GENERALIZATIONS OF A CURIOUS FAMILY OF MSTD SETS HIDDEN BY INTERIOR BLOCKS

Hùng Việt Chu

Department of Mathematics, Washington and Lee University, Lexington, VA 24450

chuh19@mail.wlu.edu

Noah Luntzlara

Department of Mathematics, University of Michigan, Ann Arbor, MI 48109

nluntzla@umich.edu

Steven J. Miller

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

sjm1@williams.edu

Lily Shao

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

ls12@williams.edu

Abstract

A set A is MSTD (more-sum-than-difference) or sum-dominant if $|A + A| > |A - A|$, and is RSD (restricted-sum dominant) if $|A\hat{+}A| > |A - A|$, where $A\hat{+}A$ is the set of sums of distinct elements in A. We study an interesting family of MSTD sets that have appeared many times in the literature (see the works of Hegarty, Martin and O'Bryant, and Penman and Wells). While these sets seem at first glance to be ad hoc, looking at them in the right way reveals a nice common structure. In particular, instead of viewing them as explicitly written sets, we write them in terms of differences between two consecutive numbers in increasing order. We denote this family by $\mathcal{F}$ and investigate many of its properties. Using $\mathcal{F}$, we are able to generate many sets A with high value of $\log|A + A|/\log|A - A|$, construct sets A with a fixed $|A+A| - |A-A|$ more economically than previous authors, and improve the lower bound on the proportion of RSD subsets of $\{0, 1, 2, \ldots, n-1\}$ to about 10^{-25} (the previous best bound was 10^{-37}). Lastly, by exhaustive computer search, we find six RSD sets with cardinality 15, which is one lower than the smallest

cardinality found to date, and find that 30 is the smallest diameter of RSD sets.

1. Introduction

1.1. Background

Given a finite set of non-negative integers A, the *sum set* is defined to be

$$A + A := \{a_i + a_j : a_i, a_j \in A\}$$

and the *difference set* to be

$$A - A := \{a_i - a_j : a_i, a_j \in A\};$$

A is said to be *sum-dominated* or *MSTD* (more sums than differences) if $|A+A| > |A-A|$, *balanced* if $|A+A| = |A-A|$, and *difference-dominated* if $|A+A| < |A-A|$. Also, we define the *restricted sum set* to be

$$A \hat{+} A := \{a_i + a_j : a_i, a_j \in A \text{ and } a_i \neq a_j\}.$$

We call a set A *restricted sum-dominant (RSD)* if $|A\hat{+}A| > |A - A|$. We could similarly define a restricted difference set by only considering differences of distinct elements, but this would amount to removing the number 0 from the difference set, decreasing the cardinality of $A - A$ by one and not substantially changing the questions about RSD sets. Thus we avoid this definition.

Since Conway gave an early example of an MSTD set in 1969[1], research on MSTDs has made incredible progress; see [2, 6, 10, 11, 13, 14, 15] for some of the earlier results and constructions. One of the most notable papers is by Martin and O'Bryant [7]. They proved the proportion of MSTD subsets of $\{0, 1, 2, \ldots, n - 1\}$ is bounded below by a positive constant as $n \to \infty$. However, the proof is probabilistic and does not give explicit constructions of MSTD sets. Later, Miller, Orosz and Scheinerman [8] gave an explicit construction of a dense family of MSTD sets (previous bounds were exponentially small). They showed that as $n \to \infty$, the proportion of MSTD subsets of $\{0, 1, 2, \ldots, n-1\}$ that are in their family is at least C/n^4 for some constant C.[2] The current record of a dense family belongs to Zhao [17] with a family of density C/n.

In this paper, we focus on a particular family of MSTD sets (which we denote by $\mathcal{F}$) that has appeared many times in the literature. These sets appear to arise at random and have no particular order, but if we look at them in the right way, they are very well-structured. In addition, our family $\mathcal{F}$ has many nice properties that we will explore, despite not being dense.

---

[1] see footnote 1 of [11]

[2] With a more refined analysis, the density can be improved to C/n^2.

We first provide some examples of sets in $\mathcal{F}$ that have been discussed in the literature. The following sets are found in [7]:

$$\begin{aligned} S_2 &= \{0, 1, 2, 4, 5, 9, 12, 13, 14, 16, 17\}, \\ S_4 &= \{0, 1, 2, 4, 5, 9, 12, 13, 17, 20, 21, 22, 24, 25\}. \end{aligned}$$

And these sets below are found in [2]:

$$\begin{aligned} A_4 &= \{0, 1, 2, 4, 5, 9, 12, 13, 14\}, \\ A_{12} &= S_2, \\ A_{15} &= \{0, 1, 2, 4, 5, 9, 12, 13, 17, 20, 21, 22, 24, 25, 29, 32, 33, 37, 40, 41, 42, 44, 45\}. \end{aligned}$$

Last but not least, the following sets are found in [12]:

$$\begin{aligned} T_j' &= \{0, 2\} \cup \{1, 9, \ldots, 1 + 8j\} \cup \{4, 12, \ldots, 4 + 8j\} \\ &\quad \cup \{5, 13, \ldots, 5 + 8j\} \cup \{6 + 8j, 8(j+1)\} \text{ (Theorem 1)}, \\ T_j &= \{0, 2\} \cup \{1, 9, \ldots, 1 + 8(j+1)\} \cup \{4, 12, \ldots, 4 + 8j\} \\ &\quad \cup \{5, 13, \ldots, 5 + 8j\} \cup \{6 + 8j, 8(j+1)\} \text{ (Theorem 4)}, \\ R_j &= \{1, 4\} \cup \{0, 12, \ldots, 12j\} \cup \{2, 14, \ldots, 2 + 12j\} \\ &\quad \cup \{7, 19, \ldots, 7 + 12j\} \cup \{8, 20, \ldots, 8 + 12j\} \cup \{3 + 12j, 6 + 12j\} \text{ (Theorem 6)}. \end{aligned}$$

These sets play important roles in the papers which initially described them. For example, A_{12} is used to prove Theorem 8 in [2], which states that there exists a positive constant lower bound for the proportion of sets with fixed cardinalities of sum sets and difference sets. The sets T_j and T_j' give explicit construction of RSD sets. The set R_j gives a set A with the highest known value of $\log|A+A|/\log|A-A|$. The study of our family $\mathcal{F}$ was motivated by trying to describe a common pattern among these remarkable MSTD sets. Although they arose in somewhat different situations, and were presented ad hoc for the purposes of each of the papers which described them, it turned out that all these sets have something in common, and all belong to $\mathcal{F}$.

The members of the family $\mathcal{F}$ have many nice properties, including (1) sets A with large values of $\log|A + A|/\log|A - A|$[3]; (2) economical construction of sets A with fixed $|A + A| - |A - A|$; (3) demonstration of Spohn's conjecture (1973); (4) compactness; (5) more constructions of RSD subsets; and (5) examples of small fringes.

