On the typical structure of sets with small sumset

M. Campos  M. Coulson  O. Serra  M. Wötzel

1Universitat Politècnica de Catalunya & Barcelona Graduate School of Mathematics

TU Ilmenau Online Mittagsseminar

March 15, 2021

*Thanks to support from MDM-2014-0445.
Introduction

- Object of study: the sumset

\[ A + B = \{ a + b : a \in A, b \in B \}, \]

where \( A, B \subset G \) are finite subsets of some group \( G \).

- Basic question: If we know that \( |A + B| \) is small, can we say anything about the structure of \( A \) and \( B \)?
Classical Inverse Results and an Observation

**Theorem (Folklore)**

Suppose $A, B \subseteq \mathbb{Z}$ are finite. Then $|A + B| = |A| + |B| - 1$, if and only if $A$ and $B$ are arithmetic progressions with the same common difference.

**Theorem (Freiman)**

If $A \subseteq \mathbb{Z}$ is finite such that $|2A| \leq K|A|$, then $A$ is contained in a generalized arithmetic progression of dimension $d(K)$ and size $f(K)|A|$. 
Classical Inverse Results and an Observation

**Theorem (Folklore)**
Suppose $A, B \subset \mathbb{Z}$ are finite. Then $|A + B| = |A| + |B| - 1$, if and only if $A$ and $B$ are arithmetic progressions with the same common difference.

**Theorem (Freiman)**
If $A \subset \mathbb{Z}$ is finite such that $|2A| \leq K|A|$, then $A$ is contained in a generalized arithmetic progression of dimension $d(K)$ and size $f(K)|A|$.

- Suppose that $P \subset \mathbb{Z}$ is an arithmetic progression of size $Ks/2$.
- If $A \subset P$ is an arbitrary subset of size $s$, then we clearly have $2A \subset 2P$ and hence $|2A| \leq |2P| \approx Ks$.
- Question: Can we get an inverse result in this direction by going to the random setting? How many sets with small doubling are exceptions to this construction?
The Typical Case - Rough Structure

- **Question:** Can we get an inverse result in this direction by going to the random setting?
- **Answer:** Yes! Campos proved the following rough structural result.

**Theorem (Campos)**

Let $s = \omega((\log n)^3)$ and $K = o(s/(\log n)^3)$.
Question: Can we get an inverse result in this direction by going to the random setting?

Answer: Yes! Campos proved the following rough structural result.

**Theorem (Campos)**

Let $s = \omega((\log n)^3)$ and $K = o(s/(\log n)^3)$. Then for almost all sets $A \subset [n]$ with $|A| = s$ and $|2A| \leq Ks$, there exists an arithmetic progression $P$ of size

$$|P| \leq \frac{1 + o(1)}{2} Ks$$

such that at most $o(s)$ points of $A$ are not contained in $P$.

Note: Campos also proved a counting result about the number of $s$-sets with doubling constant $K$. 
The Typical Case - Precise Structure

This was then used to get the following more precise structural result in the case $K = O(1)$.

**Theorem (Campos, Collares, Morris, Morrison, Souza)**

Fix $K \geq 3$ and $\epsilon > 0$. For $n$ sufficiently large, let $s \geq (\log n)^4$. 

Note that the value of $c$ is close to optimal, as there is also a lower bound of the form $\Omega(K^2 \log \frac{1}{\epsilon})$. 

This was then used to get the following more precise structural result in the case $K = O(1)$.

**Theorem (Campos, Collares, Morris, Morrison, Souza)**

Fix $K \geq 3$ and $\epsilon > 0$. For $n$ sufficiently large, let $s \geq (\log n)^4$. Then for all but an $\epsilon$ proportion of sets $A \subset [n]$ with $|A| = s$ and $|2A| \leq Ks$, it holds that $A$ is contained in an arithmetic progression $P$ of size

$$|P| \leq \frac{Ks}{2} + c(K, \epsilon),$$

where $c(K, \epsilon) = O(K^2 \log K \log(1/\epsilon))$. 

