Phase Transitions for Binomial Sets under Linear Forms

Advisor: Steven J. Miller (Fibonacci Association, Williams)
Ryan Jeong (Leland Stanford Junior University)

rsjeong@stanford.edu, sjm1@williams.edu

https://web.williams.edu/Mathematics/sjmiller/public_html/

Preprint: https://arxiv.org/pdf/2309.01801

CANT, May 21, 2025

Polymath Jr

https://geometrynyc.wixsite.com/polymathreu

Our goal is to provide research opportunities to every undergraduate who wishes to explore advanced mathematics. This online program consists of research projects in a variety of mathematical topics and runs in the spirit of the Polymath Project. Each project is mentored by an active researcher with experience in undergraduate mentoring.

Each project consists of 15-25 undergraduates, a main mentor, and graduate students and postdocs as additional mentors. The group works towards solving a research problem and writing a paper. Each participant decides what they wish to obtain from the program, and participates accordingly. The program is partially supported by NSF award DMS-2218374.

Prologue: Sumsets and Difference Sets

Let A be a finite set of nonnegative integers, with $|A|$ its size.

Definition

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

$$A + A = \{a_i + a_j : a_i, a_j \in A\},$$
$$A - A = \{a_i - a_j : a_i, a_j \in A\}.$$

Prologue: Sumsets and Difference Sets

Let A be a finite set of nonnegative integers, with $|A|$ its size.

Definition

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

$$A + A = \{a_i + a_j : a_i, a_j \in A\},$$
$$A - A = \{a_i - a_j : a_i, a_j \in A\}.$$

Expect a generic A has $|A - A| \geq |A + A|$, as addition commutes while subtraction doesn't.

Prologue: Sumsets and Difference Sets

Let A be a finite set of nonnegative integers, with $|A|$ its size.

Definition

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

$$A + A = \{a_i + a_j : a_i, a_j \in A\},$$
$$A - A = \{a_i - a_j : a_i, a_j \in A\}.$$

Expect a generic A has $|A - A| \geq |A + A|$, as addition commutes while subtraction doesn't.

Definition

Say A is **sum-dominated** if $|A + A| > |A - A|$, **difference-dominated** if $|A - A| > |A + A|$ and **balanced** otherwise.

Existence of Sum-Dominated Sets

Question

Do there exist sum-dominated sets?

Existence of Sum-Dominated Sets

Question

Do there exist sum-dominated sets?

Conway: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ is sum-dominated.

Existence of Sum-Dominated Sets

Question

Do there exist sum-dominated sets?

Conway: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ is sum-dominated.
Hegarty, Nathanson...: Infinite families of sum-dominated sets.

Existence of Sum-Dominated Sets

Question

Do there exist sum-dominated sets?

Conway: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ is sum-dominated.
Hegarty, Nathanson...: Infinite families of sum-dominated 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 Sum-Dominated Sets

Question

Do there exist sum-dominated sets?

Conway: $\{0, 2, 3, 4, 7, 11, 12, 14\}$ is sum-dominated.
Hegarty, Nathanson...: Infinite families of sum-dominated 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 Subset Model on $I_N := \{0, 1, \ldots, N-1\}$

Interpretation: Find a natural family of probability measures $\{\mathbb{P}_N\}_{N=1}^{\infty}$ on 2^{I_N} where

$$\lim_{N \to \infty} \mathbb{P}_N \left(\{A \subseteq I_N : A \text{ is MSTD}\} \right) = 0.$$

Uniform Subset Model on $I_N := \{0, 1, \ldots, N-1\}$

Interpretation: Find a natural family of probability measures $\{\mathbb{P}_N\}_{N=1}^{\infty}$ on 2^{I_N} where

$$\lim_{N\to\infty} \mathbb{P}_N\left(\{A \subseteq I_N : A \text{ is MSTD}\}\right) = 0.$$

First try: Pick A uniformly at random from 2^{I_N} (i.e., independently $k \in I_N$ in A with probability $1/2$).

