WHEN SETS CAN AND CANNOT HAVE MSTD SUBSETS

NATHAN MCNEW, STEVEN J. MILLER, VICTOR XU, AND SEAN ZHANG

ABSTRACT. A finite set of integers A is a More Sums Than Differences (MSTD) set if $|A+A| > |A-A|$. While almost all subsets of $\{0, \ldots, n\}$ are not MSTD, interestingly a small positive percentage are. We explore sufficient conditions on infinite sets of positive integers such that there are either no MSTD subsets, at most finitely many MSTD subsets, or infinitely many MSTD subsets. In particular, we prove no subset of the Fibonacci numbers is an MSTD set, establish conditions such that solutions to a recurrence relation have only finitely many MSTD subsets, and show there are infinitely many MSTD subsets of the primes.

CONTENTS

1. INTRODUCTION

For any finite set of natural numbers $A \subset \mathbb{N}$, we define the sumset

$$A + A \ := \ \{a + a' : a, a' \in A\} \tag{1.1}$$

and the difference set

$$A - A \ := \ \{a - a' : a, a' \in A\}; \tag{1.2}$$

A is called an More Sums Than Differences (MSTD) set if $|A+A| > |A-A|$ (if the two cardinalities are equal it is called balanced, and otherwise difference dominated). As addition is commutative and subtraction is not, it was natural to conjecture that MSTD sets are rare. Conway gave the first example of such a set, $\{0, 2, 3, 4, 7, 11, 12, 14\}$, and this is the smallest such set. Later authors constructed infinite families, culminating in

Date: August 17, 2016.
2000 *Mathematics Subject Classification.* 11P99 (primary), 11K99 (secondary).
The second named author was partially supported by NSF grants DMS1265673 and DMS1561945. We thank the students from the Math 21-499 Spring '16 research class at Carnegie Mellon and the participants from CANT 2016 for many helpful conversations, and Angel Kumchev for comments on an earlier draft.

the work of Martin and O'Bryant, which proved a small positive percentage of subsets of $\{0,\dots,n\}$ are MSTD as $n \to \infty$, and Zhao, who estimated this percentage at around $4.5 \cdot 10^{-4}$. See [FP, He, HM, Ma, MO, Na1, Na2, Na3, Ru1, Ru2, Zh3] for general overviews, examples, constructions, bounds on percentages and some generalizations, [MOS, MPR, MS, Zh1] for some explicit constructions of infinite families of MSTD sets, and [DKMMW, DKMMWW, MV, Zh2] for some extensions to other settings.

Much of the above work looks at finite subsets of the natural numbers, or equivalently subsets of $\{0,1,\dots,n\}$ as $n \to \infty$. We investigate the effect of restricting the initial set on the existence of MSTD subsets. In particular, given an infinite set $A = \{a_k\}_{=1}^{\infty}$, when does A have no MSTD subsets, only finitely many MSTD subsets, or infinitely many MSTD subsets?

Our first result shows that if the sequence grows sufficiently rapidly and there are no 'small' subsets which are MSTD, then there are no MSTD subsets.

Theorem 1.1. *Let $A := \{a_k\}_{k=1}^{\infty}$ be a sequence of natural numbers. If there exists a positive integer r such that*

(1) $a_k > a_{k-1} + a_{k-r}$ *for all* $k \geq r+1$, *and*
(2) A *does not contain any MSTD set* S *with* $|S| \leq 2r+1$,

then A contains no MSTD set.

We prove this in §2. As the smallest MSTD set has 8 elements (see [He]), the second condition is trivially true if $r \leq 3$. In particular, we immediately obtain the following interesting result.

Corollary 1.2. *No subset of the Fibonacci numbers $\{0,1,2,3,5,8,\dots\}$ is an MSTD set.*

The proof is trivial, and follows by taking $r = 3$ and noting

$$F_k \ = \ F_{k-1} + F_{k-2} \ > \ F_{k-1} + F_{k-3} \tag{1.3}$$

for $k \geq 4$.