Note that the value of $c$ is close to optimal, as there is also a lower bound of the form $\Omega(K^2 \log K \log(1/\epsilon))$. 

M. Wötzel (UPC & BGSMath)
The Typical Case - Precise Structure

This was then used to get the following more precise structural result in the case \( K = O(1) \).

**Theorem (Campos, Collares, Morris, Morrison, Souza)**

Fix \( K \geq 3 \) and \( \epsilon > 0 \). For \( n \) sufficiently large, let \( s \geq (\log n)^4 \). Then for all but an \( \epsilon \) proportion of sets \( A \subset [n] \) with \( |A| = s \) and \( |2A| \leq Ks \), it holds that \( A \) is contained in an arithmetic progression \( P \) of size

\[
|P| \leq \frac{Ks}{2} + c(K, \epsilon),
\]

where \( c(K, \epsilon) = O(K^2 \log K \log(1/\epsilon)) \).

Note that the value of \( c \) is close to optimal, as there is also a lower bound of the form \( \Omega(K^2 \log(1/\epsilon)) \).
Our Result

We adapt Campos’ original result to distinct sets of equal cardinality.

**Theorem (Campos, Coulson, Serra, W.)**

\[ s = \omega((\log n)^3) \text{ and } K = o(s/(\log n)^3). \]
Our Result

We adapt Campos’ original result to distinct sets of equal cardinality.

**Theorem (Campos, Coulson, Serra, W.)**

Let \( s = \omega((\log n)^3) \) and \( K = o(s/(\log n)^3) \). Then for almost all sets \( A, B \subset [n] \) with \( |A| = |B| = s \) and \( |A + B| \leq Ks \) there exist arithmetic progressions \( P, Q \) with the same common difference of size

\[
|P|, |Q| \leq \frac{1 + o(1)}{2} Ks
\]

such that \( |A \setminus P|, |B \setminus Q| = o(s) \).

More precise bounds for the hidden terms are obtained, similar but slightly weaker compared to those of Campos.

Main tool in the proof: recent version of the method of hypergraph containers.
The Method of Hypergraph Containers

- Introduced explicitly in 2015 by Balogh, Morris and Samotij, and independently by Saxton and Thomason, used to count the number of combinatorial objects avoiding some specific substructure, especially in the \textit{sparse random} setting.

General idea: If \( H \) is a hypergraph satisfying some specific degree conditions, then there exists a relatively small family \( C \) of containers such that:

1. For each independent set \( I \) in \( V \), there exists a container \( C \) with \( I \subseteq C \).
2. Each \( C \) is smaller than \( V \) by some constant factor.

As long as the degree conditions are met, one can now reapply this result to the induced hypergraph on each container.
The Method of Hypergraph Containers

- Introduced explicitly in 2015 by Balogh, Morris and Samotij, and independently by Saxton and Thomason, used to count the number of combinatorial objects avoiding some specific substructure, especially in the sparse random setting.

- General idea: If $\mathcal{H}$ is a hypergraph satisfying some specific degree conditions, then there exists a relatively small family $\mathcal{C} \subset 2^V(\mathcal{H})$ of containers such that:

  1. For each independent set $I \subseteq V(\mathcal{H})$, there exists a container $C \in \mathcal{C}$ with $I \subseteq C$.
  2. Each $C$ is smaller than $V(\mathcal{H})$ by some constant factor.

As long as the degree conditions are met, one can now reapply this result to the induced hypergraph on each container $C$. 
The Method of Hypergraph Containers

- Introduced explicitly in 2015 by Balogh, Morris and Samotij, and independently by Saxton and Thomason, used to count the number of combinatorial objects avoiding some specific substructure, especially in the sparse random setting.

- General idea: If $\mathcal{H}$ is a hypergraph satisfying some specific degree conditions, then there exists a relatively small family $C \subseteq 2^{V(\mathcal{H})}$ of containers such that:
  1. For each independent set $I \subseteq V(\mathcal{H})$, there exists a container $C \in C$ with $I \subseteq C$.
  2. Each $C$ is smaller than $V(\mathcal{H})$ by some constant factor.
The Method of Hypergraph Containers

- Introduced explicitly in 2015 by Balogh, Morris and Samotij, and independently by Saxton and Thomason, used to count the number of combinatorial objects avoiding some specific substructure, especially in the *sparse random* setting.

