The Helly number of Hamming balls and related problems

Zhihan Jin

ETH Zürich

Joint work with Noga Alon and Benny Sudakov

Summit280, July 11, 2024

Helly's theorem

Theorem (Helly 1923)

Let $K_1, K_2, \ldots, K_n$ be convex sets in $\mathbb{R}^d$. If every $d+1$ of them intersect, then all the sets intersect.

Helly's theorem

Theorem (Helly 1923)

Let $K_1, K_2, \ldots, K_n$ be convex sets in $\mathbb{R}^d$. If every $d+1$ of them intersect, then all the sets intersect.

- $d+1$ is tight: consider n hyperplanes in general position.

Helly number

- $\mathcal{F}$ is a fixed family of sets.

Question

Is there $h < \infty$ such that the following holds?
Let $K_1, K_2, \ldots, K_n \in \mathcal{F}$ be arbitrary. If every h of them intersect, then all the sets intersect.

Helly number

- $\mathcal{F}$ is a fixed family of sets.

Question

Is there $h < \infty$ such that the following holds?
Let $K_1, K_2, \ldots, K_n \in \mathcal{F}$ be arbitrary. If every h of them intersect, then all the sets intersect.

- The Helly number of family $\mathcal{F}$ is the minimum such h.

Examples of Helly numbers

- Let $\mathcal{F}$ be the family of all subtrees of a fixed tree T.
  The Helly number is 2.

Examples of Helly numbers

- Let $\mathcal{F}$ be the family of all subtrees of a fixed tree T.
  The Helly number is 2.
- Let $\mathcal{F}$ be the family of $K \cap \mathbb{Z}^d$ for all convex $K \subseteq \mathbb{R}^d$.
  Doignon showed its Helly number is 2^d.

Examples of Helly numbers

- Let $\mathcal{F}$ be the family of all subtrees of a fixed tree T.
  The Helly number is 2.
- Let $\mathcal{F}$ be the family of $K \cap \mathbb{Z}^d$ for all convex $K \subseteq \mathbb{R}^d$.
  Doignon showed its Helly number is 2^d.
- Let $\mathcal{F}$ be a family of subsets in $\mathbb{R}^d$ such that any subfamily-intersection is a union of k convex sets in $\mathbb{R}^d$.
  The Helly number is $O_{k,d}(1)$ by Alon-Kalai and Matoušek.

Hamming balls

- What about discrete objects?

Hamming balls

- What about discrete objects?
- The *Hamming distance* $\mathrm{dist}(a, b) := |\{k \in [n] : a_k \neq b_k\}|$.

Hamming balls

- What about discrete objects?
- The *Hamming distance* $\mathrm{dist}(a, b) := |\{k \in [n] : a_k \neq b_k\}|$.
- The *Hamming ball* centered at $p \in \mathbb{R}^n$ of radius t is $B(p, t) := \{q \in \mathbb{R}^n : \mathrm{dist}(p, q) \leq t\}$.

Hamming balls

- What about discrete objects?
- The *Hamming distance* $\mathrm{dist}(a, b) := |\{k \in [n] : a_k \neq b_k\}|$.
- The *Hamming ball* centered at $p \in \mathbb{R}^n$ of radius t is $B(p, t) := \{q \in \mathbb{R}^n : \mathrm{dist}(p, q) \leq t\}$.

Question

Determine the Helly number for Hamming balls in $\mathbb{R}^n$ of radius t.

Hamming balls

- What about discrete objects?
- The *Hamming distance* $\text{dist}(a, b) := |\{k \in [n] : a_k \neq b_k\}|$.
- The *Hamming ball* centered at $p \in \mathbb{R}^n$ of radius t is $B(p, t) := \{q \in \mathbb{R}^n : \text{dist}(p, q) \leq t\}$.

Question

Determine the Helly number for Hamming balls in $\mathbb{R}^n$ of radius t.