1.2. Notation and Main Results

Let nonnegative numbers $a \leq b$ be chosen. Let $[a, b] = \{x \mid a \leq x \leq b\}$ and $[a, b]_q = \{x \mid x \equiv a \pmod{q} \text{ and } a \leq x \leq b\}$.

---

[3]The highest value of $\log|A + A|/\log|A - A|$ for a set A in $\mathcal{F}$ is about 1.03059.

We use a different notation to represent sets of integers; it was first introduced by Spohn [16] (1973): Given a set $S = \{a_1, a_2, \ldots, a_n\}$, we arrange its elements in increasing order and form the sequence of differences between consecutive elements. Suppose that $a_1 < a_2 < \cdots < a_n$, then our sequence is $a_2 - a_1, a_3 - a_2, a_4 - a_3, \ldots, a_n - a_{n-1}$ and we represent

$$S = (a_1 \mid a_2 - a_1, a_3 - a_2, a_4 - a_3, \ldots, a_n - a_{n-1}).$$

For example, if $S = \{3, 2, 5, 10, 9\}$, we would arrange the elements in increasing order to get $2 < 3 < 5 < 9 < 10$, then write $S = (2 \mid 1, 2, 4, 1)$.

We call the sequence $a_2 - a_1, a_3 - a_2, a_4 - a_3, \ldots, a_n - a_{n-1}$ the *sequence of consecutive differences (SCD)*. The advantage of this notation is that differences between elements of S correspond to sums of consecutive runs in the SCD. For example, look at the SCD $1, 2, 4, 1$. We know that 7 is in the difference set because the run $1, 2, 4$ sums up to 7.

We now define the $\mathcal{F}$ *family* and *interior blocks*.

Definition 1.1. Let M^k denote the sequence $1, \underbrace{4, \ldots, 4}_{k\text{-times}}, 3$. We define $\mathcal{F}$ to be the family of sets with SCD

$$1, 1, 2, 1, M^{k_1}, M^{k_2}, \ldots, M^{k_\ell}, M_1$$

where $\ell, k_1, \ldots, k_\ell$ are positive integers, and M_1 is either $1, 1$ or $1, 1, 2$ or $1, 1, 2, 1$.

Remark 1.2. It can be verified that all the sets $S_2, S_4, A_4, A_{12}, A_{15}, T'_j, T_j, R_j$ are in $\mathcal{F}$.

Conjecture 1.3. *All sets in $\mathcal{F}$ are MSTD.*

Example 1.4. The set

$$\begin{aligned} S &= (0 \mid 1, 1, 2, 1, 4, 3, 1, 4, 4, 3, 1, 4, 4, 4, 3, 1, 1, 2, 1) \\ &= \{0, 1, 2, 4, 5, 9, 12, 13, 17, 21, 24, 25, 29, 33, 37, 40, 41, 42, 44, 45\} \end{aligned}$$

has $|S + S| - |S - S| = 86 - 83 = 3$.

In this paper, we prove that the conjecture is true for a periodic subfamily of $\mathcal{F}$.

Definition 1.5. Consider a set S with its SCD. Let B denote a consecutive subsequence of the SCD. Suppose there exists $N \in \mathbb{N}$ such that for $k \geq N$, the sets S_k with SCD constructed by repeating B for k times in place of B are MSTD sets. Then we call B an *interior block* and let $|B|$ denote the length of the interior block.

A natural question to ask is what are the possible values of $|B|$, which is addressed by the following theorem:

Theorem 1.6. *Let k and ℓ be arbitrary positive integers. The following three subfamilies of $\mathcal{F}$ consist of MSTD sets; we specify the exact values of $|S+S|-|S-S|$.*

1. *The family $\{S_{k,\ell} \mid k,\ell \in \mathbb{N}\}$, where*

$$S_{k,\ell} = (0 \mid 1,1,2,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,\ldots,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,1,1,2,1)$$
$$\underbrace{}_{\ell\text{-times}}$$

has $|S_{k,\ell} + S_{k,\ell}| - |S_{k,\ell} - S_{k,\ell}| = 2\ell$.

2. *The family $\{S'_{k,\ell} \mid k,\ell \in \mathbb{N}\}$, where*

$$S'_{k,\ell} = (0 \mid 1,1,2,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,\ldots,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,1,1,2)$$
$$\underbrace{}_{\ell\text{-times}}$$

has $|S'_{k,\ell} + S'_{k,\ell}| - |S'_{k,\ell} - S'_{k,\ell}| = 2\ell - 1$.

3. *The family $\{S''_{k,\ell} \mid k,\ell \in \mathbb{N}\}$, where*

$$S''_{k,\ell} = (0 \mid 1,1,2,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,\ldots,1,\underbrace{4,\ldots,4}_{k\text{-times}},3,1,1)$$
$$\underbrace{}_{\ell\text{-times}}$$

has $|S''_{k,\ell} + S''_{k,\ell}| - |S''_{k,\ell} - S''_{k,\ell}| = \ell$.

We call $\{S_{k,\ell}|k,\ell \in \mathbb{N}\} \cup \{S'_{k,\ell}|k,\ell \in \mathbb{N}\} \cup \{S''_{k,\ell}|k,\ell \in \mathbb{N}\}$ the $\mathcal{F}_{per}$ family, which is a periodic subfamily of the larger family $\mathcal{F}$.

Example 1.7.   1. The set

$$\begin{aligned}
S_{2,3} &= (0 \mid 1,1,2,1,4,4,3,1,4,4,3,1,4,4,3,1,1,2,1) \\
&= \{0,1,2,4,5,9,13,16,17,21,25,28,29,33,37,40,41,42,44,45\}
\end{aligned}$$

has $|S_{2,3} + S_{2,3}| - |S_{2,3} - S_{2,3}| = 85 - 79 = 6 = 2 \cdot 3$.

2. The set

$$\begin{aligned}
S'_{3,2} &= (0 \mid 1,1,2,1,4,4,4,3,1,4,4,4,3,1,1,2) \\
&= \{0,1,2,4,5,9,13,17,20,21,25,29,33,36,37,38,40\}
\end{aligned}$$

has $|S'_{3,2} + S'_{3,2}| - |S'_{3,2} - S'_{3,2}| = 72 - 69 = 3 = 2 \cdot 2 - 1$.

3. The set

$$\begin{aligned}
S''_{3,2} &= (0 \mid 1,1,2,1,4,4,4,3,1,4,4,4,3,1,1) \\
&= \{0,1,2,4,5,9,13,17,20,21,25,29,33,36,37,38\}
\end{aligned}$$

has $|S''_{3,2} + S''_{3,2}| - |S''_{3,2} - S''_{3,2}| = 67 - 65 = 2$.