- General idea: If $\mathcal{H}$ is a hypergraph satisfying some specific degree conditions, then there exists a relatively small family $C \subset 2^{V(\mathcal{H})}$ of containers such that:
  1. For each independent set $I \subset V(\mathcal{H})$, there exists a container $C \in C$ with $I \subset C$.
  2. Each $C$ is smaller than $V(\mathcal{H})$ by some constant factor.

- As long as the degree conditions are met, one can now reapply this result to the induced hypergraph on each container $C \in C$. 
After iterating, end up with a still small-ish collection of containers, each of which is very small, and it still holds that every independent set \( I \) of \( \mathcal{H} \) is contained in one of these containers.
After iterating, end up with a still small-ish collection of containers, each of which is very small, and it still holds that every independent set $I$ of $\mathcal{H}$ is contained in one of these containers.

Consider now a specific $\mathcal{H}$ whose hyperedges encode some structure (e.g. triangles in $K_n$).
After iterating, end up with a still small-ish collection of containers, each of which is very small, and it still holds that every independent set $I$ of $\mathcal{H}$ is contained in one of these containers.

Consider now a specific $\mathcal{H}$ whose hyperedges encode some structure (e.g. triangles in $K_n$).

The iteration stopped, hence the induced hypergraph on each container has few edges (i.e., few triangles).
The Method of Hypergraph Containers

- After iterating, end up with a still small-ish collection of containers, each of which is very small, and it still holds that every independent set $I$ of $\mathcal{H}$ is contained in one of these containers.

- Consider now a specific $\mathcal{H}$ whose hyperedges encode some structure (e.g. triangles in $K_n$).

- The iteration stopped, hence the induced hypergraph on each container has few edges (i.e., few triangles).

- Can now use classical stability and supersaturation results in the dense regime to get structural statements about the containers.
The Method of Hypergraph Containers

- After iterating, end up with a still small-ish collection of containers, each of which is very small, and it still holds that every independent set $I$ of $\mathcal{H}$ is contained in one of these containers.

- Consider now a specific $\mathcal{H}$ whose hyperedges encode some structure (e.g. triangles in $K_n$).

- The iteration stopped, hence the induced hypergraph on each container has few edges (i.e., few triangles).

- Can now use classical stability and supersaturation results in the dense regime to get structural statements about the containers.

- We can now exploit the structure of the containers to make more precise statements about the independent sets (i.e. triangle-free graphs) they contain.
An Asymmetric Version of the Container Lemma

- Original method works well when e.g. counting the number or determining structure of $H$-free graphs, for some fixed graph $H$. 

While it is possible to consider induced (or more generally multicolored) subgraphs with the original method, there is a quantitative loss coming with it. 

Problem: The container method does not take into account the asymmetry between the two colours, so it cannot distinguish between an induced copy of $H$ and a clique on $v^1$ vertices. 

In order to give a structure theorem about induced-$C_4$-free graphs, Morris, Samotij and Saxton recently developed an asymmetric (bipartite) version of the container method. Campos slightly modified this so that the two components can shrink at different rates and allowed non-uniform hypergraphs.
Original method works well when e.g. counting the number or determining structure of $H$-free graphs, for some fixed graph $H$.

While it is possible to consider induced (or more generally multicolored) subgraphs with the original method, there is a quantitative loss coming with it.
An Asymmetric Version of the Container Lemma

- Original method works well when e.g. counting the number or determining structure of $H$-free graphs, for some fixed graph $H$.
- While it is possible to consider induced (or more generally multicolored) subgraphs with the original method, there is a quantitative loss coming with it.
- Problem: The container method does not take into account the asymmetry between the two colours, so it cannot distinguish between an induced copy of $H$ and a clique on $v(H)$ vertices.