The case when $\mathbb{R}^n$ is replaced by $\{0, 1\}^n$ was raised by Raman, Subedi and Tewari in the study of online learning.

Main result

An example when $n = t + 1$:

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0, 1\}^{t+1}$.

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0, 1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0,1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.
- $\{0,1\}^{t+1} \setminus B(p,t) = \{\overline{p}\}$

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0,1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.
- $\{0,1\}^{t+1} \setminus B(p,t) = \{\overline{p}\}$ ········ $\overline{p}$ is the opposite point of p.

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0,1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.
- $\{0,1\}^{t+1} \setminus B(p,t) = \{\overline{p}\}$ ········· $\overline{p}$ is the opposite point of p.
- $2^{t+1} - 1$ of them intersect;

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0,1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.
- $\{0,1\}^{t+1} \setminus B(p,t) = \{\overline{p}\}$ .... $\overline{p}$ is the opposite point of p.
- $2^{t+1} - 1$ of them intersect; all of them do not intersect.

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0,1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.
- $\{0,1\}^{t+1} \setminus B(p,t) = \{\overline{p}\}$ ········· $\overline{p}$ is the opposite point of p.
- $2^{t+1} - 1$ of them intersect; all of them do not intersect.

So, the Helly number is at least 2^{t+1}.

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0,1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.
- $\{0,1\}^{t+1} \setminus B(p,t) = \{\overline{p}\}$ ········· $\overline{p}$ is the opposite point of p.
- $2^{t+1} - 1$ of them intersect; all of them do not intersect.

So, the Helly number is at least 2^{t+1}. Is it tight?

Main result

An example when $n = t + 1$:

- Consider all Hamming balls centered at $p \in \{0,1\}^{t+1}$.
- There are 2^{t+1} Hamming balls of radius t.
- $\{0,1\}^{t+1} \setminus B(p,t) = \{\overline{p}\}$ ········· $\overline{p}$ is the opposite point of p.
- $2^{t+1} - 1$ of them intersect; all of them do not intersect.

So, the Helly number is at least 2^{t+1}. Is it tight?

Theorem (Alon, J. and Sudakov 2024+)

Let $n \geq t + 1$. The Helly number for the family of Hamming balls of radius t in $\mathbb{R}^n$ is 2^{t+1}.

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.
- There exist Hamming balls $B_1, \ldots, B_h$:

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.
- There exist Hamming balls $B_1, \ldots, B_h$:
  - every $h-1$ of them intersect; all of them do not intersect.

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.
- There exist Hamming balls $B_1, \ldots, B_h$:
  - every $h-1$ of them intersect; all of them do not intersect.
- Let a_i be the center of B_i,

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.
- There exist Hamming balls $B_1, \ldots, B_h$:
  - every $h-1$ of them intersect; all of them do not intersect.
- Let a_i be the center of B_i,
  and b_i be in the intersection of all B_js, $j \neq i$.

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.
- There exist Hamming balls $B_1, \ldots, B_h$:
  - every $h-1$ of them intersect; all of them do not intersect.
- Let a_i be the center of B_i,
  and b_i be in the intersection of all B_js, $j \neq i$.
  - $\text{dist}(a_i, b_i) \geq t+1$; $\text{dist}(a_i, b_j) \leq t$ for $i \neq j$.

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.
- There exist Hamming balls $B_1, \ldots, B_h$:
  - every $h-1$ of them intersect; all of them do not intersect.
- Let a_i be the center of B_i,
  and b_i be in the intersection of all B_js, $j \neq i$.
  - $\mathrm{dist}(a_i, b_i) \geq t+1$; $\mathrm{dist}(a_i, b_j) \leq t$ for $i \neq j$.

Theorem (Alon, J. and Sudakov 2024+)

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t+1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

Reduction to a set-pair inequality

- Suppose the Helly number is at least h for some $h \geq 0$.
- There exist Hamming balls $B_1, \ldots, B_h$:
  - every $h-1$ of them intersect; all of them do not intersect.
- Let a_i be the center of B_i,
  and b_i be in the intersection of all B_js, $j \neq i$.
  - $\mathrm{dist}(a_i, b_i) \geq t+1$; $\mathrm{dist}(a_i, b_j) \leq t$ for $i \neq j$.

