On the Relative Sizes of Complements of Generalized Sumsets

Ryan Jeong
rsjeong@sas.upenn.edu

Joint work with Steven J. Miller

University of Pennsylvania

The 21st International Conference on Random Structures and Algorithms

June 12–16, 2023

More Sums Than Difference Sets (MSTD Sets)

Let $A \subseteq \{0, 1, \ldots, N\}$.

More Sums Than Difference Sets (MSTD Sets)

Let $A \subseteq \{0, 1, \ldots, N\}$.

Definition

The **sumset** $A + A$ and **difference set** $A - A$ of A are

$$A + A := \{a + a' : a, a' \in A\},$$
$$A - A := \{a - a' : a, a' \in A\}.$$

More Sums Than Difference Sets (MSTD Sets)

Let $A \subseteq \{0, 1, \ldots, N\}$.

Definition

The **sumset** $A + A$ and **difference set** $A - A$ of A are

$$A + A := \{a + a' : a, a' \in A\},$$
$$A - A := \{a - a' : a, a' \in A\}.$$

We would expect a generic set A to satisfy $|A - A| \geq |A + A|$, since addition commutes while subtraction doesn't.

More Sums Than Difference Sets (MSTD Sets)

Let $A \subseteq \{0, 1, \ldots, N\}$.

Definition

The **sumset** $A + A$ and **difference set** $A - A$ of A are

$$A + A := \{a + a' : a, a' \in A\},$$
$$A - A := \{a - a' : a, a' \in A\}.$$

We would expect a generic set A to satisfy $|A - A| \geq |A + A|$, since addition commutes while subtraction doesn't.

Definition

We say A is **MSTD** if $|A + A| > |A - A|$.

Existence of MSTD Sets

MSTD sets exist: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ (Conway, 1960s).

Existence of MSTD Sets

MSTD sets exist: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ (Conway, 1960s).

Hegarty, Nathanson, Zhao: Infinite families of MSTD sets.

Existence of MSTD Sets

MSTD sets exist: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ (Conway, 1960s).

Hegarty, Nathanson, Zhao: Infinite families of MSTD sets.

"Even though there exist sets A that have more sums than differences, such sets should be rare, and it must be true with the right way of counting that the vast majority of sets satisfies $|A - A| > |A + A|$." - Melvyn B. Nathanson, 2006.

Existence of MSTD Sets

MSTD sets exist: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ (Conway, 1960s).

Hegarty, Nathanson, Zhao: Infinite families of MSTD sets.

"Even though there exist sets A that have more sums than differences, such sets should be rare, and it must be true with the right way of counting that the vast majority of sets satisfies $|A - A| > |A + A|$." - Melvyn B. Nathanson, 2006.

Question

What is the "right way of counting"?

Uniform Model

First try: Pick A uniformly at random from $2^{\{0,1,\ldots,N\}}$.
(Equivalently, independently include each element w.p. $\frac{1}{2}$.)

Uniform Model

First try: Pick A uniformly at random from $2^{\{0,1,\ldots,N\}}$.
(Equivalently, independently include each element w.p. $\frac{1}{2}$.)

Theorem (Martin and O'Bryant, 2006)

For $N \geq 15$, there exists a constant $c > 0$ such that

$$\frac{\#\{A \subseteq \{0,1,\ldots,N\} : A \text{ is sum-dominated}\}}{2^{N+1}} > c.$$

Uniform Model

First try: Pick A uniformly at random from $2^{\{0,1,\ldots,N\}}$.
(Equivalently, independently include each element w.p. $\frac{1}{2}$.)

Theorem (Martin and O'Bryant, 2006)

For $N \geq 15$, there exists a constant $c > 0$ such that

$$\frac{\#\{A \subseteq \{0,1,\ldots,N\} : A \text{ is sum-dominated}\}}{2^{N+1}} > c.$$

Intuition:

- With high probability, $A + A$ and $A - A$ will hit everything in the middle of $[0, 2N]$ and $[-N, N]$, respectively.

Uniform Model

First try: Pick A uniformly at random from $2^{\{0,1,\ldots,N\}}$.
(Equivalently, independently include each element w.p. $\frac{1}{2}$.)