Campos slightly modified this so that the two components can shrink at different rates and allowed non-uniform hypergraphs.
An Asymmetric Version of the Container Lemma

- Original method works well when e.g. counting the number or determining structure of $H$-free graphs, for some fixed graph $H$.
- While it is possible to consider induced (or more generally multicolored) subgraphs with the original method, there is a quantitative loss coming with it.
- Problem: The container method does not take into account the asymmetry between the two colours, so it cannot distinguish between an induced copy of $H$ and a clique on $v(H)$ vertices.
- In order to give a structure theorem about induced-$C_4$-free graphs, Morris, Samotij and Saxton recently developed an asymmetric (bipartite) version of the container method.
An Asymmetric Version of the Container Lemma

- Original method works well when e.g. counting the number or determining structure of $H$-free graphs, for some fixed graph $H$.
- While it is possible to consider induced (or more generally multicolored) subgraphs with the original method, there is a quantitative loss coming with it.
- Problem: The container method does not take into account the asymmetry between the two colours, so it cannot distinguish between an induced copy of $H$ and a clique on $v(H)$ vertices.
- In order to give a structure theorem about induced-$C_4$-free graphs, Morris, Samotij and Saxton recently developed an asymmetric (bipartite) version of the container method.
- Campos slightly modified this so that the two components can shrink at different rates and allowed non-uniform hypergraphs.
We generalize the bipartite container lemma to an $r$-partite, $r$-bounded version and apply it to the following hypergraph.

For $V_1 - V_2 - V_3 \supseteq \binom{\mathbb{F}_2^n}{3}$, we consider the 3-partite and 3-uniform hypergraph $H(\mathbb{F}_2^n)$ with vertex set $V_1 \cup V_2 \cup V_3$ and hyperedges $\mathcal{H}(\mathbb{F}_2^n) = \{x - y - z \mid x \not\in y = z\}$. Note that the independent sets $I$ we are interested in will satisfy the additional condition that they have a large intersection with the third component, more precisely $|I \cap V_3| \gg K_s$.

Think: We want to count triples $A - B - C$ such that $|A| = |B| = s$, $A \not\subseteq B \subseteq C$ and $|C| = K_s$. 

M. Wötzel (UPC & BGSMath)
Proof Outline – Generalizing the Asymmetric Container Lemma

- We generalize the bipartite container lemma to an $r$-partite, $r$-bounded version and apply it to the following hypergraph.
- For $(V_1, V_2, V_3) \in 2[n] \times 2[n] \times 2[2n]$, we consider the 3-partite and 3-uniform hypergraph $\mathcal{H}(V_1, V_2, V_3)$ with vertex set $V_1 \cup V_2 \cup V_3$ and hyperedges

$$\{(x, y, z) \in V_1 \times V_2 \times V_3 : x + y = z\}.$$
We generalize the bipartite container lemma to an $r$-partite, $r$-bounded version and apply it to the following hypergraph.

For $(V_1, V_2, V_3) \in 2^n \times 2^n \times 2^{2n}$, we consider the 3-partite and 3-uniform hypergraph $\mathcal{H}(V_1, V_2, V_3)$ with vertex set $V_1 \cup V_2 \cup V_3$ and hyperedges

$$\{(x, y, z) \in V_1 \times V_2 \times V_3 : x + y = z\}.$$ 

Note that the independent sets $I$ we are interested in will satisfy the additional condition that they have a large intersection with the third component, more precisely $|I \cap V_3| \geq |V_3| - Ks$. 
Proof Outline – Generalizing the Asymmetric Container Lemma

- We generalize the bipartite container lemma to an $r$-partite, $r$-bounded version and apply it to the following hypergraph.

- For $(V_1, V_2, V_3) \in 2^n \times 2^n \times 2^{2n}$, we consider the 3-partite and 3-uniform hypergraph $\mathcal{H}(V_1, V_2, V_3)$ with vertex set $V_1 \cup V_2 \cup V_3$ and hyperedges

  \[ \{(x, y, z) \in V_1 \times V_2 \times V_3 : x + y = z\}. \]

- Note that the independent sets $I$ we are interested in will satisfy the additional condition that they have a large intersection with the third component, more precisely $|I \cap V_3| \geq |V_3| - Ks$.