Theorem (Alon, J. and Sudakov 2024+)

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t+1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

The Helly number for Hamming balls of radius t is precisely 2^{t+1}!

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) \geq t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \neq j,$$

then $m \le 2^{t+1}$.

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \textit{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \textit{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

- We will prove this by the algebraic method.

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

- We will prove this by the algebraic method.
- The dimension argument:

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \neq j,$$

then $m \le 2^{t+1}$.

- We will prove this by the algebraic method.
- The dimension argument:
  - construct m objects (say polynomials);

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \neq j,$$

then $m \le 2^{t+1}$.

- We will prove this by the algebraic method.
- The dimension argument:
  - construct m objects (say polynomials);
  - show they are linearly independent;

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \textit{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \textit{ for } i \neq j,$$

then $m \le 2^{t+1}$.

- We will prove this by the algebraic method.
- The dimension argument:
  - construct m objects (say polynomials);
  - show they are linearly independent;
  - derive that $m \le \dim(\text{ambient space})$.

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \neq j,$$

then $m \le 2^{t+1}$.

- We will prove this by the algebraic method.
- The dimension argument:
  - construct m objects (say polynomials);
  - show they are linearly independent;
  - derive that $m \le \dim(\text{ambient space})$.
- Note: dim(ambient space) usually depends on n!

Proof of a simplified version

A simplified version:

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \ne j,$$

then $m \le 2^{t+1}$.

- We will prove this by the algebraic method.
- The dimension argument:
  - construct m objects (say polynomials);
  - show they are linearly independent;
  - derive that $m \le \dim(\text{ambient space})$.
- Note: dim(ambient space) usually depends on n!
- *Our strategy*: construct more objects for each $i \in [m]$.

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\operatorname{dist}(a_i, b_i) = t + 1 \textit{ for all } i, \quad \operatorname{dist}(a_i, b_j) \leq t \textit{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

- $D_i := \{k \in [n] : a_{i,k} \neq b_{i,k}\}$;

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \ne j,$$

then $m \le 2^{t+1}$.

- $D_i := \{k \in [n] : a_{i,k} \ne b_{i,k}\}$; $|D_i| = t + 1$.

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \neq j,$$

then $m \le 2^{t+1}$.

- $D_i := \{k \in [n] : a_{i,k} \neq b_{i,k}\}$; $|D_i| = t + 1$.
- $\forall i \in [m]\ \forall S \subseteq [n] \setminus D_i$, define a multilinear polynomial

$$f_{i,S}(x) := \prod_{k \in S \cup D_i} (x_k - a_{i,k}).$$

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\operatorname{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \operatorname{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

- $D_i := \{k \in [n] : a_{i,k} \neq b_{i,k}\}$; $|D_i| = t + 1$.
- $\forall i \in [m]\ \forall S \subseteq [n] \setminus D_i$, define a multilinear polynomial

$$f_{i,S}(x) := \prod_{k \in S \cup D_i} (x_k - a_{i,k}).$$

- Claim: $f_{i,S}$s are linearly independent.

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \le t \text{ for } i \neq j,$$

then $m \le 2^{t+1}$.

- $D_i := \{k \in [n] : a_{i,k} \neq b_{i,k}\}$; $|D_i| = t + 1$.
- $\forall i \in [m] \ \forall S \subseteq [n] \setminus D_i$, define a multilinear polynomial

$$f_{i,S}(x) := \prod_{k \in S \cup D_i} (x_k - a_{i,k}).$$

- Claim: $f_{i,S}$s are linearly independent.
  - Then, $m \cdot 2^{n-t-1} \le 2^n$

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) = t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

- $D_i := \{k \in [n] : a_{i,k} \neq b_{i,k}\}$; $|D_i| = t + 1$.
- $\forall i \in [m]\ \forall S \subseteq [n] \setminus D_i$, define a multilinear polynomial

$$f_{i,S}(x) := \prod_{k \in S \cup D_i} (x_k - a_{i,k}).$$

- Claim: $f_{i,S}$s are linearly independent.
  - Then, $m \cdot 2^{n-t-1} \leq 2^n \implies m \leq 2^{t+1}$.

