+s(s+1)-\delta(\delta+1)-\sigma(\sigma+1)) \\ &= \ \frac{1}{2}(d^2+s^2-\delta^2-\sigma^2)m \\ &= \ (\sigma\delta-sd)m.\end{aligned}$$

Since $0 \le d < \delta \le \sigma < s$ implies $\sigma\delta - sd \ge 1$, $f(sq) \ge m$.

Now we want an integer k for which $f(k) = 0$. For any $c > 0$, $cR \cap [0, q-1] = [0, cm]$. Therefore, $f(q-1) = dm + 1 + sm + 1 - \delta m - 1 - \sigma m - 1 = 0$.

We have the facts that $|f(k) - f(k-1)| \le 1$, $f(q-1) = 0$, and $f(sq) = (\sigma\delta - sd)m \ge m$. Therefore, for any given difference x such that $0 \le x \le (\sigma\delta - sd)m$, there exists a k such that $q - 1 \le k \le sq$ and $f(k) = x$.

Pick an arbitrary $x \ge 0$ and let $m \ge x$. Then let $q - 1$ be large enough so that for a sufficiently large n, the proportion of q-affluent sets in $[0, n]$ is positive (see Lemma 5.4). Consider the fringe pair $(L, R; k)$ where $q - 1 \le k \le sq$ and $f(k) = x$. Then the proportion of affluent sets with fringe pair $(L, R; k)$ is positive and bounded below as n goes to infinity. □

Similarly, we can prove that for any $x < 0$, a positive proportion of sets have $|sA - dA| - |\sigma A - \delta A| = x$. However, we must use a different fringe pair.

Lemma 4.5. *Let $0 \le d < \delta \le \sigma < s$ and $s + d = \sigma + \delta$ and $L = [0, m]$, $R = [0, m] \cup \{q\}$ where $q > (s + d)m$. Then*

$$\begin{aligned} sL + dR = {} & [0, (s+d)m] \cup [q, q + (s+d-1)m] \cup \cdots \cup \\ & [(d-1)q, (d-1)q + (s+1)m] \cup [dq, dq + sm] \end{aligned}$$

with

$$|sL + dR| + |dL + sR| - |\sigma L + \delta R| - |\delta L + \sigma R| = (sd - \sigma\delta)m \le -m.$$

Furthermore,

$$sL + dR \subseteq \sigma L + \delta R \subseteq \delta L + \sigma R \subseteq dL + sR.$$

Proof. The first claim is a routine calculation. The set $sL + dR$ can be written as

$$[0, sm] + ([0, dm] \cup [q, q + (d-1)m] \cup \cdots \cup [(d-1)q, (d-1)q + m] \cup \{dq\}),$$

which can be simplified to the desired expression. Next, we calculate $|sL + dR| + |dL + sR| - |\sigma L + \delta R| - |\delta L + \sigma R|$.

$$\begin{aligned} |sL + dR| &= ((s+d)m + 1) + ((s+d-1)m + 1) + \cdots + (sm + 1) \\ &= \frac{1}{2}((s+d)m + 2)(s + d + 1) - \frac{1}{2}((s-1)m + 2)s. \end{aligned}$$

Therefore $|sL + dR| + |dL + sR| = ((s+d)m + 2)(s + d + 1) - \frac{1}{2}((s-1)m + 2)s - \frac{1}{2}((d-1)m + 2)d$. Similarly, $|\sigma L + \delta R| + |\delta L + \sigma R| = ((\sigma + \delta)m + 2)(\sigma + \delta + 1) - \frac{1}{2}((\sigma - 1)m + 2)\sigma - \frac{1}{2}((\delta - 1)m + 2)\delta$. Since $s + d = \sigma + \delta$,

$$\begin{aligned} g((s+d)q) &= \frac{1}{2}\left(((\sigma - 1)m + 2)\sigma + ((\delta - 1)m + 2)\delta \right. \\ &\quad \left. -((s-1)m + 2)s - ((d-1)m + 2)d\right) \\ &= \frac{1}{2}(\sigma^2 + \delta^2 - s^2 - d^2)m \\ &= (sd - \sigma\delta)m \le -m. \end{aligned}$$