We now present a partial result on when there are at most finitely many MSTD subsets. For an MSTD set S, we call S a special MSTD set if $|S+S| - |S-S| \geq |S|$. Note if S is a special MSTD set then if $S' = S \cup \{x\}$ for any sufficiently large x then S' is also an MSTD set. We have the following result about a sequence having at most finitely many MSTD sets (see §A for the proof).

Theorem 1.3. *Let $A := \{a_k\}_{k=1}^{\infty}$ be a sequence of natural numbers. If there exists an integer s such that the sequence $\{a_k\}$ satisfies*

(1) $a_k > a_{k-1} + a_{k-3}$ *for all* $k \geq s$, *and*
(2) $\{a_1,\dots,a_{4s+6}\}$ *has no special MSTD subsets,*

then A contains at most finitely many MSTD sets.

The above results concern situations where there are not many MSTD sets; we end with an example of the opposite behavior.

Theorem 1.4. *There are infinitely many MSTD subsets of the primes.*

We will see later that this result follows immediately from the Green-Tao Theorem [GT], which asserts that the primes contain arbitrarily long; unfortunately, such an argument is wasteful as we almost surely have to look at a longer sequence of the primes to find an MSTD set than is needed. We show in §3 that assuming the Hardy-Littlewood conjecture (see Conjecture 3.1) holds, we are able to find such subsets far earlier.

2. Subsets with no MSTD sets

We prove Theorem 1.1, establishing a sufficient condition to ensure the non-existence of MSTD subsets.

Proof of Theorem 1.1. Let $S = \{s_1, s_2, \dots, s_k\} = \{a_{g(1)}, a_{g(2)}, \dots, a_{g(k)}\}$ be a finite subset of A, where $g : \mathbb{Z}^+ \to \mathbb{Z}^+$ is an increasing function. We show that S is not an MSTD set by strong induction on $g(k)$.

We know from [He] that all MSTD sets have at least 8 elements, so S is not an MSTD set if $k \leq 7$; in particular, S is not an MSTD set if $g(k) \leq 7$.

We proceed by induction. Assume for $g(k) \geq 8$ that all S' of the form $\{s_1, \dots, s_{k-1}\}$ with $s_{k-1} < a_{g(k)}$ are not MSTD sets. The proof is completed by showing

$$S \ := \ S' \cup \{a_{g(k)}\} \ = \ \{s_1, \dots, s_{k-1}, a_{g(k)}\} \tag{2.1}$$

is not MSTD sets for any $a_{g(k)}$.

We know that S' is not an MSTD set. Also, if $k \leq 2r+1$ then $|S| \leq 2r+1$ and S is not an MSTD set by the second assumption of the theorem. If $k \geq 2r+2$, consider the number of new sums and differences obtained by adding $a_{g(k)}$. As we have at most $k+1$ new sums, the proof is completed by showing there are at least $k+1$ new differences.

Since $k \geq 2r+2$, we have $k - \lfloor \frac{k+3}{2} \rfloor \geq r$. Let $t = \lfloor \frac{k+3}{2} \rfloor$. Then $t \leq k - r$, which implies $s_t \leq s_{k-r}$. The largest difference in absolute value between elements in S is $s_{k-1} - s_1$; we now show that we have added at least $k+1$ distinct differences greater than $sk - 1 - s_1$ in absolute value, completing the proof.

$$\begin{aligned} a_{g(k)} - s_t \ &\geq \ a_{g(k)} - s_{k-r} \ = \ a_{g(k)} - a_{g(k-r)} \\ &\geq \ a_{g(k)} - a_{g(k)-r} \\ &> \ a_{g(k)-1} \ \geq \ a_{g(k)-1} - a_1 \qquad \text{(by the first assumption on } \{a_n\}) \\ &\geq \ s_{k-1} - a_1 \ \geq \ s_{k-1} - s_1. \end{aligned} \tag{2.2}$$