Theorem (Martin and O'Bryant, 2006)

For $N \geq 15$, there exists a constant $c > 0$ such that

$$\frac{\#\{A \subseteq \{0,1,\ldots,N\} : A \text{ is sum-dominated}\}}{2^{N+1}} > c.$$

Intuition:

- With high probability, $A+A$ and $A-A$ will hit everything in the middle of $[0,2N]$ and $[-N,N]$, respectively.
- We can then "rig" the fringes of $A+A$ and $A-A$ by carefully selecting the fringes of A.

Binomial Model

Second try: Keep the same model, but now let the inclusion probability be $p(N) \asymp N^{-\delta}$ for some $\delta \in (0, 1)$.

Binomial Model

Second try: Keep the same model, but now let the inclusion probability be $p(N) \asymp N^{-\delta}$ for some $\delta \in (0,1)$.

"Inclusion probability vanishes, but $|A|$ explodes."

Binomial Model

Second try: Keep the same model, but now let the inclusion probability be $p(N) \asymp N^{-\delta}$ for some $\delta \in (0,1)$.

"Inclusion probability vanishes, but $|A|$ explodes."

Conjecture (Martin and O'Bryant, 2006)

When A is drawn from the binomial model, the set A is not MSTD with high probability.

Binomial Model

Second try: Keep the same model, but now let the inclusion probability be $p(N) \asymp N^{-\delta}$ for some $\delta \in (0,1)$.

"Inclusion probability vanishes, but $|A|$ explodes."

Conjecture (Martin and O'Bryant, 2006)

When A is drawn from the binomial model, the set A is not MSTD with high probability.

Theorem (Hegarty and Miller, 2008)

- *(Fast Decay) If $\delta > \frac{1}{2}$, $\frac{|A-A|}{|A+A|} \sim 2$.*
- *(Critical Decay) If $p(N) = cN^{-1/2}$, $\frac{|A-A|}{|A+A|} \sim f(c) \searrow 1$.*
- *(Slow Decay) If $\delta < \frac{1}{2}$, $\frac{|(A+A)^c|}{|(A-A)^c|} \sim 2$.*

Generalized Sumsets

Definition

The **generalized sumset** $A_{s,d}$ of A with s sums, d differences is

$$A_{s,d} := \{a_1 + \cdots + a_s - b_1 - \cdots - b_d : a_1, \ldots, a_s, b_1, \ldots, b_d \in A\}.$$

Generalized Sumsets

Definition

The **generalized sumset** $A_{s,d}$ of A with s sums, d differences is

$$A_{s,d} := \{a_1 + \cdots + a_s - b_1 - \cdots - b_d : a_1, \ldots, a_s, b_1, \ldots, b_d \in A\}.$$

Let $s_1 + d_1 = s_2 + d_2 = h$, with $p(N) \asymp N^{-\delta}$ as before.

Generalized Sumsets

Definition

The **generalized sumset** $A_{s,d}$ of A with s sums, d differences is

$$A_{s,d} := \{a_1 + \cdots + a_s - b_1 - \cdots - b_d : a_1, \ldots, a_s, b_1, \ldots, b_d \in A\}.$$

Let $s_1 + d_1 = s_2 + d_2 = h$, with $p(N) \asymp N^{-\delta}$ as before.

Theorem (Hogan and Miller, 2013)

- *(Fast Decay) If* $\delta > \frac{h-1}{h}$, $\frac{|A_{s_1,d_1}|}{|A_{s_2,d_2}|} \sim \frac{s_2!d_2!}{s_1!d_1!}$.
- *(Critical Decay) If* $p(N) = cN^{-\frac{h-1}{h}}$, $\frac{|A_{s_1,d_1}|}{|A_{s_2,d_2}|} \sim \frac{f(c,s_1,d_1)}{f(c,s_2,d_2)}$.

Generalized Sumsets

Definition

The **generalized sumset** $A_{s,d}$ of A with s sums, d differences is

$$A_{s,d} := \{a_1 + \cdots + a_s - b_1 - \cdots - b_d : a_1, \ldots, a_s, b_1, \ldots, b_d \in A\}.$$

Let $s_1 + d_1 = s_2 + d_2 = h$, with $p(N) \asymp N^{-\delta}$ as before.