For the last claim, notice that $\sigma L + \delta R = (sL + dR) \cup [(d+1)q, (d+1)q + (s-1)m] \cup \cdots \cup [\delta q, \delta q + \sigma m]$. So $\sigma < s$ implies that $\sigma L + \delta R \supseteq sL + dR$. The chain of inclusions follows. □

Theorem 4.6. *Let $x < 0$ be a negative integer and $s + d = \sigma + \delta$ with $0 \le d < \delta \le \sigma < s$. Then the proportion of $A \subseteq [0, n]$ satisfying $|sA - dA| - |\sigma A - \delta A| = x$ is bounded below by a positive number as n goes to infinity.*

Proof. Let $L = [0, m]$ and $R = [0, m] \cup \{q\}$ with $q > (s + d)m$. Define a similar quantity

$$g(k) \;=\; |(sL+dR)\cap[0,k]|+|(dL+sR)\cap[0,k]|-|(\sigma L+\delta R)\cap[0,k]|-|(\delta L+\sigma R)\cap[0,k]|.$$

It suffices to show three things: $|g(k)-g(k-1)| \le 1$, $g(q-1) = 0$, and $g((s+d)q) = (sd - \sigma\delta)m \le -m$. We begin by claiming that $|g(k) - g(k-1)| \le 1$.

By Lemma 4.5, the chain of inclusions is

$$sL + dR \;\subseteq\; \sigma L + \delta R \;\subseteq\; \delta L + \sigma R \;\subseteq\; dL + sR,$$

and we have the implications

(1) $k \in sL + dR \implies g(k) - g(k-1) \;=\; 0$,
(2) $k \in \sigma L + \delta R, \notin sL + dR \implies g(k) - g(k-1) \;=\; -1$,
(3) $k \in \delta L + \sigma R, \notin \sigma L + \delta R \implies g(k) - g(k-1) \;=\; 0$,
(4) $k \in dL + sR, \notin \delta L + \sigma R \implies g(k) - g(k-1) \;=\; 1$,
(5) $k \notin dL + sR \implies g(k) - g(k-1) \;=\; 0$.

Now consider $g((s + d)q)$. By Lemma 4.5, $g((s + d)q) \;=\; (sd - \sigma\delta)m \le -m$.

Finally, consider $g(q - 1)$. The intersection of each set $(sL + dR)$, $(\sigma L + \delta R)$, $(\delta L + \sigma R)$, $(dL + sR)$ with $[0,\ q-1]$ is $[0,\ (s+d)m]$. Therefore $g((s+d)m) \;=\; 0$. The rest of the proof is identical to the positive arbitrary differences proof. □

Together, Theorems 4.4 and 4.6 imply Theorem 1.2.

5. Another MSTD construction and $\frac{\log|sA-dA|}{\log|\sigma A-\delta A|}$

5.1. **New Construction and Large Ratio.** We have used the fringe $L = \{0\}, R = [0, m] \cup \{q\}$ for various m and q in the previous sections. The choice $L = \{0\}$ has a nice property: if we fix the left fringe to be the singleton $\{0\}$, almost any right fringe gives an MSTD fringe pair. If we let R be a uniform random subset and k a suitably large integer (say $2 \cdot \max R$), then it is likely that $|(R+R)\cap[0,k]|+1 > 2|R\cap[0,k]|$, in which case we have an MSTD fringe pair. We make this statement precise below.

