

Triangle-free subsets of the r -distance graph of the Hypercube

Padmini Mukkamala*

Ananthkrishnan Ravi†

Abstract

Given the r -distance graph on the hypercube \mathbb{F}_2^n , where two vertices are adjacent if their Hamming distance is exactly r , we study the maximum size $T(n, r)$ of a triangle-free set of vertices. For even $r \leq n/2$, we prove

$$T(n, r) = O\left(\frac{r2^n}{n+1}\right).$$

We also obtain lower bounds in various regimes of r as a function of n .

1 Introduction

We work in the setting of the hypercube \mathbb{F}_2^n . For two vectors $x, y \in \mathbb{F}_2^n$, the Hamming distance between x and y is the number of coordinates at which they differ. We denote this by $d_H(x, y)$. Fix an integer r with $1 \leq r \leq n$. The r -distance graph $H(n, r)$ has vertex set \mathbb{F}_2^n , and we join x and y by an edge when $d_H(x, y) = r$. A *triangle* in $H(n, r)$ is a triple $\{x, y, z\}$ with

$$d_H(x, y) = d_H(x, z) = d_H(y, z) = r.$$

Our goal is to understand how large a subset of \mathbb{F}_2^n can be that does not contain triangles. Let $T(n, r)$ be the maximum size of a triangle-free set of vertices in $H(n, r)$.

Two constraints appear immediately. Triangles can occur only when r is even, and only when $3r \leq 2n$. For any triple $x, y, z \in \mathbb{F}_2^n$, the sum $d_H(x, y) + d_H(x, z) + d_H(y, z)$ is always even, since each coordinate contributes either 0 (all bits equal) or 2 (exactly one bit differs). The same coordinate-wise count also gives $d_H(x, y) + d_H(x, z) + d_H(y, z) \leq 2n$, so for a triangle this sum equals $3r$ and hence $3r \leq 2n$. So we focus on even r with $r \leq 2n/3$.

Before we move to the general setting, we look at two small cases. For $n = 3$ and $r = 2$, the set $\{000, 001, 010, 011\}$ is a maximum triangle-free set in $H(3, 2)$. For $n = 4$ and $r = 2$, the set

$$\{0000, 0001, 0101, 0111, 1000, 1010, 1110, 1111\}$$

is a maximum triangle-free set in $H(4, 2)$. These examples highlight that triangle-free sets in $H(n, r)$ need not be independent: they may contain many edges of length r while still avoiding triangles.

The problem may be viewed as a higher-order exact-distance problem on the hypercube. If we forbid all pairs at distance r , we obtain the independence number of the graph $H(n, r)$, which is the classical forbidden-distance problem; in the linear regime this is the setting of the Frankl–Rödl

*BME Budapest Institute of Technology and Economics, Budapest, Hungary. Supported by staff mobility funded by ERASMUS+ EU fund

†Delft University of Technology, Netherlands. Supported by an NWO open competition grant (OCENW.M.22.090).
E-mail: a.ravi@tudelft.nl.

graph [5, 8]. In coding-theoretic language, this asks for the largest binary code in which the distance r never occurs between two distinct codewords [3]. Questions of this type are part of the broader study of distance constraints in the hypercube; see, for example, [9]. Here we allow pairs at distance r , but forbid triples of vertices that are pairwise at distance r . Equivalently, if $\mathcal{H}(n, r)$ denotes the 3-uniform hypergraph on \mathbb{F}_2^n whose hyperedges are the triangles of $H(n, r)$, then $T(n, r)$ is the size of the largest independent set in $\mathcal{H}(n, r)$. In other words,

$$T(n, r) = \alpha(\mathcal{H}(n, r)),$$

where $\alpha(\mathcal{H})$ denotes the maximum size of a vertex set containing no hyperedge of \mathcal{H} .