Theorem (Hogan and Miller, 2013)

- *(Fast Decay) If* $\delta > \frac{h-1}{h}$, $\frac{|A_{s_1,d_1}|}{|A_{s_2,d_2}|} \sim \frac{s_2!d_2!}{s_1!d_1!}$.
- *(Critical Decay) If* $p(N) = cN^{-\frac{h-1}{h}}$, $\frac{|A_{s_1,d_1}|}{|A_{s_2,d_2}|} \sim \frac{f(c,s_1,d_1)}{f(c,s_2,d_2)}$.

Proposition (J., Miller, 2023++)

We have that $\frac{f(c,s_1,d_1)}{f(c,s_2,d_2)} \searrow 1$.

Slow Decay/Dense A Regime

Arguments in Hegarty and Miller proceed in two steps.

Slow Decay/Dense A Regime

Arguments in Hegarty and Miller proceed in two steps.

1. Estimate $\mathbb{E}[|A+A|]$, $\mathbb{E}[|A-A|]$, $\mathbb{E}[|(A+A)^c|]$, $\mathbb{E}[|(A-A)^c|]$.

Slow Decay/Dense A Regime

Arguments in Hegarty and Miller proceed in two steps.

1. Estimate $\mathbb{E}[|A+A|]$, $\mathbb{E}[|A-A|]$, $\mathbb{E}[|(A+A)^c|]$, $\mathbb{E}[|(A-A)^c|]$.
2. Show that these random variables are strongly concentrated about their expectations.

Slow Decay/Dense A Regime

Arguments in Hegarty and Miller proceed in two steps.

1. Estimate $\mathbb{E}[|A+A|]$, $\mathbb{E}[|A-A|]$, $\mathbb{E}[|(A+A)^c|]$, $\mathbb{E}[|(A-A)^c|]$.
2. Show that these random variables are strongly concentrated about their expectations.

Step (1) is easy for the complements when there are two summands:

$$\mathbb{E}[|(A+A)^c|] = \sum_{k=0}^{2N} \Pr[k \notin A+A],$$

and $k \notin A+A$ is the union of the mutually independent events

$$\{0,k\} \not\subseteq A, \{1,k-1\} \not\subseteq A, \ldots, \{\lfloor k/2 \rfloor, \lceil k/2 \rceil\} \not\subseteq A.$$

Slow Decay/Dense A Regime

Arguments in Hegarty and Miller proceed in two steps.

1. Estimate $\mathbb{E}[|A+A|]$, $\mathbb{E}[|A-A|]$, $\mathbb{E}[|(A+A)^c|]$, $\mathbb{E}[|(A-A)^c|]$.
2. Show that these random variables are strongly concentrated about their expectations.

Step (1) is easy for the complements when there are two summands:

$$\mathbb{E}[|(A+A)^c|] = \sum_{k=0}^{2N} \Pr[k \notin A+A],$$

and $k \notin A+A$ is the union of the mutually independent events

$$\{0,k\} \not\subseteq A, \{1,k-1\} \not\subseteq A, \ldots, \{\lfloor k/2 \rfloor, \lceil k/2 \rceil\} \not\subseteq A.$$

Harder to compute exclusion probabilities for three or more summands.

Main Result

We include each element in $\{0, 1, \ldots, N\}$ in A independently with probability $p(N) \asymp N^{-\delta}$, where $\delta < \frac{h-1}{h}$.

Theorem (J., Miller, 2023++)

If $s_1 + d_1 = s_2 + d_2 = h$ *and* $\delta < \frac{h-1}{h}$,

$$\frac{|A^c_{s_1,d_1}|}{|A^c_{s_2,d_2}|} \sim \left(\frac{s_1!d_1!}{s_2!d_2!}\right)^{\frac{1}{h-1}}.$$

Main Result

We include each element in $\{0, 1, \ldots, N\}$ in A independently with probability $p(N) \asymp N^{-\delta}$, where $\delta < \frac{h-1}{h}$.

