Sets Characterized by Missing Sums and Differences in $\mathbb{Z}^d$

Archit Kulkarni (Presenter), Carnegie Mellon University
David Moon (Presenter), Williams College

Thao Do, Stony Brook University
Jake Wellens, California Institute of Technology
James Wilcox, Williams College

Advisor: Steven J. Miller
SMALL REU, Williams College

Young Mathematicians Conference, 2013
Ohio State University

Introduction

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

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

Question: If A is finite, how does $|A + A|$ compare to $|A - A|$?

Introduction

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

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

Question: If A is finite, how does $|A + A|$ compare to $|A - A|$?

Typical pair of elements in $a_1, a_2 \in A$ has two differences, but only one sum. Expect sets A with $|A + A| > |A - A|$ to be rare.

Previous Results

Previous Results

Set A is *MSTD (More-Sums-Than-Differences)* if $|A+A| > |A-A|$.

Previous Results

Set A is *MSTD (More-Sums-Than-Differences)* if $|A + A| > |A - A|$.

Let ρ_n be the proportion of sets $A \subset [0, n]$ $(= \{0, \ldots, n\})$ that are MSTD.

Previous Results

Set A is *MSTD (More-Sums-Than-Differences)* if $|A + A| > |A - A|$.

Let ρ_n be the proportion of sets $A \subset [0, n]$ $(= \{0, \ldots, n\})$ that are MSTD.

Theorem [Martin and O'Bryant, 2006]

For $n \geq 14$, $\rho_n > 2 \cdot 10^{-7}$.

Previous Results

Set A is *MSTD (More-Sums-Than-Differences)* if $|A+A| > |A-A|$.

Let ρ_n be the proportion of sets $A \subset [0, n]$ $(= \{0, \ldots, n\})$ that are MSTD.

Theorem [Martin and O'Bryant, 2006]

For $n \geq 14$, $\rho_n > 2 \cdot 10^{-7}$.

Theorem [Zhao, 2010]

As $n \to \infty$, ρ_n converges to a limit $\rho > 4.28 \cdot 10^{-4}$.

Previous Results

Set A is *MSTD (More-Sums-Than-Differences)* if $|A+A| > |A-A|$.

Let ρ_n be the proportion of sets $A \subset [0, n]$ $(= \{0, \ldots, n\})$ that are MSTD.

Theorem [Martin and O'Bryant, 2006]

For $n \geq 14$, $\rho_n > 2 \cdot 10^{-7}$.

Theorem [Zhao, 2010]

As $n \to \infty$, ρ_n converges to a limit $\rho > 4.28 \cdot 10^{-4}$.

Monte Carlo simulations suggest $\rho \approx 4.5 \cdot 10^{-4}$.

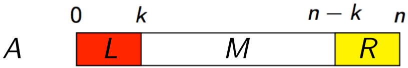
Fringe Elements

Fringe Elements

Key idea: *fringe elements are important*.

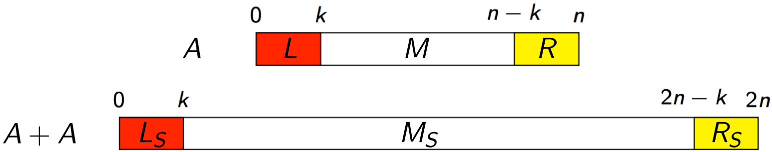
Fringe Elements

Key idea: *fringe elements are important*.



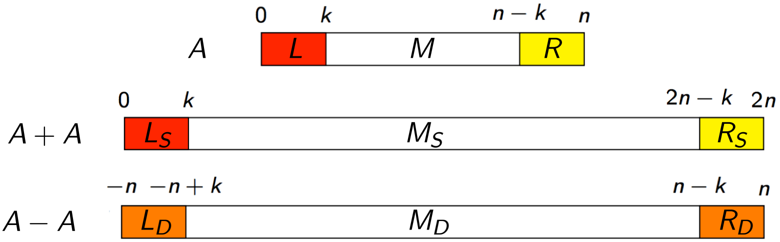
Fringe Elements

Key idea: *fringe elements are important*.



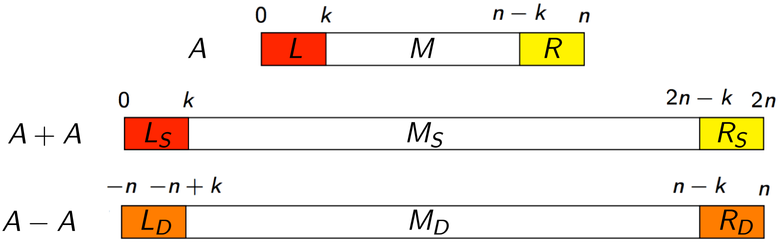
Fringe Elements

Key idea: *fringe elements are important*.



Fringe Elements

Key idea: *fringe elements are important*.



To show that ρ_n is positive, carefully fix elements in L and R so that

$$|(A+A)\cap(L_S\cup R_S)| > |(A-A)\cap(L_D\cap R_D)|.$$

With positive probability independent of n, $M_S \subset A+A$ (and therefore A is MSTD) .