Proof of a simplified version

Theorem

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) = t + 1 \textit{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \textit{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

- $D_i := \{k \in [n] : a_{i,k} \neq b_{i,k}\}$; $|D_i| = t + 1$.
- $\forall i \in [m]\ \forall S \subseteq [n] \setminus D_i$, define a multilinear polynomial

$$f_{i,S}(x) := \prod_{k \in S \cup D_i} (x_k - a_{i,k}).$$

- Claim: $f_{i,S}$s are linearly independent.
  - Then, $m \cdot 2^{n-t-1} \leq 2^n \implies m \leq 2^{t+1}$. *Nice!*

$f_{i,S}$s are linearly independent

<u>Recall</u>:

$D_i = \{k : a_{i,k} \neq b_{i,k}\}$, $f_{i,S}(x) = \prod_{k \in S \cup D_i} (x_k - a_{i,k}) \; \forall S \subseteq [n] \setminus D_i$.

$f_{i,S}$s are linearly independent

Recall:

$D_i = \{k : a_{i,k} \neq b_{i,k}\}$, $f_{i,S}(x) = \prod_{k \in S \cup D_i}(x_k - a_{i,k})\ \forall S \subseteq [n] \setminus D_i$.

Evaluation points:

$x_{i,S} \in \mathbb{R}^n : x_{i,S,k} = b_{i,k}\ \forall k \notin S, \quad x_{i,S,k} \in \mathbb{R} \setminus \{b_{i,k}\}\ \forall k \in S.$

$f_{i,S}$s are linearly independent

Recall:

$D_i = \{k : a_{i,k} \neq b_{i,k}\}$, $f_{i,S}(x) = \prod_{k \in S \cup D_i} (x_k - a_{i,k})$ $\forall S \subseteq [n] \setminus D_i$.

Evaluation points:

$x_{i,S} \in \mathbb{R}^n : x_{i,S,k} = b_{i,k}$ $\forall k \notin S$, $\quad x_{i,S,k} \in \mathbb{R} \setminus \{b_{i,k}\}$ $\forall k \in S$.

- $f_{i,S}(x_{i,S}) \neq 0$.

$f_{i,S}$s are linearly independent

Recall:

$D_i = \{k : a_{i,k} \neq b_{i,k}\}$, $f_{i,S}(x) = \prod_{k \in S \cup D_i} (x_k - a_{i,k})$ $\forall S \subseteq [n] \setminus D_i$.

Evaluation points:

$x_{i,S} \in \mathbb{R}^n : x_{i,S,k} = b_{i,k}$ $\forall k \notin S$, $\quad x_{i,S,k} \in \mathbb{R} \setminus \{b_{i,k}\}$ $\forall k \in S$.

- $f_{i,S}(x_{i,S}) \neq 0$.
- $f_{i',S'}(x_{i,S}) = 0$ if $(i, S) \neq (i', S')$ and $|S| \leq |S'|$.

$f_{i,S}$s are linearly independent

Recall:

$D_i = \{k : a_{i,k} \neq b_{i,k}\}$, $f_{i,S}(x) = \prod_{k \in S \cup D_i}(x_k - a_{i,k})$ $\forall S \subseteq [n] \setminus D_i$.

Evaluation points:

$x_{i,S} \in \mathbb{R}^n : x_{i,S,k} = b_{i,k}$ $\forall k \notin S$, $\quad x_{i,S,k} \in \mathbb{R} \setminus \{b_{i,k}\}$ $\forall k \in S$.