Since $a_{g(k)} - s_t \geq s_{k-1} - s_1$, we know that

$$a_{g(k)} - s_t, \ \dots, \ a_{g(k)} - s_2, \ a_{g(k)} - s_1$$

are t differences greater than the greatest difference in S'. As we could subtract in the opposite order, S contains at least

$$2t \ = \ 2\left\lfloor \frac{k+3}{2} \right\rfloor \ > \ 2 \cdot \frac{k+1}{2} \ = \ k+1 \tag{2.3}$$

new differences. Thus $S+S$ has at most $k+1$ more sums than $S'+S'$ but $S-S$ has at least $k+1$ more differences compared to $S'-S'$. Since S' is not an MSTD set, we see that S is not an MSTD set. □

We end with an immediate corollary.

Corollary 2.1. *Let $A := \{a_k\}_{k=1}^{\infty}$ be a sequence of natural numbers. If $a_k > a_{k-1}+a_{k-3}$ for all $k \geq 4$, then A contains no MSTD subsets.*

Proof. From [He] we know that all MSTD sets have at least 8 elements. When $r = 3$ the second condition of Theorem 1.1 holds, completing the proof. □

For another example, we consider shifted geometric progressions.

Corollary 2.2. *Let $A = \{a_k\}_{k=1}^{\infty}$ with $a_k = cr^k + d$ for all $k \geq 1$, where $0 \neq c \in \mathbb{N}$, $d \in \mathbb{N}$, and $1 < r \in \mathbb{N}$. Then A contains no MSTD subsets.*

Proof. Without loss of generality we may shift and assume $d = 0$ and $c = 1$; the result now follows immediately from simple algebra. □

3. MSTD SUBSETS OF THE PRIME NUMBERS

We now investigate MSTD subsets of the primes. While Theorem 1.4 follows immediately from the Green-Tao theorem, we first conditionally prove there are infinitely many MSTD subsets of the primes as this argument gives a better sense of what the 'truth' should be (i.e., how far we must go before we find MSTD subsets).

3.1. Admissible Prime Tuples and Prime Constellations.

We first consider the idea of prime m-tuples. A prime m-tuple $(b_1, b_2, \dots, b_m)$ represents a pattern of differences between prime numbers. An integer n matches this pattern if $(b_1+n, b_2+n, \dots, b_m+n)$ are all primes.

A prime m-tuple $(b_1, b_2, \dots, b_m)$ is called admissible if for all integers $k \geq 2$, $\{b_1, b_2, \dots, b_m\}$ does not cover all values modulo k. If a prime m-tuple is not admissible, it's easy to see that it can be matched by at most finitely many primes: whenever $n > k$, one of $(b_1 + n, b_2 + n, \dots, b_m + n)$ is divisible by k and greater than k, so it's not a prime.

It is conjectured in [HL] that all admissible m-tuples are matched by infinitely many integers.

Conjecture 3.1 (Hardy-Littlewood)**.** *Let $b_1, b_2, \dots, b_m$ be m distinct integers, and $P(x; b_1, b_2, \dots, b_m)$ the number of integers $1 \leq n \leq x$ such that $\{n+b_1, n+b_2, \dots, n+b_m\}$ consists wholly of primes. Then*

$$P(x) \ \sim\ \mathfrak{S}(b_1, b_2, \dots, b_m) \int_2^x \frac{du}{(\log u)^m}$$

when $x \to \infty$, where

$$\mathfrak{S}(b_1, b_2, \dots, b_m) \ =\ \prod_{p \geq 2} \left(\left(\frac{p}{p-1} \right)^{m-1} \frac{p-v}{p-1} \right)$$

is a product over all primes, and $v = v(p; b_1, b_2, \dots, b_m)$ is the number of distinct residues of $b_1, b_2, \dots b_m$ to modulus p.