Theorem (J., Miller, 2023++)

If $s_1 + d_1 = s_2 + d_2 = h$ *and* $\delta < \frac{h-1}{h}$,

$$\frac{|A^c_{s_1,d_1}|}{|A^c_{s_2,d_2}|} \sim \left(\frac{s_1!d_1!}{s_2!d_2!}\right)^{\frac{1}{h-1}}.$$

Takeaways:

- When comparing sizes of complements in the slow decay regime, there is always a limit in probability.

Main Result

We include each element in $\{0, 1, \ldots, N\}$ in A independently with probability $p(N) \asymp N^{-\delta}$, where $\delta < \frac{h-1}{h}$.

Theorem (J., Miller, 2023++)

If $s_1 + d_1 = s_2 + d_2 = h$ *and* $\delta < \frac{h-1}{h}$,

$$\frac{\left|A^c_{s_1,d_1}\right|}{\left|A^c_{s_2,d_2}\right|} \sim \left(\frac{s_1!d_1!}{s_2!d_2!}\right)^{\frac{1}{h-1}}.$$

Takeaways:

- When comparing sizes of complements in the slow decay regime, there is always a limit in probability.
- The limiting ratio of the main terms in the fast decay regime generalizes differently from the limiting ratio of the complements in the slow decay regime.

Main Ingredients in the Proof

The proof follows from three key insights.

Main Ingredients in the Proof

The proof follows from three key insights.

1. All of the missing elements lie in the fringes: we hit basically everything in the middle.

Main Ingredients in the Proof

The proof follows from three key insights.

1. All of the missing elements lie in the fringes: we hit basically everything in the middle.
2. The number of ways that we hit a particular sum in the fringes converges in distribution to a Poisson with rate that we can calculate.

Main Ingredients in the Proof

The proof follows from three key insights.

1. All of the missing elements lie in the fringes: we hit basically everything in the middle.
2. The number of ways that we hit a particular sum in the fringes converges in distribution to a Poisson with rate that we can calculate.
3. The number of missing elements is strongly concentrated about its expectation.

Main Ingredients in the Proof

The proof follows from three key insights.

1. All of the missing elements lie in the fringes: we hit basically everything in the middle.
2. The number of ways that we hit a particular sum in the fringes converges in distribution to a Poisson with rate that we can calculate.
3. The number of missing elements is strongly concentrated about its expectation.

Henceforth: Everything will be under the context of $A_{h,0}$, which takes values in $[0, hN]$. We also assume that we only allow distinct summands.

Step 1: The Fringes

We want to formalize “hitting everything in the middle.”

Step 1: The Fringes

We want to formalize "hitting everything in the middle."

To study the left fringe, think of generating A by a sequence of Geom(p) random variables (gaps between consecutive terms).

Step 1: The Fringes

We want to formalize "hitting everything in the middle."

To study the left fringe, think of generating A by a sequence of Geom(p) random variables (gaps between consecutive terms).

1. The $\frac{(1/p)^{\frac{1}{h-1}}}{h}$th element of A is $\sim \frac{(1/p)^{\frac{h}{h-1}}}{h}$.
2. New sums in $A_{h,0}$ from adding the k^{th} element of A beat sums with summands before the $\sim (k/h)^{\text{th}}$ element of A.

Step 1: The Fringes

We want to formalize "hitting everything in the middle."

To study the left fringe, think of generating A by a sequence of Geom(p) random variables (gaps between consecutive terms).

1. The $\frac{(1/p)^{\frac{1}{h-1}}}{h}$th element of A is $\sim \frac{(1/p)^{\frac{h}{h-1}}}{h}$.
2. New sums in $A_{h,0}$ from adding the k^{th} element of A beat sums with summands before the $\sim (k/h)^{\text{th}}$ element of A.

We therefore have the asymptotic lower bound

$$\mathbb{E}[|A_{h,0}^c|] \gtrsim h \cdot \frac{(1/p)^{\frac{h}{h-1}}}{h} - \sum_{i=1}^{(1/p)^{\frac{1}{h-1}}} i^{h-1} \gtrsim (1/p)^{\frac{h}{h-1}}.$$

Step 1: The Middle

We bound exclusion probabilities $\Pr[k \notin A_{h,0}]$ with Janson.