Remark 1.8. Theorem 1.6 answers a question raised by Spohn [16], on whether the interior block must contain at least 3 elements. The answer is no: set $\ell = 1$ and choose $k \geq 2$ for sets in Theorem 1.6. Then, we can use either the interior block 4 or $4, 4$. In other words, we have shown that for all $i \in \mathbb{N}$, there exists an interior block B with $|B| = i$. Also, this theorem provides an infinite family of sets which demonstrate Conjecture 6 in [16].[4]

There has been a lot of interest in finding sets A with large values of the ratio $\log |A + A| / \log |A - A|$. An early high ratio of about 1.0208 was given by Hegarty [2], and later, a higher ratio of about 1.02313 was found by Asada et al. [1]. The current highest is about 1.03059 found by Penman and Wells in [12], which is much higher than previous results. We observe that both examples of sets A with high ratios from [2] and [12] belong to $\mathcal{F}$. We offer examples of several sets in $\mathcal{F}$ that give higher ratios than the ones given in [1] and [2]; there are at least 22 sets A in $\mathcal{F}$ with $\log |A + A| / \log |A - A| > 1.3$.

Furthermore, the family $\mathcal{F}$ gives an economical way (in the sense of having a relatively small width between its minimum and maximum element) to construct a set A with any specific value of $|A + A| - |A - A|$. Martin and O'Bryant proved that for a given $x \in \mathbb{N}$, there exists a set $A \subseteq [0, 17|x|]$ such that $|A + A| - |A - A| = x$, which is significantly more efficient than the base expansion method.[5] With subfamilies of $\mathcal{F}$, we further improve this.

Theorem 1.9. *Given $x \in \mathbb{N}$, there exists a set $A \subseteq [0, 12 + 4x]$ such that $|A + A| - |A - A| = x$. Furthermore, it is impossible to construct $A \subseteq [0, f(x)]$ such that $|A + A| - |A - A| = x$, where $f(x)$ is sub-linear. This means that a linear growth of the interval containing A is the best we can do.*

Finally, we improve the lower bound for the proportion of RSD subsets of $[0, n-1]$ as n goes to infinity. RSD implies MSTD, and compared to MSTD sets, RSD sets are much less common: exhaustive computer search shows that there are no RSD subsets of $[0, 29]$, while there are at least $4.5 \cdot 10^5$ MSTD sets in the same interval. In [12], the lower bound on the proportion of RSD subsets of $[0, n-1]$ as $n \to \infty$ is about 10^{-37}; we improve this bound to $4.135 \cdot 10^{-25}$ by a better fringe formed by using $\mathcal{F}$.[6]

Theorem 1.10. *For $n \geq 81$, the proportion of RSD subsets of $[0, n-1]$ is at least $4.135 \cdot 10^{-25}$.*

---

[4]The repetition of certain interior blocks can cause the number of sums to be increased by a greater constant than that by which the number of differences is increased.

[5]We can generate an infinite family of MSTD sets from a given MSTD set through the base expansion method. Let A be an MSTD set, and let $A_{k,m} = \{\sum_{i=1}^{k} a_i m^{i-1} : a_i \in A\}$. If m is sufficiently large, then $|A_{k,m} \pm A_{k,m}| = |A \pm A|^k$ and $|A_{k,m}| = |A|^k$.

[6]There are exactly 6 RSD subsets of $[0, 30]$ and 16 RSD subsets of $[0, 31]$. Based on these observations, we predict the true proportion of RSD subsets of $[0, n-1]$ as $n \to \infty$ to be about $3 \cdot 10^{-9}$.

2. Proof of Theorem 1.6

We only prove (1) for conciseness, since the proofs of (2) and (3) are similar. For (1), the case when $k = 1$ is Theorem 4 in [12], so we prove the case $k \geq 2$.

Lemma 2.1. *Fix $i \geq 0$ and consider $6 + 4i$. For some fixed k and ℓ, if there is a run that sums up to $6 + 4i$ in the SCD*

$$1, 1, 2, 1, \underbrace{\underbrace{4, \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{k\text{-}times}, 3}_{\ell\text{-}times}, 1, 1, 2, 1,$$

then the run is one of the forms

(A) $1, 1, 2, 1, \underbrace{\underbrace{4, \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{k\text{-}times}, 3}_{\ell\text{-}times}, 1, 1$

(B) $2, 1, \underbrace{\underbrace{4, \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{k\text{-}times}, 3}_{j\text{-}times}$

(C) $2, 1, \underbrace{\underbrace{4, \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{k\text{-}times}, 3}_{\ell\text{-}times}, 1, 1, 2$

(D) $1, \underbrace{\underbrace{4, \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{k\text{-}times}, 3}_{j\text{-}times}, 1, 1$

for some $j \geq 1$.

Proof. We consider possible cases for where the run that sums up to $6 + 4i$ can start.

1. Case I: the run starts at the first 1. Since $1 + 1 + 2 + 1 = 5 < 6$, the run must contain $1, 1, 2, 1$.

   (a) Subcase 1: If the run ends at 4, the run sums up to $1 + 4m$ for some $m \geq 0$.

   (b) Subcase 2: If the run ends at 3, we have $4m$ for some $m \geq 0$.

   (c) Subcase 3: If the run ends with $3, 1$, we have $1 + 4m$ for some $m \geq 0$.

   (d) Subcase 4: If the run ends with $3, 1, 1$, we have $6 + 4m$ for some $m \geq 0$. We have form (A).

(e) Subcase 5: If the run ends with $3, 1, 1, 2$, we have $4m$ for some $m \geq 0$.

(f) Subcase 6: If the run ends with $3, 1, 1, 2, 1$, we have $1 + 4m$ for some $m \geq 0$.

2. Case II: the run starts at the second 1. As above, the run must contain $1, 2, 1$. Using the same argument, we see that there are no such runs that sum up to $6 + 4i$.

3. Case III: the run starts at the first 2. As above, the run must contain $2, 1$. Using the same argument, we see that to have $6 + 4i$, the run must either end at 3 or end with $3, 1, 1, 2$. We have form (B) and form (C).

4. Case IV: the run starts with $1, 4$; it must end with $3, 1, 1$.

5. Case V: the run starts at 4; there are no such runs.

6. Case VI: the run starts at 3; there are no such runs.

We have iterated through all possible cases and thus the proof is complete. $\square$

Lemma 2.2. *Let k and $\ell \in \mathbb{N}$ be chosen. Consider*

$$S = (0 \mid 1, 1, 2, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, 1, 1, 2, 1).$$

$$\underbrace{\hspace{5cm}}_{\ell\text{-}times}$$

Then the set $[1, \max(S)] \setminus (S - S)$ of missing positive differences is exactly

$$
\begin{aligned}
T \;=\; & [6, 6 + 4(k-1)]_4 \cup [14 + 4(k-1), 14 + 8(k-1)]_4 \cup \\
& \cdots \cup [6 + 8(\ell - 1) + 4(\ell - 1)(k-1), 6 + 8(\ell - 1) + 4\ell(k-1)]_4.
\end{aligned}
$$

Recall from Section 1.2 that $[a, b]_4 := \{x \mid x \equiv a \pmod{4} \text{ and } a \leq x \leq b\}$.