This hypergraph viewpoint is also useful for upper bounds. Lovász [11] introduced the theta number as an efficiently computable upper bound for the independence number of a graph, and Grötschel, Lovász, and Schrijver [10] developed the related notion of the theta body. More recently, Castro-Silva, de Oliveira Filho, Slot, and Vallentin [13] introduced a recursive theta body for hypergraphs and applied it to this setting, obtaining an upper bound for $T(n, r)$. Their results suggest that in the linear regime $r \sim n/c$, the behavior should change around the threshold $r = n/2$: for $c > 2$ the normalized upper bound is conjectured to decay exponentially fast, while at $c = 2$ only linear decay is expected [13, Conjecture 7.3].

After the first version of this paper [12] appeared on ArXiv, very recently, Balogh, Chen, and Li [1] settled the constant regime. They proved that for every fixed even $r = 2d$,

$$T(n, 2d) = \Theta\left(\frac{2^n}{n^d}\right).$$

Thus the main novelty of the present paper lies in the non-constant regimes for r , together with the upper bounds in the range $r \leq n/2$.

Since triangle-free property is a weaker condition than independence, we can exploit constructions that allow many edges of length r while still avoiding triangles. This is reflected in our lower bounds, which come from three different mechanisms. The first mechanism constructs independent sets by forbidding the distance r , using forbidden-distance codes. The second keeps a large set and then deletes a small number of vertices to destroy every triangle, via the method of alteration. The third restricts to a structured slice of the hypercube and enforces a local obstruction that prevents triangles from occurring.

We combine these three ideas to obtain the following lower bound.

Theorem 1. *Let r be even and satisfy $r \leq 2n/3$. Then $T(n, r)$ satisfies the following lower bounds.*

1. *Constant $r = 2d$. For every fixed integer $d \geq 1$, there is a constant $c_d > 0$ such that*

$$T(n, 2d) \geq c_d \frac{2^n}{n^d} \quad (n \rightarrow \infty).$$

2. *If $r = o(n)$ with $r \rightarrow \infty$. We have*

$$T(n, r) \geq 2^{n - \frac{3}{4}(r \log(n/r)) + O(r)}.$$

We further divide the case when $r = \alpha n$ into two intervals, one where $\alpha \in (0, \alpha_0)$ and the other when $\alpha \in (\alpha_0, 2/3)$. Numerically, $\alpha_0 \approx 0.162$.

3. If $r = \alpha n$ with fixed $0 < \alpha < \alpha_0$, then

$$T(n, \alpha n) \geq 2^{c(\alpha)n + o(n)},$$

$$\text{where } c(\alpha) = 1 - \frac{1}{2} \left(H_2(\alpha) + \alpha + (1 - \alpha)H_2\left(\frac{\alpha}{2(1-\alpha)}\right) \right).$$

4. If $r = \alpha n$ with fixed $\alpha_0 \leq \alpha \leq 2/3$, then

$$T(n, \alpha n) \geq 2^{(H_2(\alpha/2) - o(1))n},$$

where $H_2(t) = -t \log_2 t - (1 - t) \log_2(1 - t)$ is the binary entropy function.

(2)–(3) are asymptotic corollaries of the alteration bound proved in Lemma 2. Here, $\alpha_0 \in (0, 2/3)$ is the threshold at which the exponents of Lemma 2 and Lemma 3 agree.

We also note that when $r = \frac{2n}{3}$, the cube \mathbb{F}_2^{n-1} does not contain any triangles, so there is a triangle-free subset of size $2^n/2$. This construction extends near the endpoint $r = 2n/3$. More generally, if $r = \frac{2}{3}(n - c)$ for some integer $c \geq 0$, then any cube of codimension $c + 1$ is triangle-free, since its dimension is $n - c - 1$ and $3r = 2(n - c) > 2(n - c - 1)$. Hence, $T(n, r) \geq 2^{n-c-1}$.

In particular, when $c = o(\log n)$ this gives

$$T(n, r) \geq 2^{n - o(\log n)} = \frac{2^n}{n^{o(1)}},$$

which is stronger than the lower bounds coming from our other constructions in this regime.

For upper bounds, we again use the hypergraph viewpoint. Let $\mathcal{H}(n, r)$ be the 3-uniform hypergraph with vertex set \mathbb{F}_2^n , whose hyperedges are the triangles of $H(n, r)$. Then a set $A \subseteq \mathbb{F}_2^n$ is triangle-free in $H(n, r)$ if and only if it is an independent set in $\mathcal{H}(n, r)$.