Proposition 5.1. *Let R be a uniform random subset of $[0, r]$ with $r \ge 1$. Then*

$$\Pr(|R + R| + 1 > 2|R|) \;>\; 1 - \frac{(r+1)^3 - (r+1)^2}{2^{r+2}}.$$

Proof. Recall that $|R+R| = 2|R| - 1$ if and only if R is an arithmetic progression, and $|R+R| > 2|R|-1$ otherwise. The proof is straightforward. Write the elements of R as $x_1 < x_2 < \cdots < x_r$ (with $r = |R|$); as the claim is trivial for arithmetic progressions, we may assume R is not an arithmetic progression. We are left with proving that we cannot have $|R + R| = 2|R| - 1$ for such R. We proceed by contradiction. Note $x_1 + x_1 < \cdots < x_1 + x_r < x_2 + x_r < x_3 + x_r < \cdots < x_{r-1} + x_r < x_r + x_r$. We have just listed $2|R| - 1$ distinct elements, and thus all other sums must be in this list. In particular, $x_2 + x_{r-1}$ is less than $x_2 + x_r$ but more than $x_1 + x_{r-1}$; it must therefore equal $x_1 + x_r$, which implies $x_2 - x_1 = x_r - x_{r-1}$. We then note $x_2 + x_{r-1} < x_3 + x_{r-1} < x_3 + x_r$, and thus $x_3 + x_{r-1} = x_2 + x_r$, which

implies $x_3 - x_2 = x_r - x_{r-1}$. Arguing similarly shows all adjacent differences are equal, proving the set is an arithmetic progression.

Thus it suffices to bound the number of arithmetic progressions in $[0, r]$. There are $\binom{r+1}{2}$ pairs (i, j) such that $i < j$. For each (i, j), there are at most $\tau(j - i)$ arithmetic progressions starting with i and ending with j, where $\tau(x)$ is the number of divisors of x. Since $\tau(j - i) \le r + 1$, we have that there are at most $(r + 1)\binom{r+1}{2}$ arithmetic progressions in $[0, r]$. Thus $\Pr(|R+R|+1 > 2|R|) \ge 1-\frac{(r+1)^3-(r+1)^2}{2^{r+2}}$. □

Sending r to infinity yields an almost sure method to construct MSTD sets. One consequence of this construction is a set A for the current largest value of $\log|A + A|/\log|A - A|$. We obtained the following R by a random search through subsets of $[0, 90]$, fixing 0 and picking each element with probability 0.27. We then sifted through the set to add or discard obvious elements, finding

$$R = \{0, 1, 2, 4, 5, 9, 10, 12, 23, 26, 32, 38, 47, 53, 59, 61, 65, 76, 78, 79, 81, 85, 86, 88, 89\}. \quad (5.1)$$

Let $A = \{0\} \cup [k + 1, n - k - 1] \cup (n - R)$ for $k = 2 \cdot 89$, $n = 3k + 2$. Then we get $\log|A + A|/\log|A - A| = 1.02313$. The previous largest value of $\log|A + A|/\log|A - A|$ was 1.0208, achieved by setting

$$A = \{0, 1, 2, 4, 5, 9, 12, 13, 17, 20, 21, 22, 24, 25, 29, 32, 33, 37, 40, 41, 42, 44, 45\}$$

as found in [MO]. In fact, sets A for which $\log|A + A|/\log|A - A| > 1.0208$ are relatively common: a random search through sets $R \subseteq [0, 90]$, picking each element with probability 0.27 (this value was chosen as it yielded good results from our simulations), yielded 174 such sets out of 100,000. Since R in (5.1) was found basically through a random search, a more sophisticated method may yield a larger value for $\log|A + A|/\log|A - A|$.

It seems unlikely that $\log|A + A|/\log|A - A|$ can exceed 1.1, or even 1.05. Is there a theoretical upper bound? A result due to Ruzsa [Ru1], also shown in [GH], states that for any finite set $A \subseteq \mathbb{Z}$,

$$3/4 \le \frac{\log|A + A|}{\log|A - A|} \le 4/3.$$

The upper bound of 4/3 is still quite far above what has been seen, and it is probably the case that the upper bound is not tight.

Remark 5.2. *Our method yields values*

$$\frac{\log(4k + 4 + |R + R|)}{\log(4k + 3 + 2|R|)} \quad (5.2)$$

where $k = 2 \cdot \max R$. *Since* $R + R \subseteq [0, k]$, *we have*

$$\frac{\log(4k + 4 + |R + R|)}{\log(4k + 3 + 2|R|)} \le \frac{\log(5k + 4)}{\log(4k + 3)} \quad (5.3)$$

which converges to 1 as $k \to \infty$. *Therefore we do not expect to find large values of* $\log|A + A|/\log|A - A|$ *when we pick* R *from* $[0, r]$ *with* r *large. In fact, by* (5.3) *we know that the maximum value of* $\log|A + A|/\log|A - A|$ *found with our method has* $r < 2000$.