- Think: We want to count triples $(A, B, C)$ such that $|A| = |B| = s$, $A + B \subset C$ and $|C| \leq Ks$. 
Proof Outline – The Container Theorem

We get the following theorem.

**Theorem (CCOW)**

Let \( n, K, s \) be integers such that \( \log n \leq s \leq Ks \leq s^2 \) and \( \epsilon > 0 \).
Proof Outline – The Container Theorem

We get the following theorem.

**Theorem (CCOW)**

Let $n, K, s$ be integers such that $\log n \leq s \leq Ks \leq s^2$ and $\epsilon > 0$. There is a family $\mathcal{A} \subset 2^{[n]} \times 2^{[n]} \times 2^{[2n]}$ of triples of sets $(X, Y, Z)$ of size

$$|\mathcal{A}| \leq \exp(O(\sqrt{Ks\epsilon^{-2}}(\log n)^{3/2})),$$

such that:

1. For all $A \subset B \subset C$ with $|A| = |B| = s$ and $A \supset B \supset C$ and $|C| \geq Ks$, there is a triple $(X, Y, Z) \in \mathcal{A}$ such that $X \supset A$, $Y \supset B$ and $Z \supset C$.
2. For every $(X, Y, Z) \in \mathcal{A}$, $|Z| \geq Ks$ and either $\max(f|X| - |Y|) \geq Ks$ or there are at most $\epsilon 2^{|X|}$ pairs $(x, y) \in X \times Y$ such that $x \neq y \notin Z$.

Note: The second property tells us that $H(X, Y, Z)$ has few edges.
We get the following theorem.

**Theorem (CCOW)**

Let \(n, K, s\) be integers such that \(\log n \leq s \leq Ks \leq s^2\) and \(\epsilon > 0\). There is a family \(\mathcal{A} \subset 2^n \times 2^n \times 2^{[2n]}\) of triples of sets \((X, Y, Z)\) of size

\[|\mathcal{A}| \leq \exp\left(O\left(\sqrt{Ks}\epsilon^{-2}(\log n)^{3/2}\right)\right),\]

such that:

1. For all \(A, B \subset [n], C \subset [2n]\) with \(|A| = |B| = s\), \(A + B \subset C\) and \(|C| \leq Ks\), there is a triple \((X, Y, Z) \in \mathcal{A}\) such that \(A \subset X\), \(B \subset Y\) and \(Z \subset C\).
We get the following theorem.

**Theorem (CCOW)**

Let $n, K, s$ be integers such that $\log n \leq s \leq Ks \leq s^2$ and $\epsilon > 0$. There is a family $\mathcal{A} \subset 2^{[n]} \times 2^{[n]} \times 2^{[2n]}$ of triples of sets $(X, Y, Z)$ of size

$$|\mathcal{A}| \leq \exp\left(O(\sqrt{Ks\epsilon^{-2}(\log n)^{3/2}})\right),$$

such that:

1. For all $A, B \subset [n]$, $C \subset [2n]$ with $|A| = |B| = s$, $A + B \subset C$ and $|C| \leq Ks$, there is a triple $(X, Y, Z) \in \mathcal{A}$ such that $A \subset X$, $B \subset Y$ and $Z \subset C$.

2. For every $(X, Y, Z) \in \mathcal{A}$, $|Z| \leq Ks$ and either $\max\{|X|, |Y|\} < Ks/\log n$ or there are at most $\epsilon^2|X||Y|$ pairs $(x, y) \in X \times Y$ such that $x + y \notin Z$. 
Proof Outline – The Container Theorem

We get the following theorem.