We note that when $(b_1, b_2, \cdots, b_m)$ is an admissible m-tuple, $v(p; b_1, b_2, \dots, b_m)$ is never equal to p, so $\mathfrak{S}(b_1, b_2, \dots, b_m)$ is positive. Therefore this conjecture implies that every admissible m-tuple is matched by infinitely many integers.

3.2. **Infinitude of MSTD subsets of the primes.** We now show the Hardy-Littlewood conjecture implies there are infinitely many subsets of the primes which are MSTD sets.

Theorem 3.2. *If the Hardy-Littlewood conjecture holds for all admissible m-tuples then the primes have infinitely many MSTD subsets.*

Proof. Consider the smallest MSTD set $S = \{0, 2, 3, 4, 7, 11, 12, 14\}$. We know that $\{p, p + 2s, p + 3s, p + 4s, p + 7s, p + 11s, p + 12s, p + 14s\}$ is an MSTD set for all positive integers p, s. Set $s = 30$, we deduce that if infinitely many primes match the 8-tuple $T = (0, 60, 90, 120, 210, 330, 360, 420)$, then there are infinitely many primes on MSTD sets.

We check that T is an admissible prime 8-tuple. When $m > 8$, the eight numbers in T clearly don't cover all values modulo m. When $m \leq 8$, we show by computation that T does not cover all values modulo m.

By Conjecture 3.1, there are infinitely many integers p such that $\{p, p+60, p+90, p+120, p+210, p+330, p+360, p+420\}$ contains all primes. These are all MSTD sets, so there are infinitely many MSTD sets on primes. □

Of course, all we need is that the Hardy-Littlewood conjecture holds for one admissible m-tuple which has an MSTD subset. We may take $p = 19$, which gives an explicit MSTD subset of the primes: $\{19, 79, 109, 139, 229, 349, 379, 439\}$ (a natural question is what is the smallest MSTD subset of the primes). If one wishes, one can use the conjecture to get some lower bounds on the number of MSTD subsets of the primes at most x. The proof of Theorem 1.3 follows similarly.

Proof of Theorem 1.3. By the Green-Tao theorem, the primes contain arbitrarily long arithmetic progressions. Thus for each $N \geq 14$ there are infinitely many pairs (p, d) such that

$$\{p, p + d, p + 2d, \ldots, p + Nd\} \tag{3.1}$$

are all prime. We can then take subsets as in the proof of Theorem 3.2. □

Appendix A. Subsets with Finitely Many MSTD Sets

We start with some properties of special MSTD sets, and then prove Theorem 1.3. The arguments are similar to those used in proving Theorem 1.1.

A.1. **Special MSTD Sets.** Recall an MSTD set S is special if $|S+S| - |S-S| \geq |S|$. For any $x \geq 2\sum_{s\in S}|s|$, adding x creates $|S| + 1$ new sums and $2|S|$ new differences. Let $S^* = S \cup \{x\}$. Then

$$|S^* + S^*| - |S^* - S^*| \ \geq\ |S| + (|S| + 1) - 2|S| \ =\ 1, \tag{A.1}$$

and S^* is also an MSTD set. Hence, from one special MSTD set $S \subset \{a_n\}_{n=1}^{\infty} =: A$, we can generate infinitely many MSTD sets by adding any large integer in A.

Conversely, if a set is not a special MSTD set, then $|S + S| - |S - S| < |S|$, and by adding any large $x \geq 2\sum_{s\in S}|s|$, $S \cup \{x\}$ has at least as many differences as sums. Thus only finitely many MSTD sets can be generated by appending one integer from A to S.

Note that special MSTD sets exist. Consider the smallest MSTD set $S = \{0, 2, 3, 4, 7, 11, 12, 14\}$. Using the method of base expansion, described in [He], we are able to obtain S_3 containing $|S_3| = 8^3 = 512$ elements, such that $|S_3+S_3| = |S+S|^3 = 26^3 = 17576$, and $|S_3 - S_3| = |S-S|^3 = 25^3 = 15625$. Then $|S_3+S_3| - |S_3-S_3| > |S_3|$.