Example 2.3. We use $S_{2,3}$ as an example. We have

$$S_{2,3} - S_{2,3} \;=\; [-45, 45] \setminus \{\pm 6, \pm 10, \pm 18, \pm 22, \pm 30, \pm 34\}.$$

Note that

$$\{6, 10, 18, 22, 30, 34\} \;=\; [6, 10]_4 \cup [18, 22]_4 \cup [30, 34]_4,$$

which is the set T in Lemma 2.2.

Proof of Lemma 2.2. Pick $1 \leq i \leq \ell$. We show that $S - S$ misses

$$[6 + 8(i-1) + 4(i-1)(k-1), 6 + 8(i-1) + 4i(k-1)]_4;$$

equivalently, there are no runs that sum up to $6 + 8(i - 1) + 4(i - 1)(k - 1) + 4m$ for all $0 \leq m \leq k - 1$. We prove this by contradiction. Pick some $0 \leq m \leq k - 1$. Suppose that such a run exists; the run must be one of the forms in Lemma 2.1. Notice that

$$\begin{aligned} 6 + 8(i - 1) + 4(i - 1)(k - 1) + 4m &\leq 6 + 8(i - 1) + 4i(k - 1) \\ &\leq 6 + 8(\ell - 1) + 4\ell(k - 1). \end{aligned}$$

Since both form (A) and form (C) in Lemma 2.1 gives $6 + 8\ell + 4(k - 1)\ell$, our run must be of the form (B) or (D). We consider these two cases.

1. Case I: the run is of form (B). Then it sums up to $2 + 4(k+1)j$ for some $j \geq 1$. We have:

$$\begin{aligned} 2 + 4(k + 1)j &= 6 + 8(i - 1) + 4(i - 1)(k - 1) + 4m \\ (k + 1)(j - i + 1) &= 1 + m. \end{aligned} \tag{1}$$

   So, $1 \leq (k + 1)(j - i + 1) = 1 + m \leq k$ and so, $0 < j - i + 1 < 1$, which is a contradiction.

2. Case II: the run is of the form (D). Then it sums up to $4(k + 1)j + 2$ for some $j \geq 1$. As above, we find a contradiction. We have shown that $S - S$ misses T.

To complete the proof, we show that $S - S$ contains $[0, 9 + 4(k+1)\ell] \backslash T$. Note that close to the beginning of the SCD, we have $1 + 2 + 1 = 4$ and after that, the sequence implicitly contains consecutive differences of 4 (because $3 + 1 = 4$ and $1 + 2 + 1 = 4$.) So, $S - S$ contains all numbers in $[0, 9 + 4(k + 1)\ell]$ that are 0 mod 4. Similarly, it is not hard to see that $S - S$ contains all numbers that are either 1 mod 4 or 3 mod 4. Next, we show that all numbers that are 2 mod 4 and not in T are in $S - S$. We have $2 \in S - S$ and $10 + 4(k - 1) \in S - S$ because $10 + 4(k - 1) = (2 + 1) + 4(k - 1) + (4 + 3)$. Therefore, $\{10 + 4(k - 1) + (1 + 4k + 3)i \mid 0 \leq i \leq \ell - 1\} \subseteq S - S$. Lastly, $6 + 4(k + 1)\ell \in S - S$. because we have the run $2, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}, 3, 1, 1, 2$. $\qquad \square$

$$\underbrace{\phantom{2, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}}}_{\ell\text{-times}}$$

Corollary 2.4. *Choose k and $\ell \in \mathbb{N}$, and let*

$$S = (0 \mid 1, 1, 2, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}, 3, 1, 1, 2, 1).$$

$$\underbrace{\phantom{S = (0 \mid 1, 1, 2, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-times}}}}_{\ell\text{-times}}$$

Then $|S - S| = 19 + \ell(6k + 8)$.

Lemma 2.5. *Choose $k \geq 2$ and $\ell \in \mathbb{N}$, and let*

$$S = (0 \mid 1, 1, 2, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, 1, 1, 2, 1).$$

$$\underbrace{}_{\ell\text{-}times}$$

Then $S + S$ contains $[0, 8(k+1)\ell + 18] \backslash T$, where

$$T = \left(\bigcup_{i \in [1, \ell]} [12 + 12(i-1) + 4(i-1)(k-2), 12 + 12(i-1) + 4i(k-2)]_4 \right)$$

$$\cup \left(\bigcup_{i \in [1, \ell]} [12\ell + 4\ell(k-2) + 16 + 4(i-1)(k-2) + 12(i-1), \right.$$

$$\left. 12\ell + 4\ell(k-2) + 16 + 4i(k-2) + 12(i-1)]_4 \right).$$

Proof. Observe that $S_1 = \{i \mid 1 \leq i \leq 4(k+1)\ell + 9 \text{ and } i \equiv 1 \bmod 4\} \subseteq S$. So, all of the following sets are in $S + S$:

$$
\begin{aligned}
0 + S_1 &= \{i \mid 1 \leq i \leq 4(k+1)\ell + 9 \text{ and } i \equiv 1 \bmod 4\}, \\
1 + S_1 &= \{i \mid 2 \leq i \leq 4(k+1)\ell + 9 \text{ and } i \equiv 2 \bmod 4\}, \\
2 + S_1 &= \{i \mid 3 \leq i \leq 4(k+1)\ell + 9 \text{ and } i \equiv 3 \bmod 4\}, \\
4(k+1)\ell + 9 + S_1 &= \{i \mid 4(k+1)\ell + 10 \leq i \leq 8(k+1)\ell + 18 \text{ and } i \equiv 2 \bmod 4\}, \\
4(k+1)\ell + 8 + S_1 &= \{i \mid 4(k+1)\ell + 9 \leq i \leq 8(k+1)\ell + 17 \text{ and } i \equiv 1 \bmod 4\}, \\
4(k+1)\ell + 6 + S_1 &= \{i \mid 4(k+1)\ell + 7 \leq i \leq 8(k+1)\ell + 15 \text{ and } i \equiv 3 \bmod 4\}
\end{aligned}
$$

Thus $S + S$ contains all numbers that are either $1, 2$ or $3 \bmod 4$ in the interval $[0, 8(k+1)\ell + 18]$. Now, we focus on numbers that are divisible by 4. Observe that $S_2 = \{0, 4, 4 + 4(k+1), 4 + 8(k+1), \ldots, 4 + 4\ell(k+1), 8 + 4\ell(k+1)\} \subseteq S$. We write

$$S_2 = \{0, 8 + 4\ell(k+1)\} \cup \{4 + 4i(k+1) \mid 0 \leq i \leq \ell\}.$$

We show that all numbers divisible by 4 that are not in T are in $S + S$. The set of all numbers divisible by 4 that are not in T is

$$\{0, 4, 8, 12 + 4\ell(k+1), 8\ell + 8k\ell + 16\} \cup (4 + \{12 + 12(i-1) + 4i(k-2) \mid 1 \leq i \leq \ell\})$$

$$\cup (8 + \{12 + 12(i-1) + 4i(k-2) \mid 1 \leq i \leq \ell\})$$

$$\cup (20 + \{12\ell + 4(\ell + i)(k-2) + 12(i-1) \mid 1 \leq i \leq \ell\})$$

$$\cup (24 + \{12\ell + 4(\ell + i)(k-2) + 12(i-1) \mid 1 \leq i \leq \ell\}).$$

We know the following:

1. Because $0, 4, 8 + 4(k+1)\ell \in S$, $\{0, 4, 8, 12 + 4\ell(k+1), 8\ell + 8k\ell + 16\} \subseteq S + S$.

2. For each $1 \leq i \leq \ell$, we have $4 + (12 + 12(i-1) + 4i(k-2)) = 0 + (4 + 4i(k+1)) \in S + S$.

3. For each $1 \leq i \leq \ell$, we have $8 + (12 + 12(i-1) + 4i(k-2)) = 4 + (4 + 4i(k+1)) \in S + S$.

4. For each $1 \leq i \leq \ell$, we have $20 + 12\ell + 4(\ell+i)(k-2) + 12(i-1) = (4 + 4\ell(k+1)) + (4 + 4i(k+1)) \in S + S$.