Uniform Subset Model on $I_N := \{0, 1, \ldots, N-1\}$

Interpretation: Find a natural family of probability measures $\{\mathbb{P}_N\}_{N=1}^{\infty}$ on 2^{I_N} where

$$\lim_{N\to\infty} \mathbb{P}_N\left(\{A \subseteq I_N : A \text{ is MSTD}\}\right) = 0.$$

First try: Pick A uniformly at random from 2^{I_N}
(i.e., independently $k \in I_N$ in A with probability $1/2$).

Theorem (Martin and O'Bryant, 2006)

$$\lim_{N\to\infty} \frac{\#\{A \subseteq I_N : A \text{ is } MSTD\}}{2^{N+1}} > c_{1/2} > 0.$$

Uniform Subset Model on $I_N := \{0, 1, \ldots, N-1\}$

Interpretation: Find a natural family of probability measures $\{\mathbb{P}_N\}_{N=1}^{\infty}$ on 2^{I_N} where

$$\lim_{N\to\infty} \mathbb{P}_N \left(\{A \subseteq I_N : A \text{ is MSTD}\}\right) = 0.$$

First try: Pick A uniformly at random from 2^{I_N}
(i.e., independently $k \in I_N$ in A with probability $1/2$).

Theorem (Martin and O'Bryant, 2006)

$$\lim_{N\to\infty} \frac{\#\{A \subseteq I_N : A \text{ is MSTD}\}}{2^{N+1}} \; > \; c_{1/2} \; > \; 0.$$

Intuition:

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

Uniform Subset Model on $I_N := \{0, 1, \ldots, N-1\}$

Interpretation: Find a natural family of probability measures $\{\mathbb{P}_N\}_{N=1}^{\infty}$ on 2^{I_N} where

$$\lim_{N\to\infty} \mathbb{P}_N \left(\{A \subseteq I_N : A \text{ is MSTD}\}\right) = 0.$$

First try: Pick A uniformly at random from 2^{I_N}
(i.e., independently $k \in I_N$ in A with probability $1/2$).

Theorem (Martin and O'Bryant, 2006)

$$\lim_{N\to\infty} \frac{\#\{A \subseteq I_N : A \textit{ is MSTD}\}}{2^{N+1}} \; > \; c_{1/2} \; > \; 0.$$

Intuition:

- With high probability, $A + A$ and $A - A$ hits everything in the middle of $[0, 2N]$ and $[-N, N]$, respectively.
- "Rig" fringes of $A + A$ and $A - A$ by selecting the fringes of A.

Uniform Subset Model on $I_N := \{0, 1, \dots, N-1\}$

Interpretation: Find a natural family of probability measures $\{\mathbb{P}_N\}_{N=1}^{\infty}$ on 2^{I_N} where

$$\lim_{N\to\infty} \mathbb{P}_N\left(\{A \subseteq I_N : A \text{ is MSTD}\}\right) = 0.$$

First try: Pick A uniformly at random from 2^{I_N} (i.e., independently $k \in I_N$ in A with probability $1/2$).

Theorem (Martin and O'Bryant, 2006)

$$\lim_{N\to\infty} \frac{\#\{A \subseteq I_N : A \textit{ is MSTD}\}}{2^{N+1}} \; > \; c_{1/2} \; > \; 0.$$

Intuition:

- With high probability, $A+A$ and $A-A$ hits everything in the middle of $[0, 2N]$ and $[-N, N]$, respectively.
- "Rig" fringes of $A+A$ and $A-A$ by selecting the fringes of A.

Uniform subset model might not be the "right way of counting."

Binomial Model: Phase Transition

Second try: Similar, but the inclusion probability for taking k is $p(N) = o(1)$ and $N \cdot p(N) \ggg 1$. **Assume this from now on.**

Inclusion probability to 0, while $|A + A|, |A - A|$ grow.

Binomial Model: Phase Transition

Second try: Similar, but the inclusion probability for taking k is $p(N) = o(1)$ and $N \cdot p(N) \ggg 1$. **Assume this from now on.**

Inclusion probability to 0, while $|A + A|, |A - A|$ grow.