Sets Characterized by Missing Sums and Differences

Sets Characterized by Missing Sums and Differences

Theorem [Hegarty, 2007]

For any $s, d \in \mathbb{N}_0$, there exists a constant $c \in (0, 1)$ that, for n sufficiently large depending on (s, d), at least $c \cdot 2^{n+1}$ subsets A of $[0, n]$ satisfy $|A + A| = (2n + 1) - s$ and $|A - A| = (2n + 1) - 2d$.

Sets Characterized by Missing Sums and Differences

Theorem [Hegarty, 2007]

For any $s, d \in \mathbb{N}_0$, there exists a constant $c \in (0, 1)$ that, for n sufficiently large depending on (s, d), at least $c \cdot 2^{n+1}$ subsets A of $[0, n]$ satisfy $|A + A| = (2n + 1) - s$ and $|A - A| = (2n + 1) - 2d$.

In particular, Hegarty explicitly constructs fringe sets L and R such that $(A + A) \cap (L_S \cup R_S)$ is missing s sums and $(A - A) \cap (L_D \cup R_D)$ is missing $2d$ differences.

Sets Characterized by Missing Sums and Differences

Theorem [Hegarty, 2007]

For any $s, d \in \mathbb{N}_0$, there exists a constant $c \in (0, 1)$ that, for n sufficiently large depending on (s, d), at least $c \cdot 2^{n+1}$ subsets A of $[0, n]$ satisfy $|A + A| = (2n + 1) - s$ and $|A - A| = (2n + 1) - 2d$.

In particular, Hegarty explicitly constructs fringe sets L and R such that $(A + A) \cap (L_S \cup R_S)$ is missing s sums and $(A - A) \cap (L_D \cup R_D)$ is missing $2d$ differences.

With positive probability, $M_S \subset A + A$ and $M_D \subset A - A$ (and therefore $A + A$ and $A - A$ are missing no other sums and differences).

Exploring Higher Dimensions

Exploring Higher Dimensions

This summer, we considered finite subsets A of $\mathbb{Z}^D$
(elements of A are vectors that add and subtract component-wise).

Exploring Higher Dimensions

This summer, we considered finite subsets A of $\mathbb{Z}^D$
(elements of A are vectors that add and subtract component-wise).

A simple extension is to consider sets $A \subset [0, n]^D$.

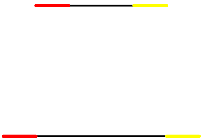
Exploring Higher Dimensions

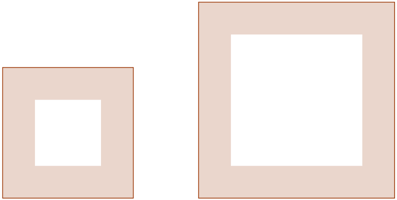
This summer, we considered finite subsets A of $\mathbb{Z}^D$
(elements of A are vectors that add and subtract component-wise).

A simple extension is to consider sets $A \subset [0, n]^D$.

Question

What is the fringe structure?





Corner Fringe

Corner Fringe

Theorem

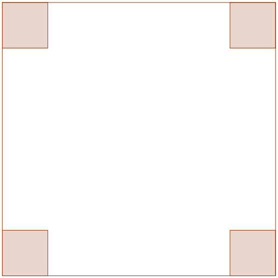
Let $D \in \mathbb{N}$. For any $(s, d) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$, there exists a constant $c_{s,d} \in (0, 1)$ such that, for n sufficiently large depending on (s, d), at least $c_{s,d} \cdot 2^{(n+1)^D}$ subsets of $[0, n]^d$ satisfy $|A + A| = (2n + 1)^D - s$ and $|A - A| = (2n + 1)^D - 2d$.

Corner Fringe

Theorem

Let $D \in \mathbb{N}$. For any $(s, d) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$, there exists a constant $c_{s,d} \in (0, 1)$ such that, for n sufficiently large depending on (s, d), at least $c_{s,d} \cdot 2^{(n+1)^D}$ subsets of $[0, n]^d$ satisfy $|A + A| = (2n + 1)^D - s$ and $|A - A| = (2n + 1)^D - 2d$.

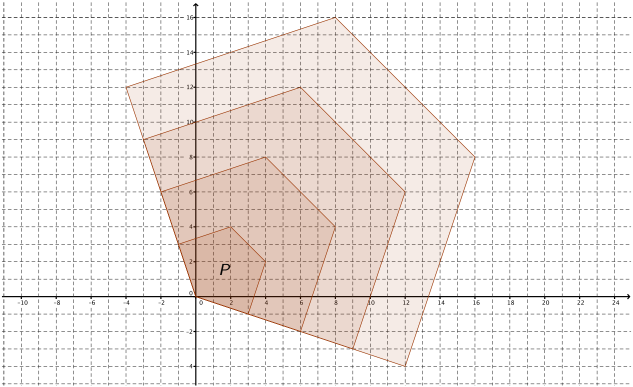
In particular, it suffices to fix a fringe about the corners:



Different Shapes

We can generalize further. Higher dimensions allow for different geometries of our set.

Let P be a convex polytope in $\mathbb{R}^d$ with vertices in $\mathbb{Z}^d$. We consider sets $A \subset L(nP) = nP \cap \mathbb{Z}^d$.