Step 1: The Middle

We bound exclusion probabilities $\Pr[k \notin A_{h,0}]$ with Janson.

It turns out that if $k \in [\tau N, (h-\tau)N]$, with $\tau \asymp \frac{\left(\log\left[Np^{\frac{h}{h-1}}\right]\right)^{\frac{1}{h-1}}}{Np^{\frac{h}{h-1}}}$,

$$\Pr[k \notin A_{h,0}] \leq e^{-Ck^{h-1}p^h}.$$

Step 1: The Middle

We bound exclusion probabilities $\Pr[k \notin A_{h,0}]$ with Janson.

It turns out that if $k \in [\tau N, (h-\tau)N]$, with $\tau \asymp \frac{\left(\log\left[Np^{\frac{h}{h-1}}\right]\right)^{\frac{1}{h-1}}}{Np^{\frac{h}{h-1}}}$,

$$\Pr[k \notin A_{h,0}] \leq e^{-Ck^{h-1}p^h}.$$

Expected number of missing elements in $[\tau N, (h-\tau)N]$ is at most

$$\sum_{k \in [\tau N, (h-\tau)N]} e^{-k^{h-1}p^h} \lesssim (1/p)^{\frac{h}{h-1}} \int_{\log\left[Np^{\frac{h}{h-1}}\right]}^{\infty} e^{-Cx^{h-1}} dx = o\left((1/p)^{\frac{h}{h-1}}\right).$$

Step 1: The Middle

We bound exclusion probabilities $\Pr[k \notin A_{h,0}]$ with Janson.

It turns out that if $k \in [\tau N, (h-\tau)N]$, with $\tau \asymp \frac{\left(\log\left[Np^{\frac{h}{h-1}}\right]\right)^{\frac{1}{h-1}}}{Np^{\frac{h}{h-1}}}$,

$$\Pr[k \notin A_{h,0}] \leq e^{-Ck^{h-1}p^h}.$$

Expected number of missing elements in $[\tau N, (h-\tau)N]$ is at most

$$\sum_{k\in[\tau N,(h-\tau)N]} e^{-k^{h-1}p^h} \lesssim (1/p)^{\frac{h}{h-1}} \int_{\log\left[Np^{\frac{h}{h-1}}\right]}^{\infty} e^{-Cx^{h-1}} dx$$
$$= o\left((1/p)^{\frac{h}{h-1}}\right).$$

Combined with Last Slide: All of the missing elements lie in $[\tau N, (h-\tau)N]^c = [0, \tau N] \cup [(h-\tau)N, hN]$, the fringes.

Step 2: Poisson Convergence

For each possible sum k in the fringes, we invoke the Stein-Chen method (Arratia, Goldstein, Gordon, 1989) to prove that the number of ways k is hit converges to a Poisson.

Step 2: Poisson Convergence

For each possible sum k in the fringes, we invoke the Stein-Chen method (Arratia, Goldstein, Gordon, 1989) to prove that the number of ways k is hit converges to a Poisson.

The expected number of missing elements in the left fringe, $[0, \tau N]$, is

$$\sum_{k=0}^{\tau N} \Pr[k \notin A_{h,0}] \stackrel{(1)}{\sim} \sum_{k=0}^{\tau N} \Pr[\mathrm{Pois}(\lambda_k) = 0] \sim \sum_{k=0}^{\tau N} e^{-\lambda_k}$$

$$\stackrel{(2)}{\sim} (1/p)^{\frac{h}{h-1}} \int_0^\infty e^{-C_h x^{h-1}}\, dx.$$

Step 2: Poisson Convergence

For each possible sum k in the fringes, we invoke the Stein-Chen method (Arratia, Goldstein, Gordon, 1989) to prove that the number of ways k is hit converges to a Poisson.