**Theorem (CCOW)**

Let $n, K, s$ be integers such that $\log n \leq s \leq Ks \leq s^2$ and $\epsilon > 0$. There is a family $\mathcal{A} \subset 2^{[n]} \times 2^{[n]} \times 2^{[2n]}$ of triples of sets $(X, Y, Z)$ of size

$$|\mathcal{A}| \leq \exp(O(\sqrt{Ks\epsilon^{-2}}(\log n)^{3/2})),$$

such that:

1. **For all** $A, B \subset [n], C \subset [2n]$ with $|A| = |B| = s$, $A + B \subset C$ and $|C| \leq Ks$, there is a triple $(X, Y, Z) \in \mathcal{A}$ such that $A \subset X$, $B \subset Y$ and $Z \subset C$.

2. **For every** $(X, Y, Z) \in \mathcal{A}$, $|Z| \leq Ks$ and either $\max\{|X|, |Y|\} < Ks/\log n$ or there are at most $\epsilon^2|X||Y|$ pairs $(x, y) \in X \times Y$ such that $x + y \notin Z$.

Note: The second property tells us that $\mathcal{H}(X, Y, [2n] \setminus Z)$ has few edges.
Proof Outline – A Stability Result for Distinct Sets of Different Sizes

- Note that while the sets $A, B$ in the end will have the same cardinality, the containers $X$ and $Y$ might not, so we need a stability result for this more general case.

Following arguments by Lev and Shao and Xu, we prove the following.

**Lemma**

Let $\varepsilon > \frac{2}{9}$. Let $X - Y - Z$ such that $\frac{j}{2} \leq Z \leq 9j$, with $|X| \geq j$ and $|Y| \geq j$, and $\max_f(X - f) \geq 1 - \frac{1}{4} \varepsilon$. Then one of the following holds:

1. There are at least $\varepsilon^2 j$ pairs $x - y \in X \times Y$ such that $x < y \notin Z$.

2. There are arithmetic progressions $P - Q$ of length at most $jZ \geq 2 \sqrt{\varepsilon j}$ with the same common difference such that $P$ contains all but at most $\varepsilon j$ points of $X$, and similarly for $Q$ and $Y$.

If the second property holds, we say that $X$ is $\varepsilon$-close to $P$. 
Proof Outline – A Stability Result for Distinct Sets of Different Sizes

- Note that while the sets $A$, $B$ in the end will have the same cardinality, the containers $X$ and $Y$ might not, so we need a stability result for this more general case.
- Following arguments by Lev and Shao and Xu, we prove the following.

**Lemma**

Let $\epsilon \leq 2^{-9}$. Let $X, Y, Z \subset \mathbb{Z}$ such that $(1 - \epsilon)|Z| \leq |X| + |Y|$ and $\max\{|X|, |Y|\} \leq (1 + 4\sqrt{\epsilon})|Z|/2$, then one of the following holds:
Proof Outline – A Stability Result for Distinct Sets of Different Sizes

■ Note that while the sets $A, B$ in the end will have the same cardinality, the containers $X$ and $Y$ might not, so we need a stability result for this more general case.

■ Following arguments by Lev and Shao and Xu, we prove the following.

Lemma

Let $\epsilon \leq 2^{-9}$. Let $X, Y, Z \subset \mathbb{Z}$ such that $(1 - \epsilon)|Z| \leq |X| + |Y|$ and $\max\{|X|, |Y|\} \leq (1 + 4\sqrt{\epsilon})|Z|/2$, then one of the following holds:

1. There are at least $\epsilon^2 |X||Y|$ pairs $(x, y) \in X \times Y$ such that $x + y \notin Z$. 
Proof Outline – A Stability Result for Distinct Sets of Different Sizes

- Note that while the sets $A$, $B$ in the end will have the same cardinality, the containers $X$ and $Y$ might not, so we need a stability result for this more general case.

- Following arguments by Lev and Shao and Xu, we prove the following.