The above analysis suggests that the value $\log|A + A|/\log|A - A|$ tends to 1 as the size of A grows. We formalize that statement below.

5.2. **Proof of Theorem 1.3.** Propositions 5.3 and 5.7 together prove the full inequality (equation (1.2)) in Theorem 1.3. Proposition 5.8 proves the second half of the theorem, equation (1.3).

Proposition 5.3. *Let $0 \le d < \delta \le \sigma < s$ with $s + d = \sigma + \delta$. For every $\epsilon > 0$ and a uniform random subset $A \subseteq [0, n]$,*

$$\lim_{n\to\infty} \Pr\left(\frac{\log|sA - dA|}{\log|\sigma A - \delta A|} \ > \ 1 + \epsilon\right) \ = \ 0. \tag{5.4}$$

The idea is that for a set A with generalized MSTD fringe pair $(L, R; k)$, $A + A$ and $A - A$ are almost always full in the middle as n goes to infinity and then k goes to infinity. We must first introduce a lemma from Zhao [Zh2, Lemma 2.13].

Lemma 5.4 (Zhao)**.** *Let $n, \overline{k}$ be positive integers with $n > 2\overline{k}$. Let A be a uniform random subset of $[0, n]$. Then*

$$\Pr([\overline{k} + 1, 2n - \overline{k} - 1] \not\subseteq A + A) \ \le \ \frac{3(3/4)^{\overline{k}/2}}{2 - \sqrt{3}}$$

and

$$\Pr([-n + \overline{k} + 1, n - \overline{k} - 1] \not\subseteq A - A]) \ \le \ 8\left(\frac{3}{4}\right)^{\overline{k}+2} + (n + 1)\left(\frac{3}{4}\right)^{(n-1)/3}.$$

We now set up the proof of Proposition 5.3. Recall that a set A is k-affluent if $[k + 1, 2n - k - 1] \subseteq A + A$ and $[-n + k + 1, n - k - 1] \subseteq A - A$. Fix $\epsilon > 0$ and let

(1)

$$\begin{aligned}\mu_n(L, R; k) \ = \ 2^{-n+1}|\{A \subseteq [0, n] : 0, n \in A,& \\ A \text{ is affluent with fringe } (L, R; k), \tfrac{\log|sA-dA|}{\log|\sigma A-\delta A|} \ > \ 1 + \epsilon\}|,&\end{aligned}$$

(2)

$$\lambda_n \ = \ 2^{-n+1}|\{A \subseteq [0, n] : 0, n \in A, \tfrac{\log|sA-dA|}{\log|\sigma A-\delta A|} \ > \ 1 + \epsilon\}|,$$

(3)

$$\Lambda_n \ = \ 2^{-n-1}|\{A \subseteq [0, n] : \tfrac{\log|sA-dA|}{\log|\sigma A-\delta A|} > 1 + \epsilon\}|.$$

Our final goal is showing that $\lim_{n\to\infty} \Lambda_n = 0$. To show this, we prove that

$$0 \ = \ \sum_{(L,R;k)} \mu_n(L, R; k) \ = \ \lim_{n\to\infty} \lambda_n \ = \ \lim_{n\to\infty} \Lambda_n,$$

where $\sum_{(L,R;k)} \mu_n(L, R; k)$ is summing over all minimal MSTD fringe pairs $(L, R; k)$.