The expected number of missing elements in the left fringe, $[0, \tau N]$, is

$$\sum_{k=0}^{\tau N} \Pr[k \notin A_{h,0}] \overset{(1)}{\sim} \sum_{k=0}^{\tau N} \Pr[\mathrm{Pois}(\lambda_k) = 0] \sim \sum_{k=0}^{\tau N} e^{-\lambda_k}$$

$$\overset{(2)}{\sim} (1/p)^{\frac{h}{h-1}} \int_0^{\infty} e^{-C_h x^{h-1}} \, dx.$$

1. The probability that each Poisson random variable is 0 dominates the bound on the total variation distance from Stein-Chen.

Step 2: Poisson Convergence

For each possible sum k in the fringes, we invoke the Stein-Chen method (Arratia, Goldstein, Gordon, 1989) to prove that the number of ways k is hit converges to a Poisson.

The expected number of missing elements in the left fringe, $[0, \tau N]$, is

$$\sum_{k=0}^{\tau N} \Pr[k \notin A_{h,0}] \overset{(1)}{\sim} \sum_{k=0}^{\tau N} \Pr[\text{Pois}(\lambda_k) = 0] \sim \sum_{k=0}^{\tau N} e^{-\lambda_k}$$

$$\overset{(2)}{\sim} (1/p)^{\frac{h}{h-1}} \int_0^\infty e^{-C_h x^{h-1}} \, dx.$$

1. The probability that each Poisson random variable is 0 dominates the bound on the total variation distance from Stein-Chen.
2. The rate λ_k is the expected number of combinations with h distinct terms that sum to k. The number of ways to partition k into h distinct parts, for large k, is $\sim C_h k^{h-1}$ (Knessl and Keller, 1990). So $\lambda_k \sim C_h p^h k^{h-1}$.

Step 2: Ratio in Expectation

If we have d minus signs: Partitions of k into h distinct "offsets." A partition of k into h distinct offsets can be realized in $\binom{h}{d}$ ways.

Step 2: Ratio in Expectation

If we have d minus signs: Partitions of k into h distinct "offsets." A partition of k into h distinct offsets can be realized in $\binom{h}{d}$ ways.

$$\mathbb{E}\left[\left|A_{s,d}^c\right|\right] \sim 2 \cdot (1/p)^{\frac{h}{h-1}} \int_0^\infty e^{-C_h \binom{h}{d} x^{h-1}}\, dx$$
$$= \frac{2 \cdot (1/p)^{\frac{h}{h-1}}}{\sqrt[h-1]{\binom{h}{d}}} \int_0^\infty e^{-C_h x^{h-1}}\, dx.$$

Step 2: Ratio in Expectation

If we have d minus signs: Partitions of k into h distinct "offsets." A partition of k into h distinct offsets can be realized in $\binom{h}{d}$ ways.

$$\mathbb{E}\left[\left|A^c_{s,d}\right|\right] \sim 2\cdot(1/p)^{\frac{h}{h-1}} \int_0^\infty e^{-C_h\binom{h}{d}x^{h-1}}\,dx$$
$$= \frac{2\cdot(1/p)^{\frac{h}{h-1}}}{\sqrt[h-1]{\binom{h}{d}}}\int_0^\infty e^{-C_h x^{h-1}}\,dx.$$

Altogether, we have that

$$\frac{\mathbb{E}\left[\left|A^c_{s_1,d_1}\right|\right]}{\mathbb{E}\left[\left|A^c_{s_2,d_2}\right|\right]} \sim \frac{\sqrt[h-1]{\binom{h}{d_2}}}{\sqrt[h-1]{\binom{h}{d_1}}} = \left(\frac{s_1!d_1!}{s_2!d_2!}\right)^{\frac{1}{h-1}}.$$

Step 2: Ratio in Expectation

If we have d minus signs: Partitions of k into h distinct "offsets." A partition of k into h distinct offsets can be realized in $\binom{h}{d}$ ways.

$$\begin{aligned}\mathbb{E}\left[\left|A^c_{s,d}\right|\right] &\sim 2\cdot(1/p)^{\frac{h}{h-1}}\int_0^\infty e^{-C_h\binom{h}{d}x^{h-1}}\,dx\\ &= \frac{2\cdot(1/p)^{\frac{h}{h-1}}}{\sqrt[h-1]{\binom{h}{d}}}\int_0^\infty e^{-C_h x^{h-1}}\,dx.\end{aligned}$$

Altogether, we have that

$$\frac{\mathbb{E}\left[\left|A^c_{s_1,d_1}\right|\right]}{\mathbb{E}\left[\left|A^c_{s_2,d_2}\right|\right]} \sim \frac{\sqrt[h-1]{\binom{h}{d_2}}}{\sqrt[h-1]{\binom{h}{d_1}}} = \left(\frac{s_1!d_1!}{s_2!d_2!}\right)^{\frac{1}{h-1}}.$$

Proves the theorem for the expectations.