Main Result

Theorem

Let P be a convex polytope in $\mathbb{R}^D$ with vertices in $\mathbb{Z}^D$, and let $s, d \in \mathbb{N}_0$ be given. Then there exists a constant $c > 0$ such that for n large, a uniformly randomly chosen subset of $L(nP)$ has s missing sums and at least $2d$ missing differences with probability at least c.

Main Result

Theorem

Let P be a convex polytope in $\mathbb{R}^D$ with vertices in $\mathbb{Z}^D$, and let $s, d \in \mathbb{N}_0$ be given. Then there exists a constant $c > 0$ such that for n large, a uniformly randomly chosen subset of $L(nP)$ has s missing sums and at least $2d$ missing differences with probability at least c.

Remark

In particular, this theorem implies the existence of a positive percentage of "more missing differences than missing sums" sets, which are in some sense an analogue of MSTD sets.

Theorem

Let P be a convex polytope in $\mathbb{R}^D$ with vertices in $\mathbb{Z}^D$, and let $s, d \in \mathbb{N}_0$ be given. Then there exists a constant $c > 0$ such that for n large, a uniformly randomly chosen subset of $L(nP)$ has s missing sums and at least $2d$ missing differences with probability at least c.

Theorem

Let P be a convex polytope in $\mathbb{R}^D$ with vertices in $\mathbb{Z}^D$, and let $s, d \in \mathbb{N}_0$ be given. Then there exists a constant $c > 0$ such that for n large, a uniformly randomly chosen subset of $L(nP)$ has s missing sums and at least $2d$ missing differences with probability at least c.

The proof idea is similar to the ideas found in previous work.

Theorem

Let P be a convex polytope in $\mathbb{R}^D$ with vertices in $\mathbb{Z}^D$, and let $s, d \in \mathbb{N}_0$ be given. Then there exists a constant $c > 0$ such that for n large, a uniformly randomly chosen subset of $L(nP)$ has s missing sums and at least $2d$ missing differences with probability at least c.

The proof idea is similar to the ideas found in previous work.

- Construct a constant fixed fringe for each nP, which guarantees that exactly s sums and $2d$ differences are missing from the fringes of the sumset and difference set.

Theorem

Let P be a convex polytope in $\mathbb{R}^D$ with vertices in $\mathbb{Z}^D$, and let $s, d \in \mathbb{N}_0$ be given. Then there exists a constant $c > 0$ such that for n large, a uniformly randomly chosen subset of $L(nP)$ has s missing sums and at least $2d$ missing differences with probability at least c.

The proof idea is similar to the ideas found in previous work.

- Construct a constant fixed fringe for each nP, which guarantees that exactly s sums and $2d$ differences are missing from the fringes of the sumset and difference set.
- Show that there is a positive probability that every sum is present in the region outside the controlled fringes of the sumset.

Theorem

Let P be a convex polytope in $\mathbb{R}^D$ with vertices in $\mathbb{Z}^D$, and let $s, d \in \mathbb{N}_0$ be given. Then there exists a constant $c > 0$ such that for n large, a uniformly randomly chosen subset of $L(nP)$ has s missing sums and at least $2d$ missing differences with probability at least c.

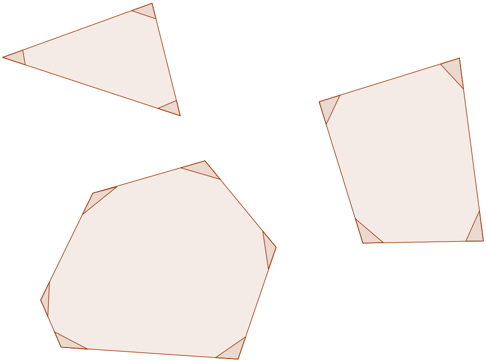
The proof idea is similar to the ideas found in previous work.

- Construct a constant fixed fringe for each nP, which guarantees that exactly s sums and $2d$ differences are missing from the fringes of the sumset and difference set.
- Show that there is a positive probability that every sum is present in the region outside the controlled fringes of the sumset.
- Conclude that we have exactly s sums missing and at least $2d$ differences missing for a positive percentage of sets.

Constructing the Fringe

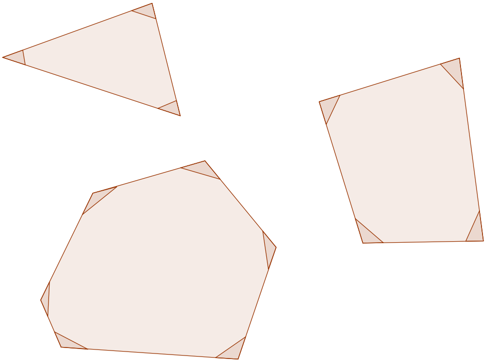
Constructing the Fringe

Like in the cube, we fix a fringe of constant size about the corners.