Lemma 5.5. *If $\mu = \lim_{n\to\infty} \sum_{(L,R;k)} \mu_n(L, R; k)$ exists, then $\lim_{n\to\infty} \lambda_n = \mu$.*

Proof. First notice that for any $\overline{k} < n$,

$$\sum_{\overline{k}} \mu_n(L, R; k) \ := \ \sum_{\substack{(L,R;k) \\ k\le\overline{k}}} \mu_n(L, R; k) \ \le \ \lambda_n.$$

We may say this because of Lemma 2.7. We want to bound above the number of sets which are counted in λ_n but not in $\sum_{\overline{k}} \mu_n(L, R; k)$. Suppose $A \subseteq [0, n]$,

$0, n \in A$, and $\log|A+A|/\log|A-A| > 1+\epsilon$ but A is not k-affluent for any $k \le \overline{k}$. Let $L = A \cap [0,\overline{k}], R = (n-A) \cap [0,\overline{k}]$. There are two cases.

In the first case, $(L, R; \overline{k})$ is not a generalized MSTD fringe pair. Since A is generalized MSTD, $A-A$ must be missing at least one element in $[-n+\overline{k}+1, n-\overline{k}-1]$. In the second case, $(L, R; \overline{k})$ is a generalized MSTD fringe pair. Then either $A+A \not\supseteq [\overline{k}+1, 2n-\overline{k}-1]$ or $A-A \not\supseteq [-n+\overline{k}+1, n-\overline{k}-1]$. So the probability that A is counted in λ_n but not in $\sum_{\overline{k}} \mu_n(L,R;k)$ is bounded above by

$$\begin{aligned}\Pr(A+A \not\supseteq [\overline{k}+1, 2n-\overline{k}-1]) + \Pr(A-A \not\supseteq [-n+\overline{k}+1, n-\overline{k}-1]) \\ \le \frac{3(3/4)^{\overline{k}/2}}{2-\sqrt{3}} + 8\left(\frac{3}{4}\right)^{\overline{k}+2} + (n+1)\left(\frac{3}{4}\right)^{(n-1)/3}\end{aligned}$$

by Lemma 5.4. Let n go to infinity to get

$$\begin{aligned}\sum_{\overline{k}} \mu(L,R;k) \le \liminf_{n\to\infty} \lambda_n \le \\ \limsup_{n\to\infty} \lambda_n \le \sum_{\overline{k}} \mu(L,R;k) + \frac{3(3/4)^{\overline{k}/2}}{2-\sqrt{3}} + 8\left(\frac{3}{4}\right)^{\overline{k}+2}.\end{aligned}$$

Let $\overline{k}$ go to infinity to get $\mu = \sum_{(L,R;k)} \mu(L,R;k) = \lim_{n\to\infty} \lambda_n$. □

Proposition 5.6. *For every $\epsilon > 0$, $\mu = 0$.*

Proof. Fix n and suppose $A \subseteq [0,n]$, $0, n \in A$, A is affluent for some $k < n/2$ and $\log|A+A|/\log|A-A| > 1+\epsilon$. Let $L = A \cap [0,k], R = (n-A) \cap [0,k]$. Then $sA - dA$ contains $[-dn+k+1, sn-k-1]$ and $\sigma A - \delta A$ contains $[-\delta n + k + 1, \sigma n - k - 1]$. This imposes strict bounds on $\log|sA-dA|/\log|\sigma A - \delta A|$. Note that the quantities $|(sR+dL) \cap [0,k]|$, $|(dL+sR) \cap [0,k]|$ are both bounded above by $k+1$. Observe the inequalities $|sA-dA| \le (s+d)n - 2k - 1 + 2k + 2$ and $|\sigma A - \delta A| \ge (\sigma+\delta)n - 2k - 1 + (s+d+1)$ (since $-dn, \dots, sn \in \sigma A + \delta A$). Thus

$$\begin{aligned}\frac{\log|sA-dA|}{\log|\sigma A - \delta A|} &\le \frac{\log((s+d)n - 2k - 1 + 2k + 2)}{\log((\sigma+\delta)n - 2k - 1 + (s+d+1))} \\ &\le \frac{\log((s+d)n+1)}{\log((s+d-1)n + s + d)}, \end{aligned} \tag{5.5}$$

which goes to 1 as n goes to infinity. There is an N such that $\log((s+d)n+1)/\log((s+d-1)n+s+d) < 1+\epsilon$ for all $n > N$. Therefore $\mu_n(L,R;k)$ is zero for all $n > N$. This implies that $\mu = \sum_{(L,R;k)} \mu(L,R;k)$ counts finitely many sets, so $\mu = 0$. □

