Sums and Differences of Correlated Random Sets

Thao Do (Presenter) - Stony Brook
thao.do@stonybrook.edu
Jake Wellens (Presenter)- Caltech
jwellens@caltech.edu
Archit Kulkarni - Carnegie-Mellon
David Moon - Williams College
Advisor: Steven Miller

2013 YMC. Columbus, Ohio. August 10, 2013

Introduction

Given $A \subset \mathbb{Z}$, let

$$A + A = \{a_1 + a_2 : a_1, a_2 \in A\},$$
$$A - A = \{a_1 - a_2 : a_1, a_2 \in A\}.$$

Introduction

Given $A \subset \mathbb{Z}$, let

$$A + A = \{a_1 + a_2 : a_1, a_2 \in A\},$$
$$A - A = \{a_1 - a_2 : a_1, a_2 \in A\}.$$

Theorem

There exists a positive constant c such that for any n large, the proportion of sets $A \subset \{0, \ldots, n\}$ with $|A + A| > |A - A|$ is greater than c. (Martin and O'Bryant 2006)

Such sets are called *More Sums Than Differences (MSTD) sets*, or *sum-dominant sets*.

Correlated Random Pairs

All of the literature to date has looked at sums and differences of a set *with itself*.

Correlated Random Pairs

All of the literature to date has looked at sums and differences of a set *with itself*.

We investigate sums and differences of *pairs* of subsets $(A, B) \subset \{0, \dots, n\}$. We select such pairs according to the dependent random process:

Correlated Random Pairs

All of the literature to date has looked at sums and differences of a set *with itself*.

We investigate sums and differences of *pairs* of subsets $(A, B) \subset \{0, \ldots, n\}$. We select such pairs according to the dependent random process:

$$P(k \in A) = p; \quad P(k \in B | k \in A) = \rho_1; \quad P(k \in B | k \notin A) = \rho_2$$

Correlated Random Pairs

All of the literature to date has looked at sums and differences of a set *with itself*.

We investigate sums and differences of *pairs* of subsets $(A, B) \subset \{0, \ldots, n\}$. We select such pairs according to the dependent random process:

$$P(k \in A) = p; \quad P(k \in B | k \in A) = \rho_1; \quad P(k \in B | k \notin A) = \rho_2$$

Let $\vec{\rho} = (p, \rho_1, \rho_2)$. We call (A, B) a *$\vec{\rho}$-correlated pair.*

Correlated Random Pairs

- $(\rho_1, \rho_2) = (1, 0) \implies (A, A)$.

Correlated Random Pairs

- $(\rho_1, \rho_2) = (1, 0) \implies (A, A)$.
- $(\rho_1, \rho_2) = (0, 1) \implies (A, A^c)$.

Correlated Random Pairs

- $(\rho_1, \rho_2) = (1, 0) \implies (A, A)$.
- $(\rho_1, \rho_2) = (0, 1) \implies (A, A^c)$.
- $\rho_1 = \rho_2, \implies (A, B)$ independent.

Probability function

Let $P(\vec{\rho}, n)$ be the probability that a $\vec{\rho}$-correlated pair $(A, B) \subset \{0, \dots, n\}$ is MSTD, that is

$$|A + B| > |\pm(A - B)| = |(A - B) \cup (B - A)|$$

Probability function

Let $P(\vec{\rho}, n)$ be the probability that a $\vec{\rho}$-correlated pair $(A, B) \subset \{0, \ldots, n\}$ is MSTD, that is

$$|A + B| > |\pm(A - B)| = |(A - B) \cup (B - A)|$$

Note: $P(1/2, 1, 0, n)$, $P(1/2, 0, 1, n)$, and $P(1/2, 1/2, 1/2, n)$ can be thought of as *proportions* of pairs (A, A), (A, A^c), resp. (A, B) which are MSTD, while other values of $P(\vec{\rho}, n)$ must be thought of as *probabilities.*