A.2. **Finitely Many MSTD Sets on a Sequence.** If a sequence $A = \{a_n\}_{n=1}^{\infty}$ contains a special MSTD set S, then we can get infinitely many MSTD subsets on the sequence just by adding sufficiently large elements of A to S. Therefore for a sequence A to have at most finitely many MSTD subsets, it is necessary that it has no special MSTD sets. Using the result from the previous subsection, we can prove Theorem 1.3.

We establish some notation before turning to the proof. We can write A as the union of $A_1 = \{a_1, \dots, a_{s-1}\}$ and $A_2 = \{a_s, a_{s+1}, \dots\}$. By Corollary 2.1, we know that A_2 contains no MSTD sets. So any MSTD set must contain some elements from A_1.

We first prove a lemma about A_2.

Lemma A.1. *Let $S' = \{s_1, \dots, s_{k-1}\}$ be a subset of A containing at least 3 elements $a_{r_1}, a_{r_2}, a_{r_3}$ in A_2, with $r_3 > r_2 > r_1$. Let $\varphi(k) > r_3$, and let $S = S' \cup \{a_{\varphi(k)}\}$. Then either S is not an MSTD set, or S satisfies $|S-S| - |S+S| > |S'-S'| - |S'+S'|$.*

Proof. We follow a similar argument as in Theorem 1.1.

If $k \le 7$, then S is not an MSTD set.

If $k \ge 8$, then $k - \lfloor \frac{k+3}{2} \rfloor \ge 3$. Let $t = \lfloor \frac{k+2}{2} \rfloor$. Then $t \le k-3$, and $s_t \le s_{k-3}$, and

$$\begin{aligned} a_{\varphi(k)} - s_t &\ge a_{\varphi(k)} - s_{k-3} = a_{\varphi(k)} - a_{\varphi(k-3)} \\ &\ge a_{\varphi(k)} - a_{\varphi(k)-3} \\ &> a_{\varphi(k)-1} = a_{\varphi(k)-1} - a_1 \qquad \text{(by assumption on } a) \\ &\ge s_{k-1} - a_1 \ge s_{k-1} - s_1. \end{aligned} \tag{A.2}$$

In the set S', the greatest difference is $s_{k-1} - s_1$. Since $a_{\varphi(k)} - s_t \ge s_{k-1} - s_1$, we know that $a_{\varphi(k)} - s_t, \dots, a_{\varphi(k)} - s_2, a_{\varphi(k)} - s_1$ are all differences greater than the greatest difference in S'.

By a similar argument, $s_t - a_{\varphi(k)}, \dots, s_2 - a_{\varphi(k)}, s_1 - a_{\varphi(k)}$ are all differences smaller than the smallest difference in S'.

So S contains at least $2t = 2\lfloor \frac{k+3}{2} \rfloor > 2 \cdot \frac{k+1}{2} = k+1$ new differences compared to S', and S satisfies

$$|S-S| - |S+S| \ > \ |S'-S'| - |S'+S'|, \tag{A.3}$$

completing the proof. □

Proof of Theorem 1.3. Let K_0 be an MSTD subset of A_1. Let K_n be an MSTD subset of A with n elements in A_2, in the form $S \cup \{a_{r_1}, \dots, a_{r_n}\}$, where S is a subset of A_1 and $s \le r_1 < r_2 < \cdots < r_n$. Let S_n be any (not necessarily MSTD) subset of A in the same form.

The lemma tells us that for any K_n with $n \ge 3$, when we add any new element $a_{r_{n+1}}$ to get S_{n+1}, either S_{n+1} is not an MSTD set, or $|S_{n+1} - S_{n+1}| - |S_{n+1} + S_{n+1}| \ge |K_n - K_n| - |K_n + K_n| + 1$.

Let

$$d \ = \ \max_{K_3}(|K_3+K_3| - |K_3-K_3|, 1). \tag{A.4}$$

Then for all $n > d+3$, consider any S_n. For $3 \le k \le n$, define S_k as the set obtained by deleting the $(n-k)$ largest elements from S_n.