To get rid of the stipulation that 0 and n are in A, we must show that $\lim_{n\to\infty} \Lambda_n = \lim_{n\to\infty} \lambda_n$. This result may be found in [Zh2, Lemma 2.15]. This concludes the proof of Proposition 5.3. The key point is showing that the proportion of affluent sets satisfying our property is the same as the proportion of all sets satisfying our property, an idea first developed in [Zh2]. The pairs (s,d) and (σ,δ) may be reversed to show that

Proposition 5.7. *Let $0 \le d < \delta \le \sigma < s$ with $s + d = \sigma + \delta$. For every $\epsilon > 0$ and a uniform random subset $A \subseteq [0, n]$,*

$$\lim_{n\to\infty} \Pr\left(\frac{\log|sA - dA|}{\log|\sigma A - \delta A|} \ < \ 1 - \epsilon\right) \ = \ 0. \tag{5.6}$$

In fact, (5.5) shows that there are no affluent sets A for which $|sA - dA|/|\sigma A - \delta A| \ge (s + d)/(s + d - 1)$. Similarly, there are no affluent sets A for which $|sA - dA|/|\sigma A - \delta A| \le (s + d - 1)/(s + d)$. Then we may say the following.

Proposition 5.8. *Let $0 \le d < \delta \le \sigma < s$ with $s + d = \sigma + \delta$. For a uniform random subset $A \subseteq [0, n]$,*

$$\lim_{n\to\infty} \Pr\left(\frac{s + d - 1}{s + d} \ < \ \frac{|sA - dA|}{|\sigma A - \delta A|} \ < \ \frac{s + d}{s + d - 1}\right) \ = \ 1. \tag{5.7}$$

Remark 5.9. *If $s + d = \sigma + \delta \ge 3$, then all of the instances of k-affluent sets in the proof above may be replaced with k-rich sets.*

6. Bi-MSTD Sets

We use similar techniques as in the previous sections to resolve the following question: *Is there a decomposition of $[0, n]$ into two disjoint MSTD sets?* We show that for sufficiently large n, such decompositions do exist and, in fact, are a positive proportion of all decompositions of $[0, n]$ as n goes to infinity. Rather than starting with an interval $[0, n]$ and decomposing it, we consider the cases in which both A and $A^c := [\min A, \max A] \setminus A$ are MSTD. We first begin with some terminology.

6.1. Definitions.

Definition 6.1. *We say that a set $A \subseteq \mathbb{Z}$ is* bi-MSTD *if A and A^c are both MSTD.*

An example of a bi-MSTD set in $[0, 19]$ is

$$A \ = \ \{0, 1, 3, 7, 8, 10, 11, 12, 15, 17, 18, 19\}. \tag{6.1}$$

Both A and its complement $A^c = \{2, 4, 5, 6, 9, 13, 14, 16\}$ are MSTD. Notice that $A^c - 2$ is the smallest MSTD set up to translation and dilation, as proved by Hegarty in [He]. By an exhaustive computer search, we determined that there are no bi-MSTD sets in $[0, 18]$. To show that a positive proportion of subsets of $[0, n]$ are bi-MSTD, we use the framework developed by Zhao in [Zh2]. We must first develop an analogue of fringe pairs for bi-MSTD sets.

Definition 6.2. *A fringe pair $(L, R; k)$ is* bi-MSTD *if*

$$\begin{aligned} |(L + L) \cap [0, k]| + |(R + R) \cap [0, k]| \ &> \ 2|(L + R) \cap [0, k]|, \\ |(L^c + L^c) \cap [0, k]| + |(R^c + R^c) \cap [0, k]| \ &> \ 2|(L^c + R^c) \cap [0, k]|, \end{aligned} \tag{6.2}$$

where $L^c = [0, k] \setminus L$ and $R^c = [0, k] \setminus R$.