Conjecture (Martin and O'Bryant, 2006)

Under the binomial model, A is not MSTD with high probability.

Binomial Model: Phase Transition

Second try: Similar, but the inclusion probability for taking k is $p(N) = o(1)$ and $N \cdot p(N) \ggg 1$. **Assume this from now on.**

Inclusion probability to 0, while $|A+A|, |A-A|$ grow.

Conjecture (Martin and O'Bryant, 2006)

Under the binomial model, A is not MSTD with high probability.

Theorem (Hegarty and Miller, 2008)

Let $1/N \lll p(N) = 0(1)$.

- *If $p(N) = o(N^{-1/2})$: $|A-A| \sim 2|A+A| \sim (N \cdot p(N))^2$.*

Binomial Model: Phase Transition

Second try: Similar, but the inclusion probability for taking k is $p(N) = o(1)$ and $N \cdot p(N) \ggg 1$. **Assume this from now on.**

Inclusion probability to 0, while $|A+A|, |A-A|$ grow.

Conjecture (Martin and O'Bryant, 2006)

Under the binomial model, A is not MSTD with high probability.

Theorem (Hegarty and Miller, 2008)

Let $1/N \lll p(N) = 0(1)$.

- *If $p(N) = o(N^{-1/2})$: $|A-A| \sim 2|A+A| \sim (N \cdot p(N))^2$.*
- *If $p(N) = cN^{-1/2}$: $|A+A| \sim g(c^2/2)N$ and $|A-A| \sim g(c^2)N$ with $g(c) = 2(e^{-x} - 1 - x)/x$.*

Binomial Model: Phase Transition

Second try: Similar, but the inclusion probability for taking k is $p(N) = o(1)$ and $N \cdot p(N) \ggg 1$. **Assume this from now on.**

Inclusion probability to 0, while $|A+A|, |A-A|$ grow.

Conjecture (Martin and O'Bryant, 2006)

Under the binomial model, A is not MSTD with high probability.

Theorem (Hegarty and Miller, 2008)

Let $1/N \lll p(N) = 0(1)$.

- *If $p(N) = o(N^{-1/2})$: $|A-A| \sim 2|A+A| \sim (N \cdot p(N))^2$.*
- *If $p(N) = cN^{-1/2}$: $|A+A| \sim g(c^2/2)N$ and $|A-A| \sim g(c^2)N$ with $g(c) = 2(e^{-x} - 1 - x)/x$.*
- *If $p(N) \ggg N^{-1/2}$: $|(A+A)^c| \sim 2|(A-A)^c| \sim \frac{4}{p(N)^2}$.*

Extension: Binomial Sets Under Linear Forms

Beyond sums and differences: What about an arbitrary fixed linear combination of h elements?

Extension: Binomial Sets Under Linear Forms

Beyond sums and differences: What about an arbitrary fixed linear combination of h elements?

Definition

A *$\mathbb{Z}$-linear form* in h variables $L : \mathbb{Z}^h \to \mathbb{Z}$ is a map given via

$$L(x_1, \ldots, x_h) \;:=\; u_1x_1 + \cdots + u_hx_h, \quad u_i \in \mathbb{Z}_{\neq 0} \text{ for all } i \in [h].$$

Extension: Binomial Sets Under Linear Forms

Beyond sums and differences: What about an arbitrary fixed linear combination of h elements?

Definition

A *$\mathbb{Z}$-linear form* in h variables $L : \mathbb{Z}^h \to \mathbb{Z}$ is a map given via

$$L(x_1, \ldots, x_h) \ := \ u_1 x_1 + \cdots + u_h x_h, \quad u_i \in \mathbb{Z}_{\neq 0} \text{ for all } i \in [h].$$

- The set $L(A)$ is a *generalized sumset* if $|u_i| = 1$ for all $i \in [h]$. The $h = 2$ case corresponds to the Hegarty-Miller model.

Extension: Binomial Sets Under Linear Forms

Beyond sums and differences: What about an arbitrary fixed linear combination of h elements?