5. For each $1 \leq i \leq \ell$, we have $24 + 12\ell + 4(\ell+i)(k-2) + 12(i-1) = (8 + 4\ell(k+1)) + (4 + 4i(k+1)) \in S + S$.

We have shown that all numbers divisible by 4 that are not in T are in $S + S$, and this completes the proof. $\qquad\square$

Lemma 2.6. *Choose $k \geq 2$ and $\ell \in \mathbb{N}$, and let*

$$S = (0 \,|\, 1, 1, 2, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, 1, 1, 2, 1).$$
$$\underbrace{}_{\ell\text{-}times}$$

Then $S + S$ contains none of the elements in

$$T = \left(\bigcup_{i \in [1,\ell]} \left[12 + 12(i-1) + 4(i-1)(k-2), 12 + 12(i-1) + 4i(k-2) \right]_4 \right)$$

$$\cup \left(\bigcup_{i \in [1,\ell]} \left[12\ell + 4\ell(k-2) + 16 + 4(i-1)(k-2) + 12(i-1), \right. \right.$$

$$\left. \left. 12\ell + 4\ell(k-2) + 16 + 4i(k-2) + 12(i-1) \right]_4 \right).$$

Proof. To complete the proof, we prove that none of the numbers in T are in $S + S$. We write out S explicitly:

$$S = \{0, 1, 2, 4, 5\} \cup \{5 + 4(j-1)(k+1) + 4i \,|\, 1 \leq j \leq \ell, 1 \leq i \leq k\}$$
$$\cup \{4 + 4i(k+1) | 1 \leq i \leq \ell\} \cup \{5 + 4i(k+1) | 1 \leq i \leq \ell\}$$
$$\cup \{6 + 4\ell(k+1), 8 + 4\ell(k+1), 9 + 4\ell(k+1)\}.$$

We consider elements in

$$\bigcup_{i \in [1,\ell]} \left[12 + 12(i-1) + 4(i-1)(k-2), 12 + 12(i-1) + 4i(k-2) \right]_4.$$

Pick $1 \le m \le \ell$ and $0 \le n \le k - 2$. Consider

$$
\begin{aligned}
12 + 12(m - 1) + 4(m - 1)(k - 2) + 4n &= 4m(k + 1) - 4k + 8 + 4n \\
&\le 4\ell + 4k\ell - 4k + 8 + 4(k - 2) \\
&= 4\ell(k + 1).
\end{aligned}
$$

Because S contains no numbers that are 3 mod 4, for a pair whose sum is $4m(k + 1) - 4k + 8 + 4n$, we cannot use numbers that are 1 mod 4. Also, because $4m(k+1) - 4k + 8 + 4n \le 4\ell(k + 1)$, we can ignore all numbers that are greater than $4\ell(k + 1)$. Hence, our set of concern is

$$
\{0, 2, 4\} \ \cup \ \{4 + 4i(k + 1) \,|\, 1 \le i \le \ell\}.
$$

If a pair that sums to $4m(k + 1) + 4(n - k) + 8$ is in $\{4 + 4i(k+1) \,|\, 1 \le i \le \ell\}$, then there exists m' and n' such that

$$
\begin{aligned}
8 + 4(m' + n')(k + 1) &= 4m(k + 1) + 4(n - k) + 8, \\
(m' + n' - m + 1)(k + 1) &= n + 1.
\end{aligned}
$$

Thus, $0 < (m' + n' - m + 1)(k + 1) = n + 1 \le k - 1$ and we get, $0 < m' + n' - m + 1 < 1$, a contradiction. Therefore, one of the number is in $\{0, 2, 4\}$. Let $4 + 4m'(k + 1)$ be the number used in $\{4 + 4i(k+1) | 1 \le i \le \ell\}$. We consider three cases corresponding to the three elements in $\{0, 2, 4\}$.

1. We have $0 + (4 + 4m'(k+1)) = 4m(k + 1) - 4k + 8 + 4n$. So, $0 < (k + 1)(m' - m + 1) = n + 2 \le k$, which implies $0 < m' - m + 1 < 1$, a contradiction.

2. We have $2 + (4 + 4m'(k+1)) = 4m(k + 1) - 4k + 8 + 4n$. So, $0 < 2(m' - m + 1)(k + 1) = 3 + 2n \le 2k - 1$, which implies $0 < m' - m + 1 < 1$, a contradiction.

3. We have $4 + (4 + 4m'(k+1)) = 4m(k + 1) - 4k + 8 + 4n$. So, $0 < (m' - m + 1)(k + 1) = n + 1 \le k - 1$, which implies $0 < m' - m + 1 < 1$, a contradiction.

Next, we consider elements in

$$
\bigcup_{i \in [1, \ell]} \big[12\ell + 4\ell(k-2) + 16 + 4(i-1)(k-2) + 12(i-1), \, 12\ell + 4\ell(k-2) + 16 + 4i(k-2) + 12(i-1) \big]_4.
$$

Pick $1 \le m \le \ell$ and $0 \le n \le k - 2$. Consider

$$
\begin{aligned}
12\ell &+ 4\ell(k - 2) + 16 + 4(m - 1)(k - 2) + 12(m - 1) + 4n \\
&= 4(\ell + m)(k + 1) + 12 - 4k + 4n \\
&\ge 4(\ell + 1)(k + 1) + 12 - 4k = 4(k + 1)\ell + 16.
\end{aligned}
$$

Thus, we cannot use any of $0, 2, 4$ in our pair. As above, the set which concerns us is

$$\{4 + 4i(k+1) \mid 1 \le i \le \ell\} \ \cup \ \{6 + 4\ell(k+1), 8 + 4\ell(k+1)\}.$$

If a pair that sums to $4(\ell+m)(k+1) + 12 - 4k + 4n$ is in $\{4 + 4i(k+1) | 1 \le i \le \ell\}$, then for some $1 \le m', n' \le \ell$ we have:

$$\begin{aligned}
4(\ell+m)(k+1) + 12 - 4k + 4n &= 8 + 4(k+1)(m' + n') \\
0 < (-\ell - m + m' + n' + 1)(k+1) &= n + 2 \le k
\end{aligned}$$

So, $0 < -\ell - m + m' + n' + 1 < 1$, a contradiction. Therefore, a number in the pair must be in $\{6 + 4\ell(k+1), 8 + 4\ell(k+1)\}$. Since $4(\ell+m)(k+1) + 12 - 4k + 4n \le 8\ell(k+1) + 4$, both numbers cannot be in $\{6 + 4\ell(k+1), 8 + 4\ell(k+1)\}$. We consider two cases corresponding to the two elements of $\{6 + 4\ell(k+1), 8 + 4\ell(k+1)\}$:

1. We have $(6 + 4\ell(k+1)) + (4 + 4m'(k+1)) = 4(\ell+m)(k+1) + 12 - 4k + 4n$. Equivalently, $0 < 2(k+1)(m' - m + 1) = 2n + 3 \le 2(k-2) + 3 = 2k - 1$ and so $0 < m' - m + 1 < 1$, a contradiction.