**Lemma**

Let $\epsilon \leq 2^{-9}$. Let $X, Y, Z \subset \mathbb{Z}$ such that $(1 - \epsilon)|Z| \leq |X| + |Y|$ and $\max\{|X|, |Y|\} \leq (1 + 4\sqrt{\epsilon})|Z|/2$, then one of the following holds:

1. There are at least $\epsilon^2|X||Y|$ pairs $(x, y) \in X \times Y$ such that $x + y \notin Z$.
2. There are arithmetic progressions $P, Q$ of length at most $|Z|/2 + 3\sqrt{\epsilon}|Z|$ with the same common difference such that $P$ contains all but at most $\epsilon|X|$ points of $X$, and similarly for $Q$ and $Y$.

If the second property holds, we say that $X$ is $(\epsilon, |Z|)$-close to $P$. 
Proof Outline – Structure of the Container Family

We can now apply the stability result to $\mathcal{A}$ from our container theorem.
We can now apply the stability result to $\mathcal{A}$ from our container theorem. We see that every container triple $(X, Y, Z) \in \mathcal{A}$ satisfies one of the following properties.

1. $|X| + |Y| \leq (1 - \epsilon)Ks$,
2. $\max\{|X|, |Y|\} > (1 + 4\sqrt{\epsilon})Ks/2$, or
3. $X$ and $Y$ are $(\epsilon, Ks)$-close to arithmetic progressions with the same common difference.
We can now apply the stability result to $\mathcal{A}$ from our container theorem. We see that every container triple $(X, Y, Z) \in \mathcal{A}$ satisfies one of the following properties.

1. $|X| + |Y| \leq (1 - \epsilon)Ks$,
2. $\max\{|X|, |Y|\} > (1 + 4\sqrt{\epsilon})Ks/2$, or
3. $X$ and $Y$ are $(\epsilon, Ks)$-close to arithmetic progressions with the same common difference.

We now want to count the number of bad pairs, that is, $A, B \subset [n]$ of size $s$ such that $|A + B| \leq Ks$ that do not have large intersection with APs of size $Ks/2$.

Specifically, since the number of good pairs is at least $\binom{Ks/2}{s}^2$, we want the number of bad pairs to be asymptotically less than this.
Proof Outline – Counting the Exceptions (i)

Suppose the container triple \((X, Y, Z) \in \mathcal{A}\) satisfies

\[|X| + |Y| \leq (1 - \epsilon)Ks.\]
Proof Outline – Counting the Exceptions (i)

Suppose the container triple \((X, Y, Z) \in \mathcal{A}\) satisfies

\[ |X| + |Y| \leq (1 - \epsilon) Ks. \]

Then

\[
\begin{align*}
\binom{|X|}{s} \binom{|Y|}{s} &\leq \left( \frac{|X| + |Y|}{2} \right)^2 \\
&\leq \left( (1 - \epsilon) Ks/2 \right)^2 \\
&\leq (1 - \epsilon)^2 s \left( \frac{Ks/2}{s} \right)^2.
\end{align*}
\]
Proof Outline – Counting the Exceptions (i)

Suppose the container triple \((X, Y, Z) \in \mathcal{A}\) satisfies

\[|X| + |Y| \leq (1 - \epsilon)Ks.\]

Then

\[
\binom{|X|}{s} \binom{|Y|}{s} \leq \left( \binom{|X| + |Y|}{2s} \right)^2 \leq \left( \frac{(1 - \epsilon)Ks/2}{s} \right)^2 \leq (1 - \epsilon)^{2s} \left( \frac{Ks/2}{s} \right)^2.
\]

Choosing \(\epsilon\) carefully while noting that \(|\mathcal{A}| \leq \exp(O(\sqrt{Ks\epsilon^{-2}(\log n)^{3/2}}))\) we can union bound over all of it and get

\[|\mathcal{A}|(1 - \epsilon)^{2s} \left(\frac{Ks/2}{s}\right)^2 = o(1) \left(\frac{Ks/2}{s}\right)^2.\]
Proof Outline – Counting the Exceptions (i)

Suppose the container triple \((X, Y, Z) \in \mathcal{A}\) satisfies

\[ |X| + |Y| \leq (1 - \epsilon)Ks. \]