Our first upper bound is a simple uniform density bound. The second uses the layered structure of the cube. On each level, we project the vertices to lower levels by taking suitable shadows. Triangle-freeness then forces these projected families to satisfy a matching-type restriction.

We then apply a theorem of Frankl on families with no three pairwise disjoint sets [7] to bound the contribution from each level.

Theorem 2. *Let r be even. Then $T(n, r)$ satisfies the following upper bounds.*

1. If $r \leq 2n/3$, then

$$T(n, r) \leq \frac{2}{3} 2^n.$$

2. If $r \leq n/2$, then

$$T(n, r) = O\left(\frac{r2^n}{n+1}\right).$$

The rest of the paper is organized as follows. Section 2 proves the lower bounds in Theorem 1. Section 3 proves Theorem 2. Section 4 contains remarks and open problems.

2 Lower bounds

Throughout we fix even $r \leq 2n/3$ and work in the graph $H(n, r)$.

2.1 Forbidden-distance codes for fixed even r

An *independent set* in $H(n, r)$ is a set $C \subseteq \mathbb{F}_2^n$ such that $d_H(x, y) \neq r$ for all distinct $x, y \in C$. In coding-theory, C is a binary code with forbidden distance r . Such sets are automatically triangle-free, so the lower bounds on the independent sets give lower bounds on $T(n, r)$.

For fixed even $r = 2d$, Bassalygo, Cohen and Zémor [2] construct large linear codes with a single forbidden weight. In particular, they show that the largest dimension of a linear code with $2d$ as a forbidden distance is $n - d \log n + O(1)$. This implies the following.

Lemma 1. *Fix $d \geq 1$. There is a constant $c_d > 0$ such that*

$$T(n, 2d) \geq c_d \frac{2^n}{n^d} \quad (n \rightarrow \infty).$$

2.2 A probabilistic bound

We now prove the alteration bound. We view triangles as hyperedges in a 3-uniform hypergraph and remove one vertex from each surviving triangle.

Lemma 2. *Let r be even and satisfy $r \leq 2n/3$. Then $H(n, r)$ has a triangle-free vertex set of size at least*

$$\frac{2\sqrt{2}}{3} \cdot \frac{2^n}{\sqrt{\binom{n}{r} \binom{r}{r/2} \binom{n-r}{r/2}}}.$$

Proof. Let $\mathcal{H}(n, r)$ be the 3-uniform hypergraph on the vertex set \mathbb{F}_2^n whose hyperedges are the triangles of $H(n, r)$. We choose each vertex independently with probability p . Let X be the number of chosen vertices and let Y be the number of hyperedges of $\mathcal{H}(n, r)$ that are fully contained in the chosen set.

Removing one vertex from each such hyperedge produces a triangle-free set of size at least $X - Y$. By linearity of expectation,

$$\mathbb{E}[X - Y] = 2^n p - p^3 \cdot |E(\mathcal{H}(n, r))|.$$

It remains to count $|E(\mathcal{H}(n, r))|$. Fix a vertex u . There are $\binom{n}{r}$ vertices v with $d_H(u, v) = r$. Fix such a neighbor v . A third vertex w forms a triangle with (u, v) exactly when w differs from u in $r/2$ of the r coordinates where u and v differ, and also differs from u in $r/2$ of the remaining $n - r$ coordinates. So there are

$$\binom{r}{r/2} \binom{n-r}{r/2}$$

choices for w . This counts each triangle 6 times, so

$$|E(\mathcal{H}(n, r))| = \frac{1}{6} 2^n \binom{n}{r} \binom{r}{r/2} \binom{n-r}{r/2}.$$

Substituting this into $\mathbb{E}[X - Y]$ and optimizing over p gives

$$p = \sqrt{\frac{2}{\binom{n}{r} \binom{r}{r/2} \binom{n-r}{r/2}}}.$$

With this choice,

$$\mathbb{E}[X - Y] \geq \frac{2\sqrt{2}}{3} \cdot \frac{2^n}{\sqrt{\binom{n}{r} \binom{r}{r/2} \binom{n-r}{r/2}}}.$$

So there exists a choice of vertices with $X - Y$ at least this large, and it is triangle-free by construction. \square

2.3 A two-bit construction

The next bound works best when r is a constant fraction of n . We place all vertices on the level $r/2$ and enforce a two-bit condition.