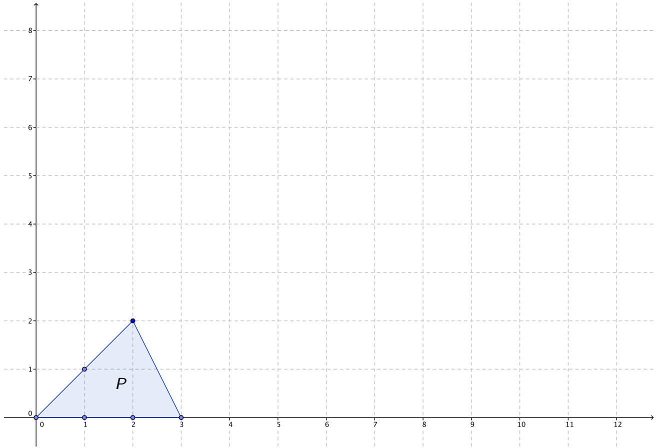
Constructing the Fringe

Like in the cube, we fix a fringe of constant size about the corners.

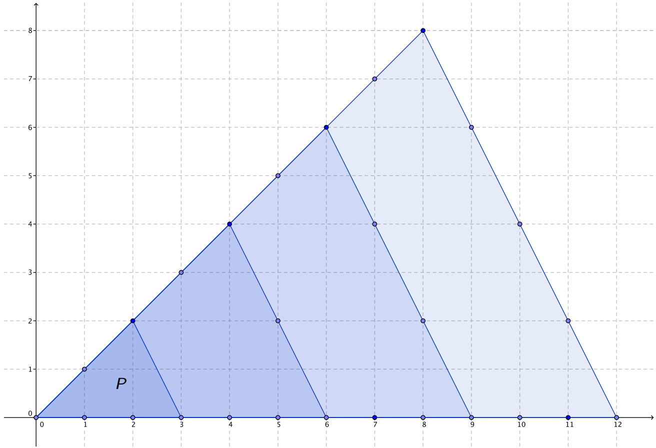


Missing sums and missing differences are controlled on the edges within the corner fringe.

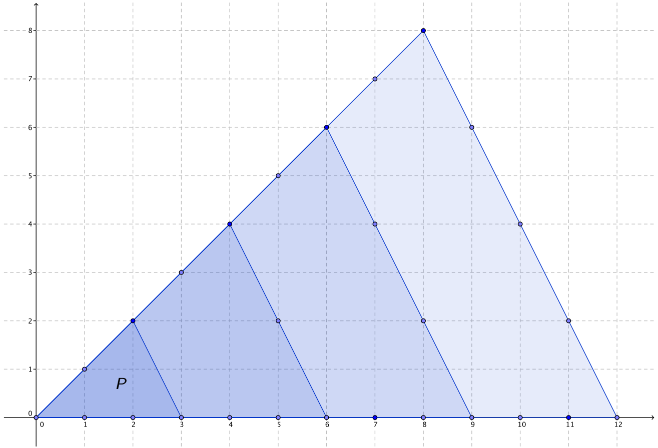
Lattice Points on the Edges



Lattice Points on the Edges



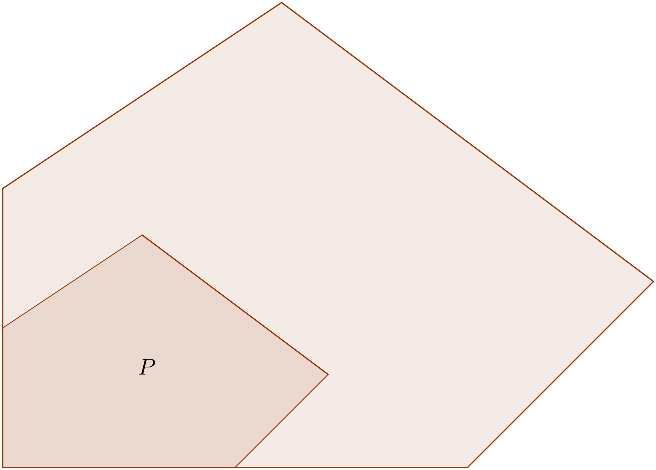
Lattice Points on the Edges



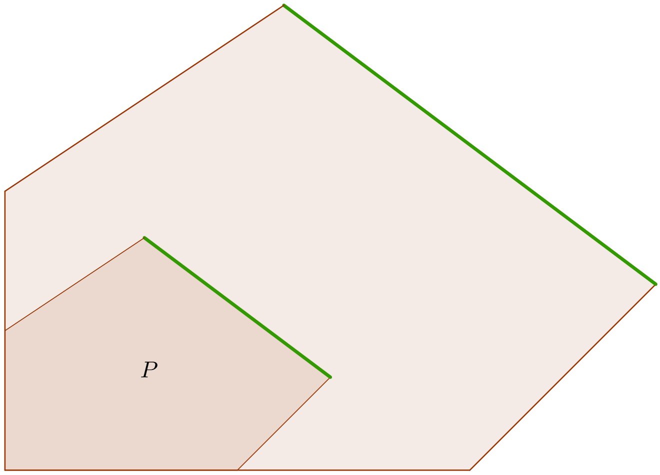
Since vertices are integer lattice points, note that:

- nP has at least $n+1$ lattice points on each edge
- Each edge forms its own 1D lattice structure