Then

\[
\binom{|X|}{s}\binom{|Y|}{s} \leq \left(\frac{|X| + |Y|}{2}\right)^2 \leq \left(\frac{(1 - \epsilon)Ks/2}{s}\right)^2 \leq (1 - \epsilon)^{2s}\binom{Ks/2}{s}^2.
\]

Choosing \(\epsilon\) carefully while noting that \(|\mathcal{A}| \leq \exp(O(\sqrt{Ks\epsilon^{-2}(\log n)^{3/2}}))\) we can union bound over all of it and get

\[
|\mathcal{A}|(1 - \epsilon)^{2s}\binom{Ks/2}{s}^2 = o(1)\binom{Ks/2}{s}^2.
\]

Note: We have actually counted the number of all pairs of \(s\)-sets of small doubling, regardless whether they are good or bad.
Finally, suppose the container triple $(X, Y, Z) \in \mathcal{A}$ satisfies

- $X$ and $Y$ are $(\epsilon, Ks)$-close to arithmetic progressions $P, Q$ with the same common difference.
Finally, suppose the container triple $(X, Y, Z) \in \mathcal{A}$ satisfies

- $X$ and $Y$ are $(\epsilon, Ks)$-close to arithmetic progressions $P, Q$ with the same common difference.

For this case, we actually want to discern between good and bad pairs of $s$-sets contained here. The moral is, if the containers are $(\epsilon, Ks)$ close to APs, then there shouldn’t be many pairs of sets within them that are not at least $(50\epsilon, Ks)$-close to the same APs.
Proof Outline – Counting the Exceptions (ii)

Finally, suppose the container triple \((X, Y, Z) \in \mathcal{A}\) satisfies

- \(X\) and \(Y\) are \((\varepsilon, Ks)\)-close to arithmetic progressions \(P, Q\) with the same common difference.

For this case, we actually want to discern between good and bad pairs of \(s\)-sets contained here. The moral is, if the containers are \((\varepsilon, Ks)\) close to APs, then there shouldn’t be many pairs of sets within them that are not at least \((50\varepsilon, Ks)\)-close to the same APs.

Specifically, for any such bad pair, it must hold that for one of the \(s\)-sets (let’s say \(A \subset X\)), there is an \(s' \geq 50\varepsilon s\) such that \(s'\) elements of \(A\) do not lie in \(P\).
Proof Outline – Counting the Exceptions (ii) cont.

But only $\epsilon|X|$ points of $X$ itself are outside of $P$, so the number of bad pairs for fixed containers is at most

$$\sum_{s' = 50\epsilon s}^{s} \binom{|X|}{s - s'}\binom{\epsilon|X|}{s'}\binom{|Y|}{s} + \binom{|X|}{s}\binom{|Y|}{s - s'}\binom{\epsilon|Y|}{s'}$$

We can now again use standard bounds for the binomial coefficient and union bound over all of $\mathcal{A}$ to get that there are $o(1)\left(\frac{Ks/2}{s}\right)^2$ bad pairs coming from containers $X, Y$ of this type.
Using these counting results and the fact that one can take a single pair of disjoint arithmetic progressions $P, Q$ with the same common difference of size $K s / 2$, we arrive again at our main result.
Using these counting results and the fact that one can take a single pair of disjoint arithmetic progressions $P, Q$ with the same common difference of size $Ks/2$, we arrive again at our main result.

**Theorem (Campos, Coulson, Serra, W.)**

Let $s = \Omega((\log n)^3)$ and $K = O(s/(\log)^3)$. Then for almost all sets $A, B \subset [n]$ with $|A| = |B| = s$ and $|A + B| \leq Ks$ there exist arithmetic progressions $P, Q$ with the same common difference of size

$$|P|, |Q| \leq \frac{1 + o(1)}{2} Ks$$

such that $|A \setminus P|, |B \setminus Q| = o(s)$.
Open Questions

- What if $A$ and $B$ are not of the same size?
- More than two distinct summands?
- Groups other than $\mathbb{Z}$?
- What about the precise structural version of CCMMS?
Thank you for your attention!