Definition

A *$\mathbb{Z}$-linear form* in h variables $L : \mathbb{Z}^h \to \mathbb{Z}$ is a map given via

$$L(x_1, \ldots, x_h) \ := \ u_1 x_1 + \cdots + u_h x_h, \quad u_i \in \mathbb{Z}_{\neq 0} \text{ for all } i \in [h].$$

- The set $L(A)$ is a *generalized sumset* if $|u_i| = 1$ for all $i \in [h]$. The $h = 2$ case corresponds to the Hegarty-Miller model.
- Natural setup to probe the additive structure of A.

Extension: Binomial Sets Under Linear Forms

Beyond sums and differences: What about an arbitrary fixed linear combination of h elements?

Definition

A *$\mathbb{Z}$-linear form* in h variables $L : \mathbb{Z}^h \to \mathbb{Z}$ is a map given via

$$L(x_1, \ldots, x_h) \;:=\; u_1 x_1 + \cdots + u_h x_h, \quad u_i \in \mathbb{Z}_{\neq 0} \text{ for all } i \in [h].$$

- The set $L(A)$ is a *generalized sumset* if $|u_i| = 1$ for all $i \in [h]$. The $h = 2$ case corresponds to the Hegarty-Miller model.
- Natural setup to probe the additive structure of A.

Core problem: Fix a $\mathbb{Z}$-linear form in h variables $L : \mathbb{Z}^h \to \mathbb{Z}$, and study

$$L(A) \;:=\; \{u_1 a_1 + \cdots + u_h a_h : a_i \in A\}$$

and how its asymptotic behavior depends on the randomness of A.

Extension: Binomial Sets Under Linear Forms

Beyond sums and differences: What about an arbitrary fixed linear combination of h elements?

Definition

A *$\mathbb{Z}$-linear form* in h variables $L : \mathbb{Z}^h \to \mathbb{Z}$ is a map given via

$$L(x_1, \ldots, x_h) \ := \ u_1 x_1 + \cdots + u_h x_h, \quad u_i \in \mathbb{Z}_{\neq 0} \text{ for all } i \in [h].$$

- The set $L(A)$ is a *generalized sumset* if $|u_i| = 1$ for all $i \in [h]$. The $h = 2$ case corresponds to the Hegarty-Miller model.
- Natural setup to probe the additive structure of A.

Core problem: Fix a $\mathbb{Z}$-linear form in h variables $L : \mathbb{Z}^h \to \mathbb{Z}$, and study

$$L(A) \ := \ \{u_1 a_1 + \cdots + u_h a_h : a_i \in A\}$$

and how its asymptotic behavior depends on the randomness of A.

- How sharp is the threshold, and what happens in the critical window?
- How does the macroscopic behavior depend on the structure of L?

Some Key Challenges

1. **No obvious candidates** for an extension of the Hegarty-Miller result to this setting, especially in the latter two regimes.

Some Key Challenges

1. **No obvious candidates** for an extension of the Hegarty-Miller result to this setting, especially in the latter two regimes.

   Can capture the amount of redundancy in L via the constant

$$\theta_L \ := \ \left|\left\{\sigma \in \mathfrak{S}_h : \left(u_{\sigma(1)}, \ldots, u_{\sigma(h)}\right) = (u_1, \ldots, u_h)\right\}\right|.$$

Some Key Challenges

1. **No obvious candidates** for an extension of the Hegarty-Miller result to this setting, especially in the latter two regimes.

   Can capture the amount of redundancy in L via the constant

   $$\theta_L \ := \ \left|\left\{\sigma \in \mathfrak{S}_h : (u_{\sigma(1)}, \ldots, u_{\sigma(h)}) = (u_1, \ldots, u_h)\right\}\right|.$$

   Expect to see that $\left|L(A)\right| \sim (Np)^h / \theta_L$ in the subcritical regime. What should we expect to see at or beyond the critical window?

Some Key Challenges

1. **No obvious candidates** for an extension of the Hegarty-Miller result to this setting, especially in the latter two regimes.