Step 3: Reductions from Martingale Machinery

For $k \in \{0, \ldots, N\}$, define the quantities

$$\Delta_k(A) = \mathbb{E}\left[|A^c_{h,0}| \;\middle|\; A \cap [0, k-1], k \in A\right] - \mathbb{E}\left[|A^c_{h,0}| \;\middle|\; A \cap [0, k-1], k \notin A\right],$$

$$C(A) = \max_{0 \le k \le N} \Delta_k(A), \qquad V(A) = p \sum_{k=0}^{N} (\Delta_k(A))^2 .$$

Step 3: Reductions from Martingale Machinery

For $k \in \{0, \ldots, N\}$, define the quantities

$$\Delta_k(A) = \mathbb{E}\left[|A^c_{h,0}| \;\middle|\; A \cap [0, k-1], k \in A\right] - \mathbb{E}\left[|A^c_{h,0}| \;\middle|\; A \cap [0, k-1], k \notin A\right],$$

$$C(A) = \max_{0 \le k \le N} \Delta_k(A), \qquad V(A) = p \sum_{k=0}^{N} (\Delta_k(A))^2 .$$

From a result of Vu (2002), for any $\lambda, V, C > 0$ such that $\lambda \le \frac{4V}{C^2}$,

$$\mathbb{P}\left(\left||A^c_{h,0}| - \mathbb{E}\left[|A^c_{h,0}|\right]\right| \ge \sqrt{\lambda V}\right) \le 2e^{-\lambda/4} + \mathbb{P}(C(A) \ge C) + \mathbb{P}(V(A) \ge V).$$

Step 3: Reductions from Martingale Machinery

For $k \in \{0, \ldots, N\}$, define the quantities

$$\Delta_k(A) = \mathbb{E}\left[|A^c_{h,0}| \;\middle|\; A \cap [0, k-1], k \in A\right] - \mathbb{E}\left[|A^c_{h,0}| \;\middle|\; A \cap [0, k-1], k \notin A\right],$$

$$C(A) = \max_{0 \le k \le N} \Delta_k(A), \qquad V(A) = p \sum_{k=0}^{N} (\Delta_k(A))^2 .$$

From a result of Vu (2002), for any $\lambda, V, C > 0$ such that $\lambda \le \frac{4V}{C^2}$,

$$\mathbb{P}\left(\left||A^c_{h,0}| - \mathbb{E}\left[|A^c_{h,0}|\right]\right| \ge \sqrt{\lambda V}\right) \le 2e^{-\lambda/4} + \mathbb{P}(C(A) \ge C) + \mathbb{P}(V(A) \ge V).$$

It will suffice to show the RHS vanishes if we take

$$\lambda \asymp \log(1/p), \quad V \asymp ((1/p)\log(1/p))^{1+\frac{h}{h-1}}, \quad C \asymp (1/p)\log(1/p),$$

since then, $\lambda \le \frac{4V}{C^2}$, $\sqrt{\lambda V} = o\left((1/p)^{\frac{h}{h-1}}\right)$, and $e^{-\lambda/4}$ vanishes.

Step 3: Sketch of Strong Concentration

Fix some $k \in \{0, \ldots, N\}$.

1. (Middle) Number of elements in $[\tilde{\tau}N, (h - \tilde{\tau})N]$ that adding k to A will add to $A_{h,0}$ is $o(1)$. (Adapt argument from Step 1, with a larger threshold $\tilde{\tau} \asymp \frac{(\log N)^{\frac{1}{h-1}}}{Np^{\frac{h}{h-1}}}$ separating the fringe from the middle.)