2. We have $(8 + 4\ell(k+1)) + (4 + 4m'(k+1)) = 4(\ell+m)(k+1) + 12 - 4k + 4n$. Equivalently, $0 < (k+1)(m' - m + 1) = n + 1 \le (k-2) + 1 = k - 1$ and so, $0 < m' - m + 1 < 1$, a contradiction.

This completes our proof.   □

Example 2.7. We use $S_{2,3}$ as an example to illustrate Lemma 2.5 and Lemma 2.6. We have

$$S_{2,3} + S_{2,3} \ = \ [0, 90] \backslash \{12, 24, 36, 52, 64, 76\},$$

and the set $\{12, 24, 36, 52, 64, 76\}$ is exactly the set T in Lemmas 2.5 and Lemma 2.6.

Corollary 2.8. *Let $k \ge 2$ and $\ell \in \mathbb{N}$ be chosen. Let*

$$S \ = \ (0 \mid 1, 1, 2, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, \ldots, 1, \underbrace{4 \ldots, 4}_{k\text{-}times}, 3, 1, 1, 2, 1).$$
$$\underbrace{\phantom{\hspace{6cm}}}_{\ell\text{-}times}$$

Then $|S + S| = 19 + \ell(6k + 10)$.

Proof of Theorem 1.6, Item 1. The proof follows immediately from Corollaries 2.4 and 2.8, because

$$|S_{k,\ell} + S_{k,\ell}| - |S_{k,\ell} - S_{k,\ell}| = [19 + \ell(6k + 10)] - [19 + \ell(6k + 8)] = 2\ell. \qquad \square$$

Remark 2.9. Theorem 1.6 is a generalization of Theorem 2 and Theorem 3 in [12]. With a little more work, we can show that sets in $\mathcal{F}_{per}$ are RSD for ℓ sufficiently large.

We also offer another family of MSTD sets formed by repeating certain interior blocks. We do not prove the theorem since it is not in the focus of the current paper. However, the proof is very similar to the proof of Theorem 1.6 but replacing "modulo 4" by "modulo k" throughout.

Theorem 2.10. *For $k \geq 4$, the following is a MSTD set:*

$$A_{k,1} \;=\; (0 \,|\, \underbrace{1, 1, \ldots, 1}_{k-2\text{-}times}, 2, 1, k \underbrace{k+1, k+1, \ldots, k+1}_{k-4\text{-}times}, 3, \underbrace{1, 1, \ldots, 1}_{k-2\text{-}times}, 2, 1),$$

and $|A_{k,1} + A_{k,1}| - |A_{k,1} - A_{k,1}| = 2$. Define $A_{k,\ell}$ to be a similarly built set with the sequence $1, k, \underbrace{k+1, k+1, \ldots, k+1}_{k-4\text{-}times}, 3$ repeated k times, then

$$|A_{k,\ell} + A_{k,\ell}| - |A_{k,\ell} - A_{k,\ell}| = 2\ell$$

Remark 2.11. If we consider Theorem 1.6 to be a generalization of Theorem 2 and Theorem 4 in [12], then Theorem 2.10 is another generalization from a different perspective. Notice that $A_{4,\ell} = S_{1,\ell}$. Sets $A_{k,\ell}$ also form a family of MSTD sets and RSD sets with interior blocks.

3. Good Properties of the Family $\mathcal{F}$

3.1. Sets A with Large $\log|A+A|/\log|A-A|$

The first application of our family $\mathcal{F}$ is that the family produces many sets A with large value of $\log|A+A|/\log|A-A|$. For convenience, we define $f(A) := \log|A+A|/\log|A-A|$. An early example of a set A with high $f(A)$ is given by Hegarty [2]. The set is

$$A_{15} \;=\; \{0, 1, 2, 4, 5, 9, 12, 13, 17, 20, 21, 22, 24, 25, 29, 32, 33, 37, 40, 41, 42, 44, 45\}.$$

In our notation,

$$A_{15} \;=\; (0 \,|\, 1, 1, 2, 1, 4, 3, 1, 4, 3, 1, 1, 2, 1, 4, 3, 1, 4, 3, 1, 1, 2, 1),$$

which is very close to a set in our family $\mathcal{F}_{per}$, namely

$$S_{1,4} \;=\; (0 \,|\, 1, 1, 2, 1, 4, 3, 1, 4, 3, 1, 4, 3, 1, 1, 2, 1).$$

It turns out that $f(S_{1,4}) > f(A_{15})$.

When analyzing the periodic subfamily $\mathcal{F}_{per}$ defined in Theorem 1.6, we find the set $S_{1,6}$ with the property of $f(S_{1,6}) = 1.023777\ldots$, which is larger than previous results in [2] $(1.0208\ldots)$ and [1] $(1.0213\ldots)$ but smaller than the current record in [12] $(1.03059\ldots)$. It is worth noting, however, the set with the highest known value of $f(A)$ $(1.3059\ldots)$ exhibited by Penman and Wells [12] is a member of the family $\mathcal{F}$:

$$(0 \mid 1, 1, 2, 1, 4, 3, 1, 4, 4, 3, 1, \underbrace{4, \ldots, 4}_{\text{3-times}}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{\text{3-times}}, 3, 1, 4, 4, 3, 1, 4, 3, 1, 1, 2, 1).$$
$$\underbrace{}_{\text{9-times}}$$

In fact, there are at least 22 sets A in $\mathcal{F}$ with $f(A) > 1.03$: these sets are of the form:

$$(0 \mid 1, 1, 2, 1, 4, 3, 1, 4, 4, 3, 1, \underbrace{4, \ldots, 4}_{\text{3-times}}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{\text{3-times}}, 3, 1, 4, 4, 3, 1, 4, 3, 1, 1, 2, 1),$$
$$\underbrace{}_{\ell\text{-times}}$$

$$(0 \mid 1, 1, 2, 1, 4, 3, 1, 4, 4, 3, 1, \underbrace{4, \ldots, 4}_{\text{3-times}}, 3, \ldots, 1, \underbrace{4, \ldots, 4}_{\text{3-times}}, 3, 1, 4, 4, 3, 1, 4, 3, 1, 1, 2).$$
$$\underbrace{}_{\ell\text{-times}}$$

3.2. Economical Way to Construct a Set A with Fixed $|A + A| - |A - A|$

We show another application of our large family $\mathcal{F}$ of MSTD sets, which is to construct sets with a fixed difference $|A + A| - |A - A|$ economically, i.e., with a relatively small width between their maximum and minimum elements.