If S_k is not an MSTD set for any $k \geq 3$, by Lemma A.1 either S_{k+1} is not an MSTD set, or $|S_{k+1} - S_{k+1}| - |S_{k+1} + S_{k+1}| > |S_k - S_k| - |S_k + S_k| \geq 0$, in which case S_{k+1} is also not an MSTD set. Assume that S_n is an MSTD set, and the previous argument shows that S_{n-1} to S_3 must all be MSTD sets, and we have

$$|S_n - S_n| - |S_n + S_n| \; > \; |S_3 - S_3| - |S_3 + S_3| + d \; > \; 0, \tag{A.5}$$

since S_3 is one of the K_3's. Then S_n is not an MSTD set, a contradiction.

Therefore the previous assumption is false, and S_n is not an MSTD set for all $n > d+3$. This means that every MSTD set on A is one of $K_0, K_1, K_2, \ldots, K_{d+3}$.

For $n \geq 0$, let k_n be the number of all possible K_n. We can show by induction that

- k_n is finite for all $n \geq 0$, and
- every K_n is not a special MSTD set.

We have 4 base cases.

We know that A_1 is finite, so k_0, which is the number of MSTD subsets of A_1, is finite.

Any K_0 is a subset of $\{a_1, \ldots, a_{s-1}\}$, which is a subset of $A' = \{a_1, \ldots, a_{4s+6}\}$. So such K_0 is not a special MSTD set.

Consider the index $4s$. We claim that

$$a_{4s} \; > \; \sum_{a \in A_1} a. \tag{A.6}$$

This is because

$$\begin{aligned} \sum_{a \in A_1} a \; &< \; s \cdot a_s \\ &< \; \frac{s}{2}\left(a_s + a_{s+2}\right) \; < \; \frac{s}{2} \cdot a_{s+3} \\ &< \; \frac{s}{4}\left(a_{s+3} + a_{s+5}\right) \; < \; \frac{s}{4} \cdot a_{s+6} \ldots \\ &< \; a_{s+3\log_2(s)} \\ &< \; a_{s+3s} \; = \; a_{4s} \end{aligned}$$

Therefore for all $r_1 \geq 4s$,

$$a_{r_1} \; \geq \; a_{4s} \; > \; \sum_{a \in A_1} a. \tag{A.7}$$

Consider any S_1 with $r_1 \geq 4s$. It contains a set of elements $S = \{s_1, \ldots, s_m\}$ in A_1 and a_{r_1} in A_2. We know that $\sum_{s \in S} s < a_{r_1}$, and we also know that S is not a special MSTD set. So $S_1 \; = \; S \cup \{a_{r_1}\}$ is not an MSTD set.

Therefore for S_1 to be an MSTD set, r_1 must be smaller than $4s$. We conclude that k_1 is finite.

Then any K_1 is a subset of $\{a_1, \ldots, a_{4s}\}$, which is a subset of A'. Hence, such K_1 is not a special MSTD set.

Consider the index $4s+3$. For all $r_2 \geq 4s+3$,

$$a_{r_2} - a_{r_2-1} \; > \; a_{r_2-3} \; \geq \; a_{4s} \; > \; \sum_{a \in A_1} a. \tag{A.8}$$

Consider any S_2 with $r_2 \geq 4s+3$. It contains some elements $S = \{s_1, \dots, s_m\}$ in A_1 and r_1, r_2 in A_2. We have $a_{r_2} - a_{r_1} \geq a_{r_2} - a_{r_2-1}$. We also have $a_{r_2} - a_{r_2-1} > \sum_{s\in S} s$.

Therefore $a_{r_2} > \left(\sum_{s\in S} s\right) + a_{r_1}$, and $S \cup \{a_{r_1}\}$ is not a special MSTD set. Hence, $S_2 = S \cup \{r_1, r_2\}$ is not an MSTD set.