- $f_{i,S}(x_{i,S}) \neq 0$.
- $f_{i',S'}(x_{i,S}) = 0$ if $(i, S) \neq (i', S')$ and $|S| \leq |S'|$.
- Upper triangular in the order of $|S|$.

$f_{i,S}$s are linearly independent

Recall:

$D_i = \{k : a_{i,k} \neq b_{i,k}\}$, $f_{i,S}(x) = \prod_{k \in S \cup D_i}(x_k - a_{i,k})\ \forall S \subseteq [n] \setminus D_i$.

Evaluation points:

$x_{i,S} \in \mathbb{R}^n : x_{i,S,k} = b_{i,k}\ \forall k \notin S, \quad x_{i,S,k} \in \mathbb{R} \setminus \{b_{i,k}\}\ \forall k \in S.$

- $f_{i,S}(x_{i,S}) \neq 0$.
- $f_{i',S'}(x_{i,S}) = 0$ if $(i, S) \neq (i', S')$ and $|S| \leq |S'|$.
- Upper triangular in the order of $|S|$.
- $f_{i,S}$**s are linearly independent!**

Binary error correcting codes

Definition

Let $n \geq d \geq 0$. A *binary error correcting code(ECC)* is a subset of $\{0,1\}^n$ s.t. the pairwise Hamming distance is at least d.

Binary error correcting codes

Definition

Let $n \geq d \geq 0$. A *binary error correcting code(ECC)* is a subset of $\{0,1\}^n$ s.t. the pairwise Hamming distance is at least d.

Let $A_{n,d}$ be the size of the largest ECC.

Binary error correcting codes

Definition

Let $n \geq d \geq 0$. A *binary error correcting code(ECC)* is a subset of $\{0,1\}^n$ s.t. the pairwise Hamming distance is at least d.

Let $A_{n,d}$ be the size of the largest ECC.

Theorem (Hamming bound)

$A_{n,d} \leq H_{n,d} := \frac{2^n}{\sum_{i=0}^{(d-1)/2} \binom{n}{i}}$.

Binary error correcting codes

Definition

Let $n \geq d \geq 0$. A *binary error correcting code(ECC)* is a subset of $\{0,1\}^n$ s.t. the pairwise Hamming distance is at least d.

Let $A_{n,d}$ be the size of the largest ECC.

Theorem (Hamming bound)

$A_{n,d} \leq H_{n,d} := \frac{2^n}{\sum_{i=0}^{(d-1)/2} \binom{n}{i}}$.

- Tightness: trivial codes, Hamming codes and Golay codes.

Binary error correcting codes

Definition

Let $n \geq d \geq 0$. A *binary error correcting code(ECC)* is a subset of $\{0,1\}^n$ s.t. the pairwise Hamming distance is at least d.

Let $A_{n,d}$ be the size of the largest ECC.

Theorem (Hamming bound)

$A_{n,d} \leq H_{n,d} := \frac{2^n}{\sum_{i=0}^{(d-1)/2} \binom{n}{i}}$.