Proof of Theorem 1.9. Fix $x \in \mathbb{N}$. If x is even, pick $k = 1$ and $\ell = x/2$ for sets in Item 1 of Theorem 1.6; then we find a set $A = S_{k,\ell}$ with

$$|S_{k,\ell} + S_{k,\ell}| - |S_{k,\ell} - S_{k,\ell}| = 2\ell = x$$

and $\max A = 9 + 8\ell = 9 + 4x$, $\min A = 0$.

If x is odd, picking $k = 1$ and $\ell = (x+1)/2$ for sets in Item 2 of Theorem 1.6, we find a set $A = S'_{k,\ell}$ with

$$|S'_{k,\ell} + S'_{k,\ell}| - |S'_{k,\ell} - S'_{k,\ell}| = 2\ell - 1 = x$$

and $\max A = 8 + 8\ell = 12 + 4x$, $\min A = 0$.

Hence, for any positive integer x, there exists $A \subseteq [0, 12 + 4x]$ such that $|A + A| - |A - A| = x$. $\qquad \square$

Remark 3.1. Linear growth of the interval containing A is the best we can do. To see this, assume that the theorem is true for $A_x \subseteq [0, \phi(x)]$, where $\phi(x)$ is sub-linear.

We have:

$$\lim_{x\to\infty} \frac{|A_x + A_x| - |A_x - A_x|}{\phi(x)} \;=\; \lim_{x\to\infty} \frac{x}{\phi(x)} \;=\; \infty, \tag{2}$$

which is a contradiction, since for all sets $A_x \subseteq [0, \phi(x)]$ for large enough x,

$$\frac{|A_x + A_x| - |A_x - A_x|}{\phi(x)} \leq \frac{|A_x + A_x|}{\phi(x)} < \frac{2\phi(x) + 1}{\phi(x)} \leq 3.$$

3.3. Small Fringe Size Generator – Proof of Theorem 1.10

Many classes of MSTD sets can be generated by finding a good fringe pair, i.e. the two sets of elements on the leftmost and rightmost sides of the interval $[0, n]$. Examples can be found the proofs of Theorem 8 in [2], Theorem 1.4 in [1], Theorem 1 in [7], Theorem 1.1 in [8] and Theorem 17 in [12]. Often, when shifted close to each other, two sets in a fringe pair form an MSTD set. However, these fringe pairs have been found by brute force and there has not been a systematic way to generate fringes. It turns out that $\mathcal{F}$ can be a good fringe generator; we demonstrate this by improving the lower bound for the proportion of RSD sets of $\{0, 1, \ldots, n - 1\}$ mentioned in Theorem 17 [2].

In particular, Pennman and Wells used a fringe pair of size 120 generated by the fringe pair used in [7]. The method is to repeat blocks of sets, which inefficiently creates a small lower bound of about 10^{-37}. The authors mentioned that Zhao's techniques can be modified to improve the result; however, this task requires a substantial computation. We believe that this is true since RSD sets are much rarer than MSTD sets[7]. As Zhao's technique relies on extensive search for fringe pairs, the technique is much less effective when applied to RSD sets. Therefore, a feasible and simple way to improve the bound is to find a better fringe pair. Here is a fringe pair generated by $\mathcal{F}$ (we use L and U to match the notations with [12]):

$$
\begin{aligned}
L \;&=\; (0 \mid 1, 1, 2, 1, 4, 3, 1, 4, 3, 1, 4, 3, 1, 4, 3, 1, 1, 1) \\
&=\; \{0, 1, 2, 4, 5, 9, 12, 13, 17, 20, 21, 25, 28, 29, 33, 36, 37, 38, 39\}, \\
U \;&=\; (n - 41) + (0 \mid 1, 1, 1, 1, 4, 3, 1, 4, 3, 1, 4, 3, 1, 4, 3, 1, 1, 2, 1) \\
&=\; (n - 41) + \{0, 1, 2, 3, 4, 8, 11, 12, 16, 19, 20, 24, 27, 28, 32, 35, 36, 37, 39, 40\} \\
&=\; n - \{41, 40, 39, 38, 37, 33, 30, 29, 25, 22, 21, 17, 14, 13, 9, 6, 5, 4, 2, 1\}.
\end{aligned}
$$

Observe that the fringe pair is formed by the MSTD set $S_{1,5}$. We have:

$$
\begin{aligned}
L \hat{+} L \;&=\; [0, 78] \backslash \{0, 8, 78\}, \\
U \hat{+} U \;&=\; [n - 41, n + 38], \\
U \hat{+} U \;&=\; [2n - 82, 2n - 2] \backslash \{2n - 2, 2n - 4, 2n - 12, 2n - 82\}.
\end{aligned}
$$

---

[7]Exhaustive computer search shows that there are no RSD subsets of $[0, 29]$, while there are at least $4.5 \cdot 10^5$ MSTD sets in the same interval.

Notice that $U - L$ misses $\pm(n - 12), \pm(n - 20), \pm(n - 28), \pm(n - 36)$. Hence, $S - S$ misses at least 8 numbers. If we can guarantee that $S \hat{+} S$ misses only 7 elements in $\{0, 8, 78, \pm(2n - 82), \pm(2n - 12), \pm(2n - 4), \pm(2n - 2)\}$, then S is RSD. Following the proof of Theorem 17 in [12], we find a lower bound of

$$(1 - 8(2^{-19} + 2^{-20})) \cdot 2^{-(40+41)} = 4.135 \cdot 10^{-25}. \tag{3}$$

This improvement comes from the reduction in fringe size from 120 to 81. Can we find a better bound for the proportion of RSD subsets of $\{0, 1, 2, \ldots, n - 1\}$ as $n \to \infty$? Since there are no RSD subsets in $[0, 29]$, if we look for a better fringe pair, which is built from an RSD set, the fringe must be of size at least 31. Then the best lower bound that can be achieved by this method is about $2^{-31} \approx 10^{-10}$.

4. Observation: Interior Block Sizes and the Growth of $|A + A| - |A - A|$

Spohn [16] was the first to conceptualize interior blocks existing within MSTD sets. He noted that the repetition of interior blocks may increase the cardinality of the sum set by more than the cardinality of the difference set. For a set A having an interior block B_A, let T_A measure how much more the sum set increases compared to the difference set when B_A is repeated. We observe the relationship between T_A and $|B_A|$. Theorem 1.6 gives us the theorem.