Main Results on Correlated Pairs

Theorem

For any $\vec{\rho} \in [0,1]^3$*, the limit*

$$\lim_{n\to\infty} P(\vec{\rho}, n) =: P(\vec{\rho})$$

exists. Moreover, as long as $p \notin \{0,1\}$ *and* $(\rho_1, \rho_2) \neq (0,0), (1,1)$*, then* $P(\vec{\rho})$ *is strictly positive.*

The function $P(\vec{\rho})$

Proof idea:
To show that $P(\vec{\rho})$ exists, we build on the idea of Zhao (2010) and count MSTD pairs by their *minimal fringe profiles*.

The function $P(\vec{\rho})$

Proof idea:
To show that $P(\vec{\rho})$ exists, we build on the idea of Zhao (2010) and count MSTD pairs by their *minimal fringe profiles*.

We call a pair a *rich MSTD pair* if the sumset wins over the difference set near the edges, while both the sumset and the difference set contain all the middle elements.

The function $P(\vec{\rho})$

Proof idea:
To show that $P(\vec{\rho})$ exists, we build on the idea of Zhao (2010) and count MSTD pairs by their *minimal fringe profiles*.

We call a pair a *rich MSTD pair* if the sumset wins over the difference set near the edges, while both the sumset and the difference set contain all the middle elements.

We show that as $n \to \infty$, a $\vec{\rho}$-pair $(A, B) \subset \{0, \ldots, n\}$ which is an MSTD pair is rich with probability 1.

The function $P(\vec{\rho})$

Proof idea:
To show that $P(\vec{\rho})$ exists, we build on the idea of Zhao (2010) and count MSTD pairs by their *minimal fringe profiles*.

We call a pair a *rich MSTD pair* if the sumset wins over the difference set near the edges, while both the sumset and the difference set contain all the middle elements.

We show that as $n \to \infty$, a $\vec{\rho}$-pair $(A, B) \subset \{0, \dots, n\}$ which is an MSTD pair is rich with probability 1.

Thus, by summing the probabilities that the edges of (A, B) have a given MSTD fringe profile and that (A, B) is rich over all such minimal fringe profiles, we can get the limit $P(\vec{\rho})$.

The function $P(\vec{\rho})$

Proof idea:
To show that $P(\vec{\rho})$ exists, we build on the idea of Zhao (2010) and count MSTD pairs by their *minimal fringe profiles*.

We call a pair a *rich MSTD pair* if the sumset wins over the difference set near the edges, while both the sumset and the difference set contain all the middle elements.

We show that as $n \to \infty$, a $\vec{\rho}$-pair $(A, B) \subset \{0, \ldots, n\}$ which is an MSTD pair is rich with probability 1.

Thus, by summing the probabilities that the edges of (A, B) have a given MSTD fringe profile and that (A, B) is rich over all such minimal fringe profiles, we can get the limit $P(\vec{\rho})$.

The function $P(\vec{\rho})$

Proof idea: To show that $P(\vec{\rho})$ is positive, we only need to exhibit a single MSTD fringe profile F such that, (for sufficiently large n) with positive probability, (A, B) is rich with fringe profile F.

The function $P(\vec{\rho})$

Proof idea: To show that $P(\vec{\rho})$ is positive, we only need to exhibit a single MSTD fringe profile F such that, (for sufficiently large n) with positive probability, (A, B) is rich with fringe profile F.

As long as $\rho_1 p \neq 0$, we can use the standard example given by Martin and O'Bryant (2006).

The function $P(\vec{\rho})$

Proof idea: To show that $P(\vec{\rho})$ is positive, we only need to exhibit a single MSTD fringe profile F such that, (for sufficiently large n) with positive probability, (A, B) is rich with fringe profile F.

As long as $\rho_1 p \neq 0$, we can use the standard example given by Martin and O'Bryant (2006).