Step 3: Sketch of Strong Concentration

Fix some $k \in \{0, \ldots, N\}$.

1. (Middle) Number of elements in $[\tilde{\tau}N, (h-\tilde{\tau})N]$ that adding k to A will add to $A_{h,0}$ is $o(1)$. (Adapt argument from Step 1, with a larger threshold $\tilde{\tau} \asymp \frac{(\log N)^{\frac{1}{h-1}}}{Np^{\frac{h}{h-1}}}$ separating the fringe from the middle.)
2. (Fringes) Number of elements in $[\tilde{\tau}N, (h-\tilde{\tau})N]^c$ that adding k to A will add to $A_{h,0}$ is $\lesssim (1/p)\log(1/p)$. (Adding k to A cannot add more new sums to $A_{h,0}$ than $\asymp$ (number of fringe elements of $A)^{h-1}$; Chernoff yields size of fringes to be $\asymp [(1/p)\log(1/p)]^{\frac{1}{h-1}}$.)

Step 3: Sketch of Strong Concentration

Fix some $k \in \{0, \ldots, N\}$.

1. (Middle) Number of elements in $[\tilde{\tau}N, (h-\tilde{\tau})N]$ that adding k to A will add to $A_{h,0}$ is $o(1)$. (Adapt argument from Step 1, with a larger threshold $\tilde{\tau} \asymp \frac{(\log N)^{\frac{1}{h-1}}}{Np^{\frac{h}{h-1}}}$ separating the fringe from the middle.)
2. (Fringes) Number of elements in $[\tilde{\tau}N, (h-\tilde{\tau})N]^c$ that adding k to A will add to $A_{h,0}$ is $\lesssim (1/p)\log(1/p)$. (Adding k to A cannot add more new sums to $A_{h,0}$ than $\asymp$ (number of fringe elements of $A)^{h-1}$; Chernoff yields size of fringes to be $\asymp [(1/p)\log(1/p)]^{\frac{1}{h-1}}$.)

A union bound with vanishing sum of error probabilities gives

$$\mathbb{P}(C(A) \geq C) = \mathbb{P}\left(\max_{0\leq k\leq N} \Delta_k(A) \geq \kappa_1(1/p)\log(1/p)\right) \xrightarrow{N\to\infty} 0.$$

Step 3: Sketch of Strong Concentration

Fix some $k \in \{0, \dots, N\}$.

1. (Middle) Number of elements in $[\tilde{\tau}N, (h - \tilde{\tau})N]$ that adding k to A will add to $A_{h,0}$ is $o(1)$. (Adapt argument from Step 1, with a larger threshold $\tilde{\tau} \asymp \frac{(\log N)^{\frac{1}{h-1}}}{Np^{\frac{h}{h-1}}}$ separating the fringe from the middle.)
2. (Fringes) Number of elements in $[\tilde{\tau}N, (h - \tilde{\tau})N]^c$ that adding k to A will add to $A_{h,0}$ is $\lesssim (1/p)\log(1/p)$. (Adding k to A cannot add more new sums to $A_{h,0}$ than $\asymp$ (number of fringe elements of $A)^{h-1}$; Chernoff yields size of fringes to be $\asymp [(1/p)\log(1/p)]^{\frac{1}{h-1}}$.)

A union bound with vanishing sum of error probabilities gives

$$\mathbb{P}(C(A) \geq C) = \mathbb{P}\left(\max_{0 \leq k \leq N} \Delta_k(A) \geq \kappa_1 (1/p)\log(1/p)\right) \xrightarrow{N \to \infty} 0.$$

Furthermore, $V(A) \lesssim p\tilde{\tau}N\,[(1/p)\log(1/p)]^2 \asymp ((1/p)\log(1/p))^{1+\frac{h}{h-1}}$, so

$$\mathbb{P}(V(A) \geq V) = \mathbb{P}(V(A) \geq \kappa_2\,((1/p)\log(1/p))^{1+\frac{h}{h-1}}) \xrightarrow{N \to \infty} 0.$$

Thank You for Listening!

This research was supported, in part, by NSF Grant DMS1947438. We thank Williams College and the University of Pennsylvania for their support.