- Tightness: trivial codes, Hamming codes and Golay codes.
- Approximate tightness: BCH codes when d is a constant.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\operatorname{dist}(a_i, b_i) \geq t + 1 \textit{ for all } i, \quad \operatorname{dist}(a_i, b_j) \leq t \textit{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ *and* $s \geq 1$. *If* $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\text{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \text{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:
  - Let $n = t + s$, $(a_i)_i$ be an ECC in $\{0,1\}^{t+s}$ with distance s. Take $b_i := \overline{a_i}$ $\forall i$.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:
  - Let $n = t + s$, $(a_i)_i$ be an ECC in $\{0,1\}^{t+s}$ with distance s. Take $b_i := \overline{a_i}\ \forall i$.
  - $\mathrm{dist}(a_i, b_i) = t + s$, $\mathrm{dist}(a_i, b_j) = t + s - \mathrm{dist}(a_i, a_j) \leq t\ \forall i \neq j$.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:
  - Let $n = t + s$, $(a_i)_i$ be an ECC in $\{0,1\}^{t+s}$ with distance s. Take $b_i := \overline{a_i}\ \forall i$.
  - $\mathrm{dist}(a_i, b_i) = t + s$, $\mathrm{dist}(a_i, b_j) = t + s - \mathrm{dist}(a_i, a_j) \leq t\ \forall i \neq j$.
- Our theorem can be seen as a generalized Hamming bound.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:
  - Let $n = t + s$, $(a_i)_i$ be an ECC in $\{0,1\}^{t+s}$ with distance s. Take $b_i := \overline{a_i}$ $\forall i$.
  - $\mathrm{dist}(a_i, b_i) = t + s$, $\mathrm{dist}(a_i, b_j) = t + s - \mathrm{dist}(a_i, a_j) \leq t$ $\forall i \neq j$.
- Our theorem can be seen as a generalized Hamming bound.
- Tight due to **perfect codes**:

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:
  - Let $n = t + s$, $(a_i)_i$ be an ECC in $\{0,1\}^{t+s}$ with distance s. Take $b_i := \overline{a_i}\ \forall i$.
  - $\mathrm{dist}(a_i, b_i) = t + s$, $\mathrm{dist}(a_i, b_j) = t + s - \mathrm{dist}(a_i, a_j) \leq t\ \forall i \neq j$.
- Our theorem can be seen as a generalized Hamming bound.
- Tight due to **perfect codes**:
  - trivial codes, Hamming codes, Golay codes.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:
  - Let $n = t + s$, $(a_i)_i$ be an ECC in $\{0,1\}^{t+s}$ with distance s. Take $b_i := \overline{a_i}$ $\forall i$.
  - $\mathrm{dist}(a_i, b_i) = t + s$, $\mathrm{dist}(a_i, b_j) = t + s - \mathrm{dist}(a_i, a_j) \leq t$ $\forall i \neq j$.
- Our theorem can be seen as a generalized Hamming bound.
- Tight due to **perfect codes**:
  - trivial codes, Hamming codes, Golay codes.
  - When $s = 1, 2, 3, 4$, or $s \in \{7, 8\}$ and $t = 16$.

A more general set-pair inequality

Theorem (Alon, J. and Sudakov 2024+)

Let $n > t \geq 0$ and $s \geq 1$. If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t + s \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq H_{t+s,s}$.

- This setting generalizes error correcting codes:
  - Let $n = t + s$, $(a_i)_i$ be an ECC in $\{0,1\}^{t+s}$ with distance s. Take $b_i := \overline{a_i}$ $\forall i$.
  - $\mathrm{dist}(a_i, b_i) = t + s$, $\mathrm{dist}(a_i, b_j) = t + s - \mathrm{dist}(a_i, a_j) \leq t$ $\forall i \neq j$.
- Our theorem can be seen as a generalized Hamming bound.
- Tight due to **perfect codes**:
  - trivial codes, Hamming codes, Golay codes.
  - When $s = 1, 2, 3, 4$, or $s \in \{7, 8\}$ and $t = 16$.
- Tight up to $O_s(1)$ due to BCH codes.

Further directions

Theorem (Alon, J. and Sudakov 2024+)

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) \geq t + 1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

Further directions

Theorem (Alon, J. and Sudakov 2024+)

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ *satisfies*

$$\mathrm{dist}(a_i, b_i) \geq t+1 \textit{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \textit{ for } i \neq j,$$

then $m \leq 2^{t+1}$.

The skew version:

Question

If $a_1, \ldots, a_m, b_1, \ldots, b_m \in \mathbb{R}^n$ satisfies

$$\mathrm{dist}(a_i, b_i) \geq t+1 \text{ for all } i, \quad \mathrm{dist}(a_i, b_j) \leq t \text{ for } 1 \leq i < j \leq m,$$

then $m \leq ???$

The End

Questions? Comments?

The  End

Questions? Comments?