Recall from Definition 2.3 that a set A which has MSTD fringe pair $(L, R; k)$ and satisfies the property

$$A + A \ \supseteq \ [k + 1, 2n - k - 1]$$

is called *rich*, and a rich set with an MSTD fringe pair is MSTD. Therefore, if A has a bi-MSTD fringe pair and satisfies

$$\begin{aligned} A + A \ &\supseteq \ [k + 1, 2n - k - 1], \\ A^c + A^c \ &\supseteq \ [k + 1, 2n - k - 1], \end{aligned} \tag{6.3}$$

then A is bi-MSTD. We call such sets *bi-rich*. We now show that the set of $A \subseteq [0,n]$ which are bi-rich has density one as n goes to infinity.

6.2. **Proof of Theorem 1.4.**

Lemma 6.3. *Let $0 < 2k < n$ and $A \subseteq [0,n]$ be a uniform random subset (dropping the condition that $0, n \in A$). Then*

$$\Pr([k+1, 2n-k-1] \not\subseteq A+A \text{ or } [k+1,2n-k-1] \not\subseteq A^c + A^c) \ \leq \ \frac{6(3/4)^{k/2}}{2-\sqrt{3}}.$$

Proof. We apply the union bound. The expression above is simply two times the expression given in Lemma 5.4. □

Note that for $k = 30$,

$$\Pr([k+1, 2n-k-1] \not\subseteq A+A \text{ or } [k+1,2n-k-1] \not\subseteq A^c + A^c) \ \leq \ 0.299.$$

Thus a positive proportion of $A \subseteq [0,n]$ are k-bi-rich (in fact, most of them are). Now we simply need to find a bi-MSTD pair $(L, R; k)$ with $k = 30$. We give an example below:

$$\begin{aligned} L &= \{0,1,2,5,8,10,11,12,14,15,16,18,23,25,26,28,29\}, \\ R &= \{0,1,3,4,8,10,11,13,14,15,17,19,20,22,23,24,28\}. \end{aligned} \tag{6.4}$$

With this fringe pair, we have shown that for some sufficiently large $N > 60$, for any $n > N$, the proportion of subsets of $[0,n]$ which are bi-MSTD is bounded below by a positive number. Since the bi-MSTD set in (6.1) is contained in $[0,19]$, the same holds for $n \geq 19$.

7. Future work

We end with a list of some interesting additional questions to pursue.

(1) We were able to construct a fringe pair for which $|sA - dA| > |\sigma A - \delta A|$ for $s + d = \sigma + d + 2 \leq q$ and $s > \sigma$. Can we construct a fringe pair to satisfy $|sA - dA| > |\sigma A - \delta A|$ for $s + d = \sigma + d + 2 \leq q$ for any sequence $\{(s, d, \sigma, \delta)_i : s + d = \sigma + \delta = i, 2 \leq i \leq q\}$ without using base expansion?

(2) As n goes to infinity, $|sA - dA|/|\sigma A - \delta A|$ is between $(s + d - 1)/(s + d)$ and $(s + d)/(s + d - 1)$ almost all the time. For any $\epsilon > 0$, is there a positive proportion of A such that $t < |sA - dA|/|\sigma A - \delta A| < t + \epsilon$ for $(s+d-1)/(s+d) < t < (s+d)/(s+d-1) - \epsilon$? What if instead of uniform random subsets $A \subseteq [0,n]$, we have the probability of $j \in A \subseteq [0,n]$ decaying with n?

(3) We showed that a positive proportion of subsets of $[0,n]$ are bi-MSTD, which is equivalent to showing that a positive proportion of decompositions of $[0,n]$ into two sets have the property that both sets are MSTD. Can we decompose $[0,n]$ into three sets which are MSTD? For any finite number k, is there a sufficiently large n for which there is a k-decomposition into MSTD sets?

(4) We found bi-MSTD fringe pairs by random searches and were not able to come up with any deterministic algorithm to produce bi-MSTD fringe pairs. Are there clean families of bi-MSTD fringe pairs? What about generalized bi-MSTD fringe pairs?