Then for S_2 to be an MSTD set, r_2 must be smaller than $4s+3$. We conclude that k_2 is finite.

We also know that any K_2 is a subset of $\{a_1, \dots, a_{4s+1}\}$, which is a subset of A'. Therefore such K_2 is not a special MSTD set.

Consider the index $4s+6$. For all $r_3 \geq 4s+6$,

$$a_{r_3-3} - a_{r_3-4} \ > \ a_{r_3-6} \ \geq \ a_{4s} \ > \ \sum_{a\in A_1} a. \tag{A.9}$$

Consider any S_3 with $r_3 \geq 4s+6$. We write S_3 as $S \cup \{a_{r_1}, a_{r_2}, a_{r_3}\}$. If $|S| < 5$, we know that $|S_3| < 8$, and S_3 is not an MSTD set. We can then assume that $|S| \geq 5$. We have 2 cases.

In the first case, $r_2 \leq r_3 - 3$, so $a_{r_3} - a_{r_2} - a_{r_1} \geq a_{r_3} - a_{r_3-3} - a_{r_3-4} \ \geq \ a_{r_3-1} - a_{r_3-4} \ \geq \ a_{r_3-2} > a_{r_3-6} \ > \ \sum_{s\in S} s$.

We know that $S \cup \{a_{r_1}, a_{r_2}\}$ is not a special MSTD set. So adding a_{r_3} with $a_{r_3} > \left(\sum_{s\in S} s\right) + a_{r_1} + a_{r_2}$ creates a non-MSTD set.

In the second case, $r_2 > r_3 - 3$, so $a_{r_3} - a_{r_2} \geq a_{r_3} - a_{r_3-1} > \sum_{s\in S} s$. Similarly, $a_{r_2} - a_{r_1} > a_{r_3-2} - a_{r_3-3} > \sum_{s\in S} s$.

Therefore the differences between $a_{r_1}, a_{r_2}, a_{r_3}$ are large relative to the elements in S, and $S_3 + S_3$ consists of 4 copies of $S+S$ ($S+S$, $a_{r_1}+S+S$, $a_{r_2}+S+S$, $a_{r_3}+S+S$) plus 5 or 6 more elements, and $S_3 - S_3$ consists of 7 copies of $S - S$ plus 4 or 6 more elements.

We have $|S_3 + S_3| = 4|S+S| + c_1$ and $|S_3 - S_3| = 7|S-S| + c_2$. $|S-S|$ is at least $2|S|-1$. Since S is not a special MSTD set, we know that $|S+S| < |S-S| + |S|$.

Then

$$\frac{|S+S|}{|S-S|} \ < \ 1 + \frac{|S|}{|S-S|} \ \leq \ 1 + \frac{|S|}{2|S|-1} \ \leq \ 1 + \frac{5}{9} \ = \ \frac{14}{9} \ < \ \frac{7}{4}. \tag{A.10}$$

This tells us that $4|S+S| < 7|S-S|$. To determine c_1, c_2, we consider 2 cases.

- $a_{r_2} - a_{r_1} = a_{r_3} - a_{r_2}$. In this case there are 5 more sums and 4 more differences. Since $4|S+S| + 5 \leq 7|S-S| + 4$, we have $|S_3 + S_3| \leq |S_3 - S_3|$. Then S_3 is not an MSTD set.
- $a_{r_2} - a_{r_1} \neq a_{r_3} - a_{r_2}$, and we have 6 more sums and 6 more differences. Since $4|S+S| + 6 < 7|S-S| + 6$, we know that $|S_3 + S_3| < |S_3 - S_3|$, and S_3 is not an MSTD set.

Hence, for S_3 to be an MSTD set, r_3 must be smaller than $4s+6$. We conclude that k_3 is finite.

In addition, any K_3 is a subset of $A' = \{a_1, \dots, a_{4s+6}\}$, so such K_3 is not a special MSTD set.