If $\rho_1 = 0$, but $\rho_2 p > 0$, we can use the fringe profile $L = R = \{1, 2, 3, 5, 7, 8\}$, $L' = R' = L^c$. (This means that the left and the right edges of A look like

$$\{1, 2, 3, 5, 7, 8\}$$

and

$$\{n-1, n-2, n-3, n-5, n-7, n-8\}$$

respectively, while B has complementary fringes).

The function $P(\vec{\rho})$

Proof idea: To show that $P(\vec{\rho})$ is positive, we only need to exhibit a single MSTD fringe profile F such that, (for sufficiently large n) with positive probability, (A, B) is rich with fringe profile F.

As long as $\rho_1 p \neq 0$, we can use the standard example given by Martin and O'Bryant (2006).

If $\rho_1 = 0$, but $\rho_2 p > 0$, we can use the fringe profile $L = R = \{1, 2, 3, 5, 7, 8\}$, $L' = R' = L^c$. (This means that the left and the right edges of A look like

$$\{1, 2, 3, 5, 7, 8\}$$

and

$$\{n-1, n-2, n-3, n-5, n-7, n-8\}$$

respectively, while B has complementary fringes).

The function $P(\vec{\rho})$

Theorem

The function $P(\vec{\rho})$ is continuous on $[0,1]^3$.

The function $P(\vec{\rho})$

Theorem

The function $P(\vec{\rho})$ is continuous on $[0,1]^3$.

Theorem

For any ρ_1, ρ_2, the function $P(p, \rho_1, \rho_2)$ is a differentiable function of $p \in [0,1]$.

Maximizing $P(\vec{\rho})$

As $P(\vec{\rho})$ is continuous on the compact space $[0, 1]^3$, it must attain a maximum.

Maximizing $P(\vec{\rho})$

As $P(\vec{\rho})$ is continuous on the compact space $[0,1]^3$, it must attain a maximum.

[MO] and others have estimated with Monte Carlo experiments that $P(1/2, 1, 0) \approx 0.00045$. Zhao has shown $P(1/2, 1, 0) > 4.286 \cdot 10^{-4}$. From our exhaustive searches, we estimate that $P(1/2, 0, 1) \approx 0.03$. So we conjecture that (A, A^c) beats (A, A) in the limit.

Maximizing $P(\vec{\rho})$

As $P(\vec{\rho})$ is continuous on the compact space $[0,1]^3$, it must attain a maximum.

[MO] and others have estimated with Monte Carlo experiments that $P(1/2,1,0) \approx 0.00045$. Zhao has shown $P(1/2,1,0) > 4.286 \cdot 10^{-4}$. From our exhaustive searches, we estimate that $P(1/2,0,1) \approx 0.03$. So we conjecture that (A, A^c) beats (A, A) in the limit.

For each n, $P_n(\vec{\rho})$ denotes the proportion of *MSTD* pair of subsets of $[1,\dots,n]$. P_n is a polynomial of p, ρ_1, ρ_2 based on the sizes of all *MSTD* pairs and their intersection.

Maximizing $P(\vec{\rho})$

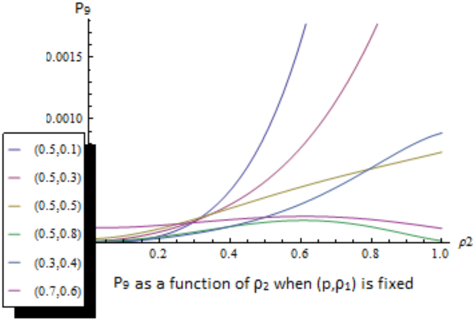
As $P(\vec{\rho})$ is continuous on the compact space $[0, 1]^3$, it must attain a maximum.

[MO] and others have estimated with Monte Carlo experiments that $P(1/2, 1, 0) \approx 0.00045$. Zhao has shown $P(1/2, 1, 0) > 4.286 \cdot 10^{-4}$. From our exhaustive searches, we estimate that $P(1/2, 0, 1) \approx 0.03$. So we conjecture that (A, A^c) beats (A, A) in the limit.