(5) We have shown that any arithmetic progression can be decomposed into two MSTD sets, and such decompositions have positive density among all decompositions of that arithmetic progression as n goes to infinity. What are the sets which can be decomposed into two MSTD sets? With positive density?
(6) In our proof that a positive proportion of subsets in $[0, n]$ are bi-MSTD, we showed that the proportion of bi-rich sets goes to one as n goes to infinity. How quickly does the density go to one? How quickly does the proportion of MSTD sets that are bi-MSTD increase with n?

References

[GH] D. Glasscock and G. Harcos, *Sumset estimates in abelian groups*, Central European University (masters thesis of Glasscock, advised by Harcos), (2011). Available online at https://people.math.osu.edu/glasscock.4/MSThesis.pdf.

[He] P. V. Hegarty, *Some explicit constructions of sets with more sums than differences* (2007), Acta Arithmetica **130** (2007), no. 1, 61–77.

[HM1] P. V. Hegarty and S. J. Miller, *When almost all sets are difference dominated*, Random Structures and Algorithms **35** (2009), no. 1, 118–136.

[ILMZ] G. Iyer, O. Lazarev, S. J. Miller and L. Zhang, *Generalized more sums than differences sets,* Journal of Number Theory **132** (2012), no. 5, 1054–1073

[MO] G. Martin and K. O'Bryant, *Many sets have more sums than differences*, in Additive Combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc., Providence, RI, 2007, pp. 287–305.

[MOS] S. J. Miller, B. Orosz and D. Scheinerman, *Explicit constructions of infinite families of MSTD sets*, Journal of Number Theory **130** (2010) 1221–1233.

[MS] S. J. Miller and D. Scheinerman, *Explicit constructions of infinite families of mstd sets,* Additive Number Theory, Springer, 2010, pp. 229-248.

[MV] S. J. Miler and K. Vissuet, *Most Subsets are Balanced in Finite Groups*, Combinatorial and Additive Number Theory, CANT 2011 and 2012 (Melvyn B. Nathanson, editor), Springer Proceedings in Mathematics & Statistics (2014), 147–157.

[Na1] M. B. Nathanson, *Sums of finite sets of integers*, The American Mathematical Monthly, Vol. 79, No. 9 (Nov., 1972), pp. 1010-1012.

[Na2] M. Nathanson, *Additive Number Theory: The Classical Bases*, Graduate Texts in Mathematics, Springer-Verlag, New York, 1996.

[Na3] M. B. Nathanson, *Problems in additive number theory, 1*, Additive combinatorics, 263–270, CRM Proc. Lecture Notes **43**, Amer. Math. Soc., Providence, RI, 2007.

[Na4] M. B. Nathanson, *Sets with more sums than differences*, Integers : Electronic Journal of Combinatorial Number Theory **7** (2007), Paper A5 (24pp).

[Ru1] I. Z. Ruzsa, *On the cardinality of $A+A$ and $A-A$*, Combinatorics year (Keszthely, 1976), vol. 18, Coll. Math. Soc. J. Bolyai, North-Holland-Bolyai Tàrsulat, 1978, 933–938.

[Ru2] I. Z. Ruzsa, *On the number of sums and differences*, Acta Math. Sci. Hungar. **59** (1992), 439–447.

[Zh1] Y. Zhao, *Constructing MSTD sets using bidirectional ballot sequences*, Journal of Number Theory **130** (2010), no. 5, 1212–1220.

[Zh2] Y. Zhao, *Counting MSTD sets in finite abelian groups*, Journal of Number Theory **130** (2010), no. 10, 2308–2322.

[Zh3] Y. Zhao, *Sets characterized by missing sums and differences*, Journal of Number Theory **131** (2011), no. 11, 2107–2134.

E-mail address: maa2@williams.edu

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

E-mail address: Sarah.Manski12@kzoo.edu

Department of Mathematics & Computer Science, Kalamazoo College, Kalamazoo, MI 49006

E-mail address: sjm1@williams.edu, Steven.Miller.MC.96@aya.yale.edu

Department of Mathematics and Statistics, Williams College, Williamstown, MA 01267

E-mail address: hs002012@pomona.edu

Department of Mathematics, Pomona College, Claremont, CA 91711