   Can capture the amount of redundancy in L via the constant

   $$\theta_L := \left|\{\sigma \in \mathfrak{S}_h : (u_{\sigma(1)}, \ldots, u_{\sigma(h)}) = (u_1, \ldots, u_h)\}\right|.$$

   Expect to see that $|L(A)| \sim (Np)^h/\theta_L$ in the subcritical regime. What should we expect to see at or beyond the critical window?

2. **Significant dependency** arises when computing exclusion probabilities for the $h \geq 3$ regime.

Some Key Challenges

1. **No obvious candidates** for an extension of the Hegarty-Miller result to this setting, especially in the latter two regimes.

   Can capture the amount of redundancy in L via the constant

   $$\theta_L \ := \ \left|\left\{\sigma \in \mathfrak{S}_h : \left(u_{\sigma(1)}, \ldots, u_{\sigma(h)}\right) = (u_1, \ldots, u_h)\right\}\right|.$$

   Expect to see that $\left|L(A)\right| \sim (Np)^h/\theta_L$ in the subcritical regime. What should we expect to see at or beyond the critical window?

2. **Significant dependency** arises when computing exclusion probabilities for the $h \geq 3$ regime.

   Simple illustrative example: $k \notin A + A$ 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.

Some Key Challenges

1. **No obvious candidates** for an extension of the Hegarty-Miller result to this setting, especially in the latter two regimes.

   Can capture the amount of redundancy in L via the constant

$$\theta_L := \left|\left\{\sigma \in \mathfrak{S}_h : (u_{\sigma(1)}, \dots, u_{\sigma(h)}) = (u_1, \dots, u_h)\right\}\right|.$$

   Expect to see that $\left|L(A)\right| \sim (Np)^h/\theta_L$ in the subcritical regime. What should we expect to see at or beyond the critical window?

2. **Significant dependency** arises when computing exclusion probabilities for the $h \geq 3$ regime.

   Simple illustrative example: $k \notin A + A$ is the union of the mutually independent events

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

   Harder to compute exclusion probabilities for three or more summands.

3. **Lower bounds:** We show a Poisson convergence result at a local scale. How to show some sort of separation from the weak limit?

Main Result 1: Global Phase Transition

Theorem (Jeong, Miller, 2023+)

(i) *(subcritical) If $p(N) \ll N^{-\frac{h-1}{h}}$, then*

$$|L(A)| \sim \frac{(N \cdot p(N))^h}{\theta_L}.$$

Main Result 1: Global Phase Transition

Theorem (Jeong, Miller, 2023+)

(i) *(subcritical) If* $p(N) \ll N^{-\frac{h-1}{h}}$, *then*

$$|L(A)| \sim \frac{(N \cdot p(N))^h}{\theta_L}.$$

(ii) *(critical) If* $p(N) = cN^{-\frac{h-1}{h}}$ *for fixed* $c > 0$, *then there exists a rational function* $R(x_0, \ldots, x_h)$ *and an increasing function* $g_{u_1,\ldots,u_h}$ *such that*

$$|L(A)| \sim g_{u_1,\ldots,u_h}(R(c, \ldots, x_h)) \cdot N.$$

Main Result 1: Global Phase Transition

Theorem (Jeong, Miller, 2023+)

(i) *(subcritical) If* $p(N) \ll N^{-\frac{h-1}{h}}$, *then*

$$|L(A)| \sim \frac{\left(N \cdot p(N)\right)^h}{\theta_L}.$$

(ii) *(critical) If* $p(N) = cN^{-\frac{h-1}{h}}$ *for fixed* $c > 0$, *then there exists a rational function* $R(x_0, \ldots, x_h)$ *and an increasing function* $g_{u_1,\ldots,u_h}$ *such that*

$$|L(A)| \sim g_{u_1,\ldots,u_h}\left(R(c, \ldots, x_h)\right) \cdot N.$$

(iii) *(supercritical) If* $p(N) \gg N^{-\frac{h-1}{h}}$, *then*

$$|L(A)^c| \sim \frac{2 \cdot \Gamma\left(\frac{1}{h-1}\right) \sqrt[h-1]{(h-1)! \cdot \theta_L \cdot \prod_{i=1}^{h} |u_i|}}{(h-1) \cdot p(N)^{\frac{h}{h-1}}}.$$