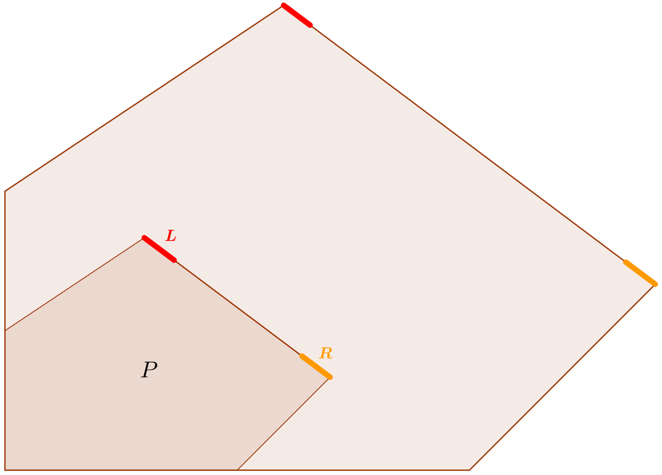
Controlling Missing Sums



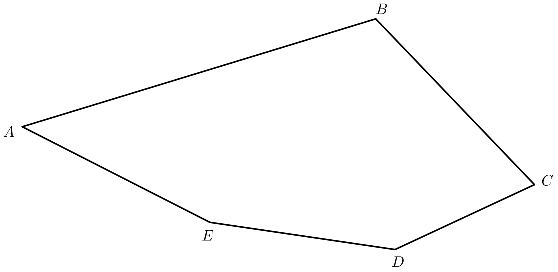
Controlling Missing Sums



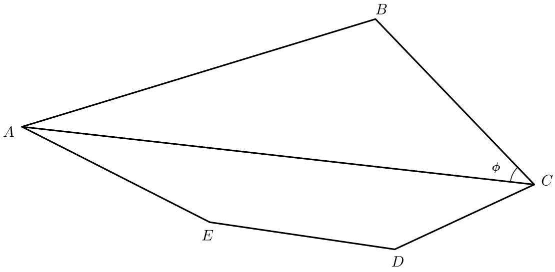
Controlling Missing Sums



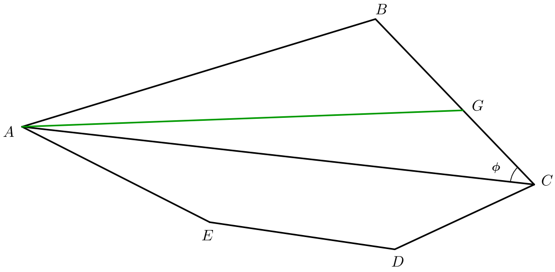
Controlling Missing Differences



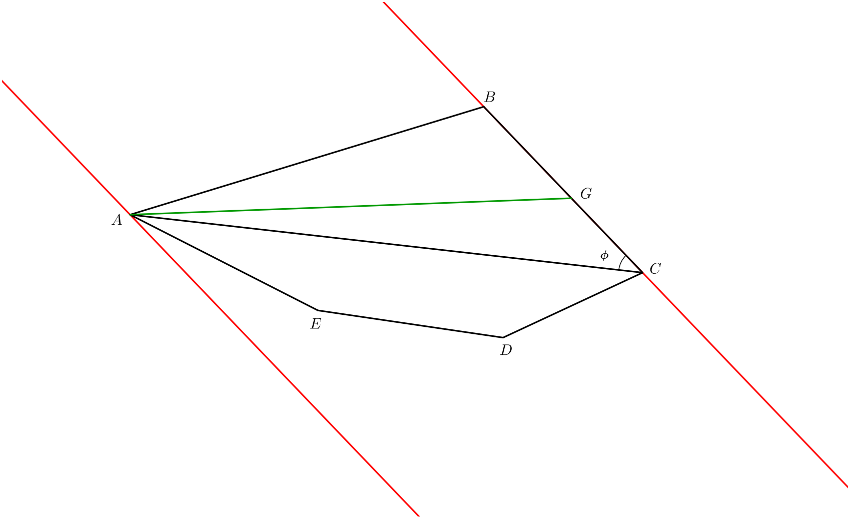
Controlling Missing Differences



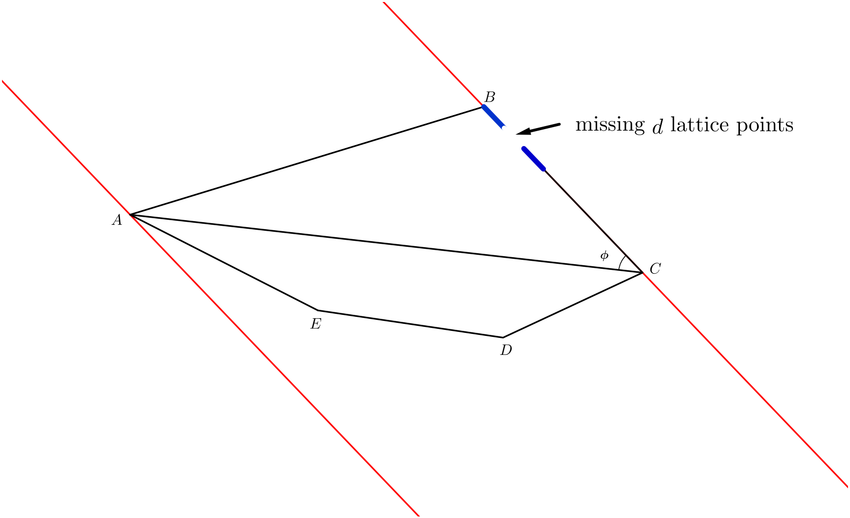
Controlling Missing Differences



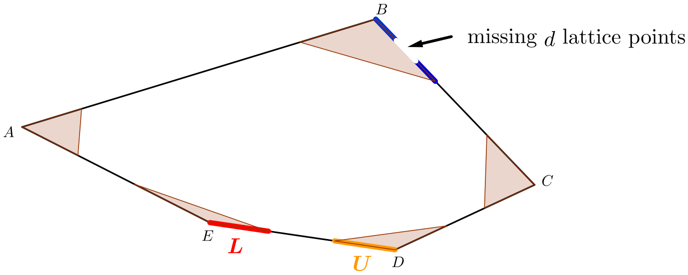
Controlling Missing Differences



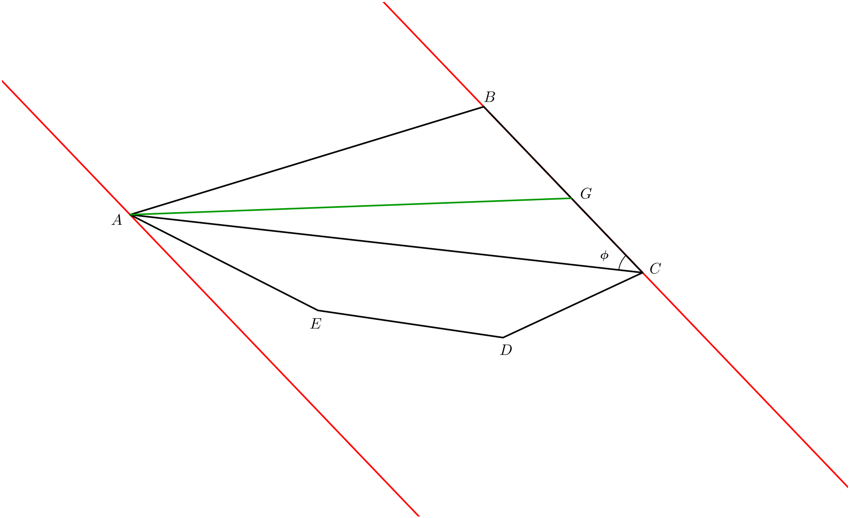
Controlling Missing Differences



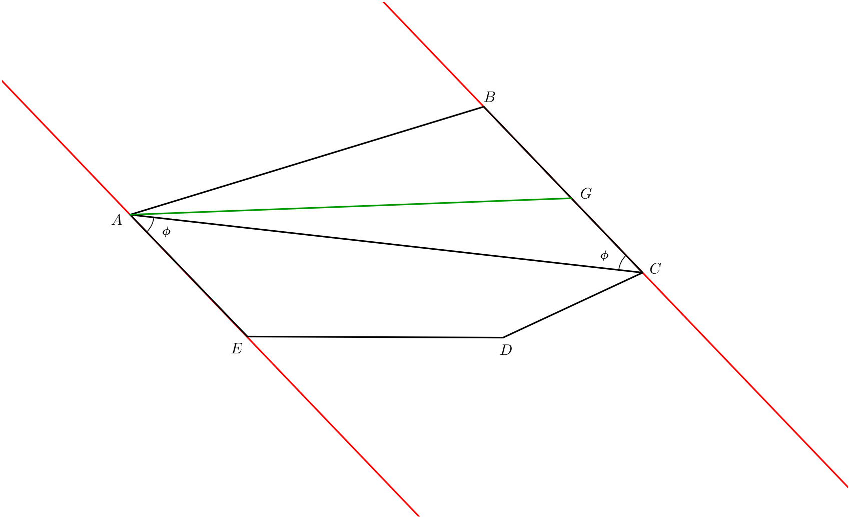
Putting It All Together