For each n, $P_n(\vec{\rho})$ denotes the proportion of *MSTD* pair of subsets of $[1, \ldots, n]$. P_n is a polynomial of p, ρ_1, ρ_2 based on the sizes of all *MSTD* pairs and their intersection.

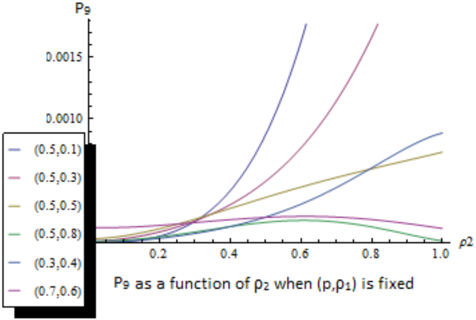
We fix $n = 9$, do an exhaustive search to find all MSTD pairs and calculate P_9.

Fix (p, ρ_1)



P_9 as a function of ρ_2 when (p, ρ_1) is fixed

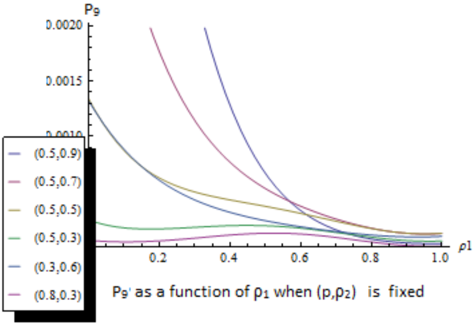
Fix (p, ρ_1)



P_9 as a function of ρ_2 when (p,ρ_1) is fixed

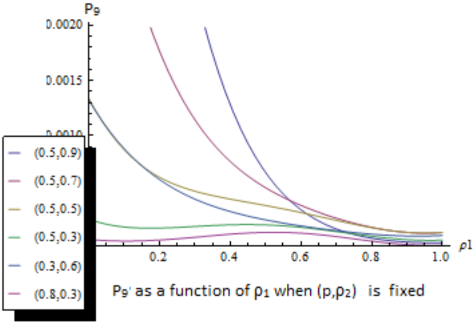
Conjecture 1: For any fixed (p, ρ_1) with ρ_1 not too big ($\rho_1 \leq 0.4$) then P as a function of ρ_2 is strictly increasing in $[0, 1]$ and reaches its maximum at $\rho_2 = 1$.

Fix (p, ρ_2)



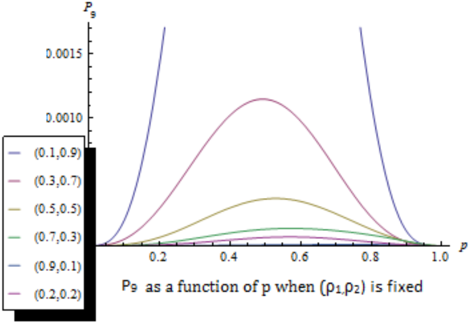
P_9' as a function of ρ_1 when (p, ρ_2) is fixed

Fix (p, ρ_2)



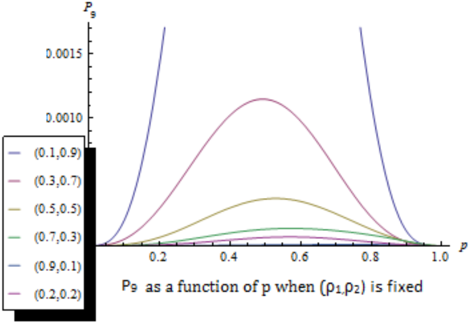
Conjecture 2: For any fixed (p, ρ_2) with ρ_2 not too small ($\rho_2 \geq .5$) then P as a function of ρ_1 is strictly decreasing in $[0, 1]$ and reaches its maximum at $\rho_1 = 0$.