Theorem 4.1. *The following results about $T_A/|B_A|$ are true.*

1. *There exists a set A such that $T_A/|B_A| = 0$.*

2. *For any $\varepsilon > 0$, there exists a set A with $0 < T_A/|B_A| < \varepsilon$.*

3. *For any $0 \le \varepsilon < 0.2$, there exists a set A such that $T_A/|B_A| > 1 + \varepsilon$.*

Proof.

1. Consider the set $S_{1,1}$ in Theorem 1.6. Notice that $|S_{2,1} + S_{2,1}| - |S_{2,1} - S_{2,1}| = |S_{1,1} + S_{1,1}| - |S_{1,1} - S_{1,1}| = 2$ and so, $T_{S_{1,1}} = 0$. In other words, repeating 4 does not change $|S_{1,1} + S_{1,1}| - |S_{1,1} - S_{1,1}|$. This proves (1).

2. Pick $\varepsilon > 0$ and choose k such that $2/(k + 2) < \varepsilon$. Consider the set $S_{k,1}$ in Theorem 1.6. We have $(|S_{k,2} + S_{k,2}| - |S_{k,2} - S_{k,2}|) - (|S_{k,1} + S_{k,1}| - |S_{k,1} + S_{k,1}|) = 2$ and so, $T_{S_{k,1}} = 2$. Hence, $0 < T_{S_{k,1}}/|B_{S_{k,1}}| = 2/(k + 2) < \varepsilon$. This proves (2).

3. Finally, Theorem 12 in [12] shows that $(|Q_3 + Q_3| - |Q_3 - Q_3|) - (|Q_2 + Q_2| - |Q_2 - Q_2|) = 6$, while the interior block is 1,4,4,4,3 (of size 5.) Hence, $T_{Q_2}/|B_{Q_2}| = 6/5 = 1.2$. This proves (3). $\square$

We care about the relationship between the interior block size and the growth of the sum set with respect to the difference set, because this relationship seems to be closely related to sets A with large $f(A)$. The previous record A_{15} in [2] has $T_{A_{15}}/|B_{A_{15}}| = 2/3$, the highest known at that time. The new record Q_{10} in [12] has $T_{Q_{10}}/|B_{Q_{10}}| = 6/5$, which is much higher and this explains why the current record $(1.03059\ldots)$ is much higher than the old record of $(1.0208\ldots)$.

5. Smallest Cardinality for RSD sets

Hegarty [2] proved that the smallest MSTD sets have size 8, and there is exactly one such set up to affine transformation. The method is to reduce the problem to finite computations and run through all possible cases by computer. As commented in [2], this method is not feasible in finding all possible MSTD sets of cardinality 9 since there are many pairs of possible equal differences for a set of 9 random numbers. However, Penman and Wells [2] proved that the list of size-9 MSTD sets given by [2] is exhaustive (up to affine transformation). They also observed that the smallest cardinality of RSD sets must be in the interval $[10, 16]$. We narrow this range of possible values for the size of the smallest RSD sets.

Theorem 5.1. *The smallest cardinality of RSD sets is in the interval $[10, 15]$. Furthermore there are no RSD subsets of $[0, 29]$ and the smallest diameter of an RSD set is 30.*

There are exactly 6 RSD sets in $[0, 30]$, and they all have cardinality 15:

$$
\begin{aligned}
C_1 &= \{0, 1, 2, 3, 6, 8, 13, 16, 18, 23, 24, 26, 28, 29, 30\}, \\
C_2 &= \{0, 1, 2, 3, 6, 9, 14, 15, 17, 22, 23, 26, 28, 29, 30\}, \\
C_3 &= \{0, 1, 2, 4, 5, 8, 9, 14, 18, 21, 22, 26, 27, 28, 30\}, \\
C_4 &= 30 - C_1, \\
C_5 &= 30 - C_2, \\
C_6 &= 30 - C_3.
\end{aligned}
$$

For all $1 \leq i \leq 6$, $|C_i + C_i| - |C_i - C_i| = 1$.

6. Open Questions

We end with these open questions:

1. What are the possible values of $T_A/|B_A|$ over all sets A? Is there a set A such that $T_A/|B_A| > 1.2$? This may lead to an increase in the highest known value of $\log|A + A|/\log|A - A|$.

2. Does $T_A/|B_A| > 0$ imply that $|B_A| \geq 3$?

3. Is Conjecture 1.3 correct?

4. Can we formalize the concept of interior blocks? When do interior blocks exist?

5. Can we find a better bound for the proportion of RSD subsets of $[0, n-1]$ as $n \to \infty$?

6. What is the size of the smallest RSD sets? Is there a better way to find out this number than Hegarty's method (which requires large computing power)?

References

[1] M. Asada, S. Manski, S. J. Miller, and H. Suh, Fringe pairs in generalized MSTD sets, *Int. J. Number Theory* **13** (2017), 2653-2675.

[2] P. V. Hegarty, Some explicit constructions of sets with more sums than differences, *Acta Arith.* **130** (2007), 61-77.

[3] P. V. Hegarty and S. J. Miller, When almost all sets are difference dominated, *Random Structures Algorithms* **35** (2009), 118-136.

[4] G. Iyer, O. Lazarev, S. J. Miller, and L. Zhang, Finding and counting MSTD sets, in *Combinatorial and additive number theory–CANT 2011 and 2012*, Springer-Verlag, New York, 2014, pp. 79-98.

[5] G. Iyer, O. Lazarev, S. J. Miller, and L. Zhang, Generalized more sums than differences sets, *J. Number Theory* **132** (2012), 1054-1073.

[6] J. Marica, On a conjecture of Conway, *Canad. Math. Bull.* **12** (1969), 233-234.

[7] G. Martin and K. O'Bryant, Many sets have more sums than differences, in *Additive Combinatorics*, Providence, RI, 2007, pp. 287-305.

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

[9] S. J. Miller, S. Pegado, and L. Robinson, Explicit Constructions of Large Families of Generalized More Sums Than Differences Sets, *Integers* **12** (2012), #A30.

[10] M. B. Nathanson, Problems in additive number theory I, in *Additive combinatorics*, Providence, RI, 2007, pp. 263-270.

[11] M. B. Nathanson, Sets with more sums than differences, *Integers* **7** (2007), #A5.

[12] D. Penman and M. Wells, On sets with more restricted sums than differences, *Integers* **13** (2013), #A57.

[13] I. Z. Ruzsa, On the cardinality of $A + A$ and $A - A$, in *Combinatorics Year*, North-Holland-Bolyai Tàrsulat, Keszthely, 1978, pp. 933-938.

[14] I. Z. Ruzsa, Sets of sums and differences, in *Séminaire de Théorie des Nombres de Paris*, Birkhäuser, Boston, 1984, pp. 267-273.

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

[16] W. G. Spohn, On Conway's conjecture for integer sets, *Canad. Math. Bull* **14** (1971), 461-462.

[17] Y. Zhao, Constructing MSTD sets using bidirectional ballot sequences, *J. Number Theory* **130** (2010), 1212-1220.

[18] Y. Zhao, Sets characterized by missing sums and differences, *J. Number Theory* **131** (2011), 2107-2134.