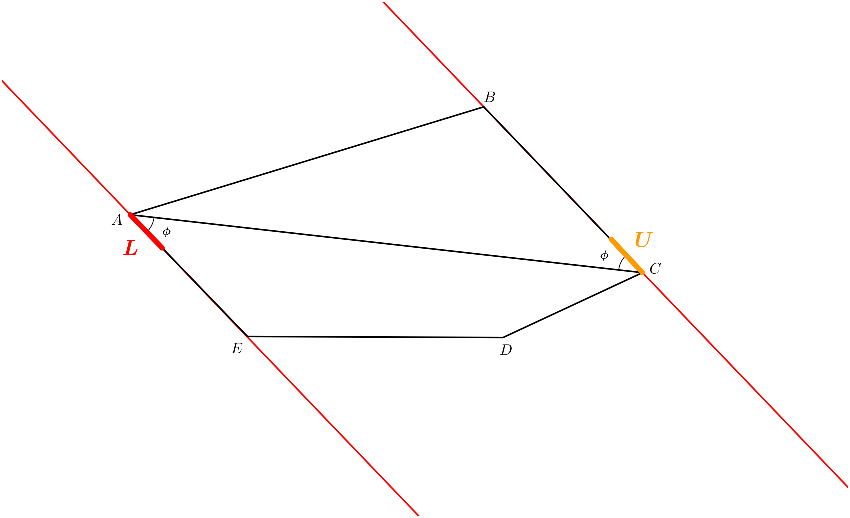
Special Case



Special Case



Special Case



Probability that a Sum is Missing

Lemma

Let A be a uniformly randomly chosen subset of $L(P)$ with our fixed fringe. Let $\mathbf{k}$ be given in $L((P+P) \setminus F_s)$. Then

$$\mathbb{P}[\mathbf{k} \notin P+P] \leq \left(\frac{3}{4}\right)^{|L(P \cap (\mathbf{k}-P))|/2}.$$

Probability that a Sum is Missing

Lemma

Let A be a uniformly randomly chosen subset of $L(P)$ with our fixed fringe. Let $\mathbf{k}$ be given in $L((P+P)\setminus F_s)$. Then

$$\mathbb{P}[\mathbf{k} \notin P+P] \leq \left(\frac{3}{4}\right)^{|L(P\cap(\mathbf{k}-P))|/2}.$$

Proof.