Assume that k_n is finite for some $n \geq 3$. Equivalently, assume that there exists t_n such that if $r_n \geq t_n$, then any set containing a_{r_n} is not an MSTD set. We can show that k_{n+1} is finite.

Let t_{n+1} be the index such that for all $r_{n+1} \geq t_{n+1}$,

$$a_{r_{n+1}} > \sum_{x<r_n} a_x. \tag{A.11}$$

Consider any $S \subseteq A_1$, and let $S_n = S \cup \{a_{r_1}, \cdots, a_{r_n}\}$ for any $\{a_{r_1}, \cdots, a_{r_n}\}$. Consider adding any $a_{r_{n+1}}$ with $r_{n+1} \geq t_{n+1}$ to S_n. We have two cases.

- If $r_n < t_n$, then S_n is, by inductive hypothesis, not a special MSTD set. So adding $a_{r_{n+1}} > \sum_{x\in S_n} x$ creates a non-MSTD set.
- If $r_n \geq t_n$, then S_n is, by inductive hypothesis, not an MSTD set. So $|S_n - S_n| - |S_n + S_n| > 0$. Since $n \geq 3$, we can apply Lemma A.1, and either S_{n+1} is not an MSTD set, or $|S_{n+1} - S_{n+1}| - |S_{n+1} + S_{n+1}| > |S_n - S_n| - |S_n + S_n| > 0$, in which case S_{n+1} is still not an MSTD set.

We conclude that for all MSTD sets S_{n+1}, we must have $r_{n+1} < t_{n+1}$. So k_{n+1} is finite.

Consider any MSTD set $K_{n+1} = S_n \cup \{a_{r_{n+1}}\}$. Applying lemma A.1 again, we have $|K_{n+1} - K_{n+1}| - |K_{n+1} + K_{n+1}| > |S_n - S_n| - |S_n + S_n|$. We know, from inductive hypothesis, that S_n is not a special MSTD set. Therefore all possible K_{n+1} are not special MSTD sets.

By induction, k_n is finite for all $n \geq 0$, and all K_n are not special MSTD sets.

We have previously shown that every MSTD set on A is one of $K_0, K_1, K_2, \ldots, K_{d+3}$. We now know that there are finitely many sets for each group, so there are finitely many MSTD sets on A. □

References

[DKMMW] T. Do, A. Kulkarni, S. J. Miller, D. Moon and J. Wellens, *Sums and Differences of Correlated Random Sets*, Journal of Number Theory **147** (2015), 44–68.

[DKMMWW] T. Do, A. Kulkarni, S. J. Miller, D. Moon, J. Wellens and J. Wilcox, *Sets Characterized by Missing Sums and Differences in Dilating Polytopes*, Journal of Number Theory **157** (2015), 123–153.

[FP] G. A. Freiman and V. P. Pigarev, *The relation between the invariants R and T*, In: Number theoretic studies in the Markov spectrum and in the structural theory of set addition (Russian), pp. 172–174. Kalinin. Gos. Univ., Moscow (1973).

[GT] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Annals of Mathematics **167** (2008), no. 2, 481–547.

[HL] G. H. Hardy and J. Littlewood, *Some Problems of 'Partitio Numerorum'; III: On the Expression of a Number as a Sum of Primes*, Acta Math. **44**, 1–70.

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

[HM] 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

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

[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.

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

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

[Na3] 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, *Sets of sums and differences*. In: Séminaire de Théorie des Nombres de Paris 1982-1983, pp. 267–273. Birkhäuser, Boston (1984).

[Ru3] 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: nmcnew@towson.edu

DEPARTMENT OF MATHEMATICS, TOWSON UNIVERSITY, TOWSON, MD 21252

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: vzx@andrew.cmu.edu

DEPARTMENT OF MATHEMATICS, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PA 15213

E-mail address: xiaoronz@andrew.cmu.edu

DEPARTMENT OF MATHEMATICS, CARNEGIE MELLON UNIVERSITY, PITTSBURGH, PA 15213