Fix (ρ_1, ρ_2)



P_9 as a function of p when (ρ_1, ρ_2) is fixed

Fix (ρ_1, ρ_2)



P_9 as a function of p when (ρ_1, ρ_2) is fixed

Conjecture 3: For any fixed (ρ_1, ρ_2), P as a function of p in $(0, 1)$ has a maximum at $1/2$.

A and *A* complement

From Conjectures 1 and 2, it makes sense that the maximum of P is at $\rho_1 = 0, \rho_2 = 1$ or when it is the case of A and A^c.

A and A complement

From Conjectures 1 and 2, it makes sense that the maximum of P is at $\rho_1 = 0, \rho_2 = 1$ or when it is the case of A and A^c.

In this case we know that $p = 1/2$ is a critical point of $P(p, 0, 1)$ as $P(p, 0, 1) = P(1 - p, 0, 1)$.

A and A complement

From Conjectures 1 and 2, it makes sense that the maximum of P is at $\rho_1 = 0, \rho_2 = 1$ or when it is the case of A and A^c.

In this case we know that $p = 1/2$ is a critical point of $P(p, 0, 1)$ as $P(p, 0, 1) = P(1 - p, 0, 1)$.

Conjecture 4: The maximum of the function $P(p, \rho_1, \rho_2)$ in $[0, 1]^3$ occurs at $P(1/2, 0, 1) \approx 0.03$.

Some notation

- Big O: $f(n) = O(g(n))$ if $\exists c, n_0 > 0$ s.t $f(n) > cg(n)$ for all $n > n_0$.

Some notation

- Big O: $f(n) = O(g(n))$ if $\exists c, n_0 > 0$ s.t $f(n) > cg(n)$ for all $n > n_0$.
- Big Θ: $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$.

Some notation

- Big O: $f(n) = O(g(n))$ if $\exists c, n_0 > 0$ s.t $f(n) > cg(n)$ for all $n > n_0$.
- Big Θ: $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$.
- Little o: $f(n) = o(g(n))$ if $\lim_{n\to\infty} \dfrac{g(n)}{f(n)} = \infty$.

Some notation

- Big O: $f(n) = O(g(n))$ if $\exists c, n_0 > 0$ s.t $f(n) > cg(n)$ for all $n > n_0$.
- Big Θ: $f(n) = \Theta(g(n))$ if $f(n) = O(g(n))$ and $g(n) = O(f(n))$.
- Little o: $f(n) = o(g(n))$ if $\lim_{n\to\infty} \dfrac{g(n)}{f(n)} = \infty$.
- $X \sim f(N)$ if for any $\epsilon_1, \epsilon_2 > 0$ there exists $N_{\epsilon_1,\epsilon_2} > 0$ such that for all $N > N_{\epsilon_1,\epsilon_2}$

$$P\Big(X \notin [(1-\epsilon_1)f(N), (1+\epsilon_1)f(N)]\Big) < \epsilon_2$$

Taking $p \to 0$

In previous section, we know that for any fixed (p, ρ_1, ρ_2) there is a positive percentage of MSTD pairs.

Taking $p \to 0$

In previous section, we know that for any fixed (p, ρ_1, ρ_2) there is a positive percentage of MSTD pairs.

Here we let some of p, ρ_1, ρ_2 vary and depend on n.

Taking $p \to 0$

In previous section, we know that for any fixed (p, ρ_1, ρ_2) there is a positive percentage of MSTD pairs.

Here we let some of p, ρ_1, ρ_2 vary and depend on n.

Hegarty-Miller investigated this for $(\rho_1, \rho_2) = (1, 0)$ and $p = p(n) = o(1), n^{-1} = o(p(n))$. The first condition indicates p decays with n while the second one guarantees the expected size of A grow with n.

Theorem (Hegarty-Miller)

Given a function $p : \mathbb{N} \to (0, 1)$ such that $p(N) = o(1)$ and $N^{-1} = o(p(N))$. As $N \to \infty$, the probability A as a subset of $[1, \ldots, N]$ is *MSTD* tends to 0. Let $\mathcal{S} = |A + A|, \mathcal{D} = |A - A|$ and $\mathcal{S}^C, \mathcal{D}^C$ denote their complements.