If $\mathbf{x}, \mathbf{y} \in L(P)$ with $\mathbf{x}+\mathbf{y}=\mathbf{k}$, then $\mathbf{x}, \mathbf{y} \in L(\mathbf{k}-P)$ too.

Probability that a Sum is Missing

Lemma

Let A be a uniformly randomly chosen subset of $L(P)$ with our fixed fringe. Let $\mathbf{k}$ be given in $L((P+P)\setminus F_s)$. Then

$$\mathbb{P}[\mathbf{k} \notin P+P] \leq \left(\frac{3}{4}\right)^{|L(P\cap(\mathbf{k}-P))|/2}.$$

Proof.

If $\mathbf{x}, \mathbf{y} \in L(P)$ with $\mathbf{x}+\mathbf{y}=\mathbf{k}$, then $\mathbf{x}, \mathbf{y} \in L(\mathbf{k}-P)$ too.
Partition $L(P \cap \mathbf{k}-P)$ into independent pairs of points adding to $\mathbf{k}$.

Probability that a Sum is Missing

Lemma

Let A be a uniformly randomly chosen subset of $L(P)$ with our fixed fringe. Let $\mathbf{k}$ be given in $L((P+P) \setminus F_s)$. Then

$$\mathbb{P}[\mathbf{k} \notin P + P] \leq \left(\frac{3}{4}\right)^{|L(P \cap (\mathbf{k} - P))|/2}.$$

Proof.

If $\mathbf{x}, \mathbf{y} \in L(P)$ with $\mathbf{x} + \mathbf{y} = \mathbf{k}$, then $\mathbf{x}, \mathbf{y} \in L(\mathbf{k} - P)$ too.
Partition $L(P \cap \mathbf{k} - P)$ into independent pairs of points adding to $\mathbf{k}$.
If the sum $\mathbf{k}$ is missing, each pair is missing at least one point.

Probability that a Sum is Missing

Lemma

Let A be a uniformly randomly chosen subset of $L(P)$ with our fixed fringe. Let $\mathbf{k}$ be given in $L((P+P)\setminus F_s)$. Then

$$\mathbb{P}[\mathbf{k} \notin P+P] \leq \left(\frac{3}{4}\right)^{|L(P\cap(\mathbf{k}-P))|/2}.$$

Proof.

If $\mathbf{x},\mathbf{y} \in L(P)$ with $\mathbf{x}+\mathbf{y}=\mathbf{k}$, then $\mathbf{x},\mathbf{y} \in L(\mathbf{k}-P)$ too.
Partition $L(P \cap \mathbf{k} - P)$ into independent pairs of points adding to $\mathbf{k}$.
If the sum $\mathbf{k}$ is missing, each pair is missing at least one point.
Probability for each pair is at most $3/4$. Exponentiate by # of pairs.

Probability that a Sum is Missing

Lemma

Let A be a uniformly randomly chosen subset of $L(P)$ with our fixed fringe. Let $\mathbf{k}$ be given in $L((P+P) \setminus F_s)$. Then

$$\mathbb{P}[\mathbf{k} \notin P+P] \leq \left(\frac{3}{4}\right)^{|L(P \cap (\mathbf{k}-P))|/2}.$$

Proof.

If $\mathbf{x}, \mathbf{y} \in L(P)$ with $\mathbf{x}+\mathbf{y}=\mathbf{k}$, then $\mathbf{x}, \mathbf{y} \in L(\mathbf{k}-P)$ too.
Partition $L(P \cap \mathbf{k}-P)$ into independent pairs of points adding to $\mathbf{k}$.
If the sum $\mathbf{k}$ is missing, each pair is missing at least one point.
Probability for each pair is at most $3/4$. Exponentiate by $\#$ of pairs.