Main Result 1: Global Phase Transition

Theorem (Jeong, Miller, 2023+)

(i) *(subcritical) If* $p(N) \ll N^{-\frac{h-1}{h}}$, *then*

$$|L(A)| \sim \frac{(N \cdot p(N))^h}{\theta_L}.$$

(ii) *(critical) If* $p(N) = cN^{-\frac{h-1}{h}}$ *for fixed* $c > 0$, *then there exists a rational function* $R(x_0, \ldots, x_h)$ *and an increasing function* $g_{u_1,\ldots,u_h}$ *such that*

$$|L(A)| \sim g_{u_1,\ldots,u_h}\left(R(c, \ldots, x_h)\right) \cdot N.$$

(iii) *(supercritical) If* $p(N) \gg N^{-\frac{h-1}{h}}$, *then*

$$|L(A)^c| \sim \frac{2 \cdot \Gamma\left(\frac{1}{h-1}\right) \sqrt[h-1]{(h-1)! \cdot \theta_L \cdot \prod_{i=1}^{h} |u_i|}}{(h-1) \cdot p(N)^{\frac{h}{h-1}}}.$$

Upshot: Universal threshold for random additive combinatorial structures, pinpointing when random sets achieve additive completeness and when multiplicities proliferate.

Settles a conjecture of Hegarty and Miller (Random Structures & Algorithms, 2009).

Main Result 2: Local Phase Transition

Key Tool: Stein's method based on dependency graphs to estimate zero probabilities of the number of distinct ways that a value under $L(A)$ is achieved.

Main Result 2: Local Phase Transition

Key Tool: Stein's method based on dependency graphs to estimate zero probabilities of the number of distinct ways that a value under $L(A)$ is achieved.

Follow-Up Question: Say that L takes values in $[-d_L N, s_L N]$. Is there something more interesting to be said about weak Poisson convergence?

Main Result 2: Local Phase Transition

Key Tool: Stein's method based on dependency graphs to estimate zero probabilities of the number of distinct ways that a value under $L(A)$ is achieved.

Follow-Up Question: Say that L takes values in $[-d_L N, s_L N]$. Is there something more interesting to be said about weak Poisson convergence?

Theorem (Jeong, Miller, 2023+)

Fix a candidate value $k \in [-d_L N, s_L N]$ of $L(A)$. Let W_k denote the number of distinct ways to represent k in $L(A)$. Let $\mu_k = \mathbb{E}[W_k]$. Let $C > 0$ be a constant.

(i) *If $p(N) \ll N^{-\frac{h-2}{h-1}}$, then uniformly over $k \in [-d_L N, s_L N]$, $d_{TV}\left(\mathcal{L}(W_k), Pois(\mu_k)\right) \ll 1$.*

(ii) *If $p(N) \gtrsim N^{-\frac{h-2}{h-1}}$, then uniformly over $k \in \left[(-d_L + C)N, (s_L - C)N\right]$, $d_{TV}\left(\mathcal{L}(W_k), Pois(\mu_k)\right) = \Omega(1)$.*

Main Result 2: Local Phase Transition

Key Tool: Stein's method based on dependency graphs to estimate zero probabilities of the number of distinct ways that a value under $L(A)$ is achieved.

Follow-Up Question: Say that L takes values in $[-d_L N, s_L N]$. Is there something more interesting to be said about weak Poisson convergence?

Theorem (Jeong, Miller, 2023+)

Fix a candidate value $k \in [-d_L N, s_L N]$ of $L(A)$. Let W_k denote the number of distinct ways to represent k in $L(A)$. Let $\mu_k = \mathbb{E}[W_k]$. Let $C > 0$ be a constant.