Theorem (Hegarty-Miller)

Given a function $p : \mathbb{N} \to (0,1)$ such that $p(N) = o(1)$ and $N^{-1} = o(p(N))$. As $N \to \infty$, the probability A as a subset of $[1, \dots, N]$ is *MSTD* tends to 0. Let $\mathcal{S} = |A + A|, \mathcal{D} = |A - A|$ and $\mathcal{S}^C, \mathcal{D}^C$ denote their complements.

(i) $p = o(N^{-1/2})$: Then $\mathcal{D} \sim 2\mathcal{S} \sim (N.p)^2$

Theorem (Hegarty-Miller)

Given a function $p : \mathbb{N} \to (0,1)$ such that $p(N) = o(1)$ and $N^{-1} = o(p(N))$. As $N \to \infty$, the probability A as a subset of $[1, \dots, N]$ is *MSTD* tends to 0. Let $\mathcal{S} = |A+A|, \mathcal{D} = |A-A|$ and $\mathcal{S}^C, \mathcal{D}^C$ denote their complements.

(i) $p = o(N^{-1/2})$: Then $\mathcal{D} \sim 2\mathcal{S} \sim (N.p)^2$

(ii) $p = c.N^{-1/2}$ for $c \in (0, \infty)$. Let $g(x) = 2(e^{-x} - (1-x))/x$:

$$\mathcal{S} \sim g\left(\frac{c^2}{2}\right) N \quad \text{and} \quad \mathcal{D} \sim g(c^2) N$$

Theorem (Hegarty-Miller)

Given a function $p : \mathbb{N} \to (0, 1)$ such that $p(N) = o(1)$ and $N^{-1} = o(p(N))$. As $N \to \infty$, the probability A as a subset of $[1, \ldots, N]$ is *MSTD* tends to 0. Let $\mathcal{S} = |A + A|, \mathcal{D} = |A - A|$ and $\mathcal{S}^C, \mathcal{D}^C$ denote their complements.

(i) $p = o(N^{-1/2})$: Then $\mathcal{D} \sim 2\mathcal{S} \sim (N.p)^2$

(ii) $p = c.N^{-1/2}$ for $c \in (0, \infty)$. Let $g(x) = 2(e^{-x} - (1 - x))/x$:

$$\mathcal{S} \sim g\left(\frac{c^2}{2}\right) N \quad \text{and} \quad \mathcal{D} \sim g(c^2)N$$

(iii) $N^{-1/2} = o(p)$: $\mathcal{S}^C \sim 2.\mathcal{D}^C \sim \frac{4}{p^2}$

Our result

We prove a similar result:

Let $\hat{p} = p^2(2\rho_1 - \rho_1^2) + 2p(1-p)\rho_2$ be depend on N. Then

(i) $\hat{p} = o(N^{-1})$: Then $\mathcal{D} \sim 2\mathcal{S} \sim N^2.\hat{p}$

(ii) $\hat{p} = c.N^{-1}$ for $c \in (0, \infty)$. Let $g(x) = 2(e^{-x} - (1-x))/x$:

$$\mathcal{S} \sim g\left(\frac{c^2}{2}\right) N \quad \text{and} \quad \mathcal{D} \sim g(c^2)N$$

(iii) $N^{-1} = o(\hat{p})$: $\mathbb{E}(\mathcal{S}^C) = \mathbb{E}(2.\mathcal{D}^C) = 4/\hat{p}$

Notes in Our result

In our result, if we let $\rho_1 = 1, \rho_2 = 0$ then $\hat{p} = p^2$, consistent with the result in Hegarty-Miller.

Notes in Our result

In our result, if we let $\rho_1 = 1, \rho_2 = 0$ then $\hat{p} = p^2$, consistent with the result in Hegarty-Miller.