□

Probability that All Middle Sums are Present

We wish to find a constant $c > 0$ so that

$$\sum_{\mathbf{k} \in (L(P)+L(P)) \setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P] < 1 - c.$$

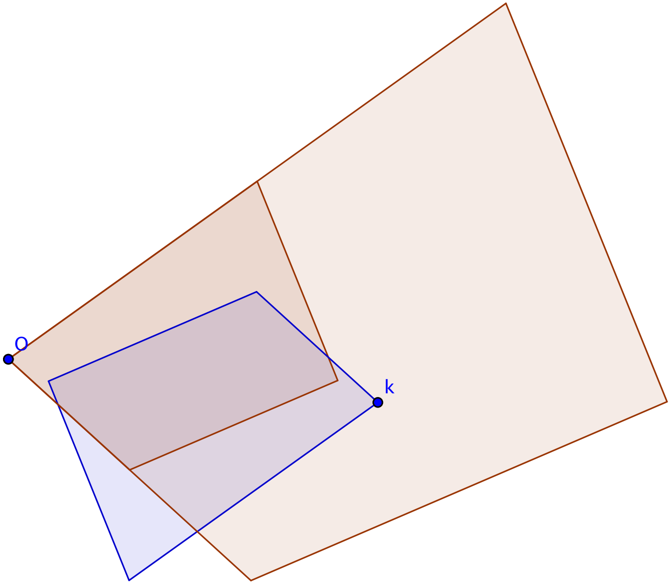
This will imply that all middle sums are present with probability at least c.

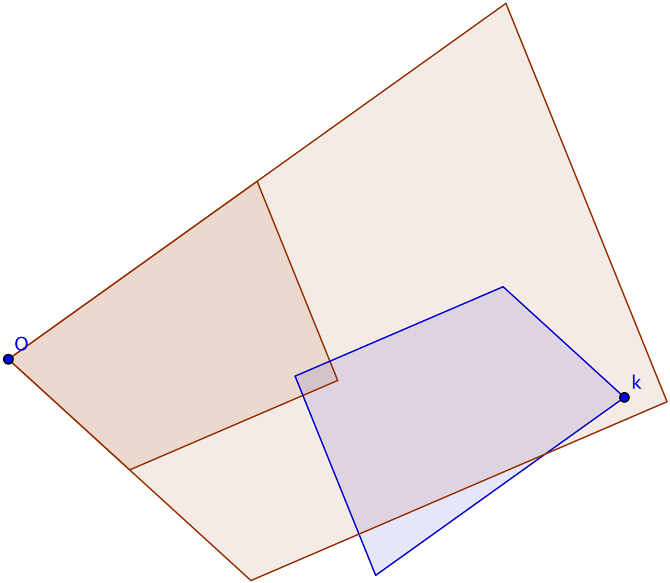
Probability that All Middle Sums are Present

We wish to find a constant $c > 0$ so that

$$\sum_{\mathbf{k} \in (L(P)+L(P)) \setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P] < 1 - c.$$

This will imply that all middle sums are present with probability at least c. First, we must get a handle on the area of $P \cap (\mathbf{k} - P)$.


O
k


O
k

To handle the sum

$$\sum_{\mathbf{k}\in(L(P)+L(P))\setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P],$$

we partition the region $L(P) + L(P)$ into three regions:

To handle the sum

$$\sum_{\mathbf{k} \in (L(P)+L(P)) \setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P],$$

we partition the region $L(P) + L(P)$ into three regions:

- Our fixed fringe.

To handle the sum

$$\sum_{\mathbf{k} \in (L(P)+L(P)) \setminus F_s} \mathbb{P}[\mathbf{k} \notin P+P],$$

we partition the region $L(P)+L(P)$ into three regions:

- Our fixed fringe.
- A middle region where $P \cap (\mathbf{k}-P)$ is so large that the probability of any given sum to be missing is less than $\frac{\epsilon}{|L(P+P)|}$.

To handle the sum

$$\sum_{\mathbf{k}\in(L(P)+L(P))\setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P],$$

we partition the region $L(P) + L(P)$ into three regions:

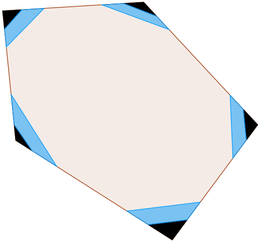
- Our fixed fringe.
- A middle region where $P \cap (\mathbf{k} - P)$ is so large that the probability of any given sum to be missing is less than $\frac{\epsilon}{|L(P+P)|}$.
- An intermediate region, close enough to the vertices that $P \cap (\mathbf{k} - P)$ is simple enough that its area can be estimated.

To handle the sum

$$\sum_{\mathbf{k} \in (L(P)+L(P)) \setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P],$$

we partition the region $L(P) + L(P)$ into three regions:

- Our fixed fringe.
- A middle region where $P \cap (\mathbf{k} - P)$ is so large that the probability of any given sum to be missing is less than $\frac{\epsilon}{|L(P+P)|}$.
- An intermediate region, close enough to the vertices that $P \cap (\mathbf{k} - P)$ is simple enough that its area can be estimated.



Thus, we can fix a large enough fringe so that

$$\sum_{\mathbf{k} \in (L(P)+L(P)) \setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P] < 1 - c,$$

so for n large, all middle sums are present with probability at least c.

Thus, we can fix a large enough fringe so that

$$\sum_{\mathbf{k}\in(L(P)+L(P))\setminus F_s} \mathbb{P}[\mathbf{k} \notin P + P] < 1 - c,$$

so for n large, all middle sums are present with probability at least c.

Since our fixed fringe guarantees s missing fringe sums and at least $2d$ missing differences, we have found a positive proportion of subsets with exactly s missing sums and at least $2d$ missing differences overall.

Thanks

We would like to thank

- Our advisor, Professor Steven Miller.
- Our co-authors Thao Do, Jake Wellens, and James Wilcox.
- Williams College and the NSF (Grant DMS0850577) for funding our research.