(i) *If $p(N) \ll N^{-\frac{h-2}{h-1}}$, then uniformly over $k \in [-d_L N, s_L N]$, $d_{TV}\left(\mathcal{L}(W_k), Pois(\mu_k)\right) \ll 1$.*

(ii) *If $p(N) \gtrsim N^{-\frac{h-2}{h-1}}$, then uniformly over $k \in \left[(-d_L + C)N, (s_L - C)N\right]$, $d_{TV}\left(\mathcal{L}(W_k), Pois(\mu_k)\right) = \Omega(1)$.*

Upshot: This threshold $N^{-\frac{h-2}{h-1}}$ here dominates the previous threshold $N^{-\frac{h-1}{h}}$. Thus, "de-Poissonization" does not explain the global phase transition of $L(A)$.

Aside: In MSTD literature, number of representations of k captured by graphs on vertices $[N]$ and edges corresponding to representations. Can express the above theorem as a threshold result on a certain random h-regular hypergraph.

Main Ingredients in the Proof

1. **Asymptotic enumeration**
2. **Stein's method and Janson inequalities**
3. **Kim–Vu martingale machinery**

Main Ingredients in the Proof

1. **Asymptotic enumeration:** Derive asymptotic formulae for the number of distinct representations for a candidate value $k \in [-d_L N, s_L N]$ in $L(A)$.

   With some work, this can be reduced to computing the number of partitions of k into h parts, for which the generating function is well-known to be the Gaussian binomial coefficient $\binom{k+h}{h}_q$.

   Recent analytic result on partition enumeration from Stanley and Zanello gives uniform asymptotics on the coefficients of $\binom{k+h}{h}_q$ for **fixed** h (previous known results held for $h \nearrow \infty$).

2. **Stein's method and Janson inequalities**

3. **Kim–Vu martingale machinery**

Main Ingredients in the Proof

1. **Asymptotic enumeration**

2. **Stein's method and Janson inequalities:** The possible ways in which a candidate sum $-d_L N + k$ may be generated in $L(A)$ correspond to a collection of rare and weakly dependent events.

   Critical window: Use the Stein-Chen method based on dependency graphs to get sharp estimates for zero probabilities for the number of representations in $L(A)$.

   Supercritical regime: Use Janson's inequalities to get vanishing bounds on zero probabilities for the number of representations in $L(A)$. Aggregate these approximations to argue that most everything is hit.

3. **Kim–Vu martingale machinery**

Main Ingredients in the Proof

1. **Asymptotic enumeration**
2. **Stein's method and Janson inequalities**
3. **Kim–Vu martingale machinery:** The second moment method fails to guarantee strong concentration bounds.

   We instead use Kim–Vu martingales to promote moment estimates to high-probability results.

Future Work

1. **Generalizations**
   - **Integer lattice:** Can we extend this to the integer lattice $\mathbb{Z}^d$ on d dimensions, or to other lattices?
   - **Other groups:** What if we change the ambient group that we work on?
   - **Nonlinearity:** What happens if L is a higher-degree polynomial (e.g., $L(x_1, x_2) = x_1^2 + x_2$)? Our work settles the degree-1 regime. Do phase transitions persist?
   - **Additional dependency:** What changes if we introduce dependencies into the binomial model itself, rather than modifying the context under which it operates?

Future Work

1. **Generalizations**
   - **Integer lattice:** Can we extend this to the integer lattice $\mathbb{Z}^d$ on d dimensions, or to other lattices?
   - **Other groups:** What if we change the ambient group that we work on?
   - **Nonlinearity:** What happens if L is a higher-degree polynomial (e.g., $L(x_1, x_2) = x_1^2 + x_2$)? Our work settles the degree-1 regime. Do phase transitions persist?
   - **Additional dependency:** What changes if we introduce dependencies into the binomial model itself, rather than modifying the context under which it operates?
2. **Extending weak convergence results**
   - **Joint Poisson limit:** Can one describe the joint distribution of $L(A)$ candidate values, not just marginals?
   - **Point processes:** Is there a local point process limit of some sort?

Acknowledgments

This research was supported, in part, by NSF Grant DMS1947438, the University of Pennsylvania and Williams College.

We also thank Professors Eyvi Palsson and Robin Pemantle for productive conversations.