If $\rho_1 = \rho_2 = p$ then the critical phase happens when $p^2 = \Theta(1/N)$ or $p = \Theta(N^{-1/2})$.

Notes in Our result

In our result, if we let $\rho_1 = 1, \rho_2 = 0$ then $\hat{p} = p^2$, consistent with the result in Hegarty-Miller.

If $\rho_1 = \rho_2 = p$ then the critical phase happens when $p^2 = \Theta(1/N)$ or $p = \Theta(N^{-1/2})$.

The interesting case is A and A^C: $\hat{p} = 2p(1-p) = \Theta(1/N)$. If we let $p = o(1)$, $p = \Theta(1/N)$ which implies the expected number of elements of A is $pN =$ constant.

The minimal MSTD pair

Hegarty (2007) proved the smallest MSTD set has size 8.

The minimal MSTD pair

Hegarty (2007) proved the smallest MSTD set has size 8. We prove

Theorem

The smallest MSTD pair has size $(3, 5)$ or $(4, 4)$.

The minimal MSTD pair

Hegarty (2007) proved the smallest MSTD set has size 8. We prove

Theorem

The smallest MSTD pair has size $(3, 5)$ or $(4, 4)$.

Examples of minimal size MSTD pair:

$$A = \{1, 2, 5, 7\}, \quad B = \{1, 3, 6, 7\}$$

$$A = \{3, 4, 6\}, \quad B = \{1, 2, 5, 7, 8\}$$

$$A = \{3, 5, 6\}, \quad B = \{1, 2, 4, 7, 8\}.$$

Proof of Minimal MSTD pair

It is enough to prove that there is no MSTD-pair of size $(1, k), (2, k), (3, 3)$ or $(3, 4)$ for any positive integer k.

Proof of Minimal MSTD pair

It is enough to prove that there is no MSTD-pair of size $(1, k), (2, k), (3, 3)$ or $(3, 4)$ for any positive integer k.

Lemma

If A, B is a MSTD pair then there exist $a_1 < a_2 < a_3 \in A$ and $b_1 > b_2 > b_3 \in B$ such that $a_1 + b_1 = a_2 + b_2 = a_3 + b_3$.

Proof of Minimal MSTD pair

It is enough to prove that there is no MSTD-pair of size $(1, k), (2, k), (3, 3)$ or $(3, 4)$ for any positive integer k.

Lemma

If A, B is a MSTD pair then there exist $a_1 < a_2 < a_3 \in A$ and $b_1 > b_2 > b_3 \in B$ such that $a_1 + b_1 = a_2 + b_2 = a_3 + b_3$.

Idea of the proof: Consider all sums and differences $a \pm b$ where $a \in A, b \in B$. Each collapsed sum implies one collapsed difference.

Proof of Minimal MSTD pair

It is enough to prove that there is no MSTD-pair of size $(1, k), (2, k), (3, 3)$ or $(3, 4)$ for any positive integer k.

Lemma

If A, B is a MSTD pair then there exist $a_1 < a_2 < a_3 \in A$ and $b_1 > b_2 > b_3 \in B$ such that $a_1 + b_1 = a_2 + b_2 = a_3 + b_3$.

Idea of the proof: Consider all sums and differences $a \pm b$ where $a \in A, b \in B$. Each collapsed sum implies one collapsed difference.

Corollary: There is no MSTD pair of size $(1, k)$ and $(2, k)$ for $k > 0$.

Proof of Minimal MSTD pair

It is enough to prove that there is no MSTD-pair of size $(1, k), (2, k), (3, 3)$ or $(3, 4)$ for any positive integer k.

Lemma

If A, B is a MSTD pair then there exist $a_1 < a_2 < a_3 \in A$ and $b_1 > b_2 > b_3 \in B$ such that $a_1 + b_1 = a_2 + b_2 = a_3 + b_3$.

Idea of the proof: Consider all sums and differences $a \pm b$ where $a \in A, b \in B$. Each collapsed sum implies one collapsed difference.

Corollary: There is no MSTD pair of size $(1, k)$ and $(2, k)$ for $k > 0$.

We use some tedious checking to eliminate the case $(3, 3)$ and $(3, 4)$.

Summary of Results

- We prove for each $\vec{\rho} = (p, \rho_1, \rho_2)$ the limiting probability $P(\vec{\rho})$ of picking an MSTD $\vec{\rho}$-correlated pair exists and positive (except in some extreme cases).

Summary of Results

- We prove for each $\vec{\rho} = (p, \rho_1, \rho_2)$ the limiting probability $P(\vec{\rho})$ of picking an MSTD $\vec{\rho}$-correlated pair exists and positive (except in some extreme cases).

- The function $P(\vec{\rho})$ is continuous and differentiable.

Summary of Results

- We prove for each $\vec{\rho} = (p, \rho_1, \rho_2)$ the limiting probability $P(\vec{\rho})$ of picking an MSTD $\vec{\rho}$-correlated pair exists and positive (except in some extreme cases).
- The function $P(\vec{\rho})$ is continuous and differentiable.
- We show that $P(\vec{\rho})$ approaches zero and characterize the phase transition when we let $\vec{\rho}$ decay with n.

Summary of Results

- We prove for each $\vec{\rho} = (p, \rho_1, \rho_2)$ the limiting probability $P(\vec{\rho})$ of picking an MSTD $\vec{\rho}$-correlated pair exists and positive (except in some extreme cases).
- The function $P(\vec{\rho})$ is continuous and differentiable.
- We show that $P(\vec{\rho})$ approaches zero and characterize the phase transition when we let $\vec{\rho}$ decay with n.
- We find the minimal size of an MSTD pair (A, B).

Future Research

- Prove Conjecture 4: $\sup P(\vec{\rho}) = P(1/2, 0, 1)$.

Future Research

- Prove Conjecture 4: $\sup P(\vec{\rho}) = P(1/2, 0, 1)$.
- Find an efficient way to calculate values of P and investigate more analytic properties of P.

Future Research

- Prove Conjecture 4: $\sup P(\vec{\rho}) = P(1/2, 0, 1)$.
- Find an efficient way to calculate values of P and investigate more analytic properties of P.
- Prove the strong concentration of $\mathcal{S}^C$ and $\mathcal{D}^C$ in the case of slow decay (i.e. when $N^{-1/2} = o(\hat{p})$).

Future Research

- Prove Conjecture 4: $\sup P(\vec{\rho}) = P(1/2, 0, 1)$.
- Find an efficient way to calculate values of P and investigate more analytic properties of P.
- Prove the strong concentration of $\mathcal{S}^C$ and $\mathcal{D}^C$ in the case of slow decay (i.e. when $N^{-1/2} = o(\hat{p})$).
- Prove the uniqueness of the MSTD pairs of size $(4, 4)$ and $(3, 5)$, up to translation/dilation.

Acknowledgements

We would like to thank our advisor, Steven J. Miller, our co-authors David Moon and Archit Kulkarni, the rest of the team at the Williams College SMALL REU 2013, and the National Science Foundation.

This research was funded by NSF grant DMS0850577.

References

- P. Hegarty, Some explicit constructions of sets with more sums than differences. Acta Arithmetica **130** (2007), no. 1, 61–77.
- P. Hegarty and S. Miller, When almost all sets are difference dominated, *Random Structures and Algorithms* 35 (2009), no. 1, 118-136.
- G. Martin and K. O'Bryant, Many sets have more sums than differences, Additive Combinatorics, CRM Proc. Lecture Notes, vol. 43, Amer. Math. Soc., Providence, RI, 2007, pp. 287-305.
- Y. Zhao, Sets Characterized by Missing Sums and Differences, *Journal of Number Theory*, 131 (2010), pp. 2107-2134.