Lemma 3. *Let r be even. Then $H(n, r)$ has a triangle-free vertex set of size at least*

$$\binom{n}{r/2} - \binom{n-2}{r/2}.$$

Proof. Identify a vertex of \mathbb{F}_2^n with the subset of $[n]$ given by its 1-coordinates. Work on the level $\binom{[n]}{r/2}$. Let \mathcal{F} be the family of all $(r/2)$ -subsets that contain 1 or contain 2; or equivalently their first or the second bit is 1. Then

$$|\mathcal{F}| = \binom{n}{r/2} - \binom{n-2}{r/2}.$$

We claim that \mathcal{F} is triangle-free in $H(n, r)$. Suppose $A, B, C \in \mathcal{F}$ form a triangle. Since all three sets have size $r/2$, the condition $d_H(A, B) = r$ is equivalent to $A \cap B = \emptyset$. So a triangle in $H(n, r)$ inside this level would force A, B, C to be pairwise disjoint. That is impossible in \mathcal{F} , because among three sets each containing 1 or 2, two of them must share one of these two elements. So \mathcal{F} spans no triangle. \square

2.4 Proof of Theorem 1

Proof of Theorem 1. The four parts of Theorem 1 are proved by the three mechanisms established above: the forbidden-distance bound in Lemma 1, the alteration bound in Lemma 2, and the two-bit construction in Lemma 3. Hence $T(n, r)$ is at least the maximum of the relevant corresponding quantities.

The lower bound obtained from the alteration bound can be analyzed in several ways based on the relationship of r and n . We note two different upper bounds for binomial coefficients, namely, $\binom{n}{r} \leq (\frac{en}{r})^r$ for all values of r , and when $r = \alpha n$ then

$$\binom{n}{r} \leq \frac{1}{\sqrt{2\pi n\alpha(1-\alpha)}} 2^{nH_2(\alpha)} \leq c \cdot 2^{nH_2(\alpha)},$$

where $H_2(t) = -t \log_2 t - (1-t) \log_2 (1-t)$ is the binary entropy function, and c is a constant. The first bound gives a triangle-free set of size

$$\geq c_1 \cdot \frac{2^{n r^{3r/4}}}{n^{3r/4} e^r},$$

while the entropy bound for $r = \alpha n$ gives us,

$$\geq c_2 \cdot 2^{n(1-\frac{1}{2}H_2(\alpha)-\frac{\alpha}{2}-\frac{1-\alpha}{2}H_2(\frac{\alpha}{2(1-\alpha)}))}$$

where $\frac{\alpha}{2(1-\alpha)} < 1$ when $\alpha < \frac{2}{3}$.

For analyzing the lower bound obtained from the two-bit construction, we note that,

$$|\mathcal{F}| \geq \binom{n}{r/2} \left(1 - \frac{(n-r/2)^2}{n^2}\right)$$

When $r = \alpha n$, it is known that $\binom{n}{r} \geq \frac{2^{nH_2(\alpha)}}{n+1}$. This gives us a triangle-free set of size

$$\geq c_3 \cdot 2^{nH_2(\alpha/2) - \log_2(n+1)}$$

for some constant c_3 that depends on α .

We now record which construction dominates in the main asymptotic regimes. For fixed even $r = 2d$, the forbidden-distance bound is stronger than the alteration bound. For $r = o(n)$ with $r \rightarrow \infty$, we apply the alteration bound.

For $r = \alpha n$ with fixed $0 < \alpha \leq 2/3$, we must compare the two bounds given by the alteration method and the two-bit construction, namely,

$$c_2 \cdot 2^{n(1-\frac{1}{2}H_2(\alpha)-\frac{\alpha}{2}-\frac{1-\alpha}{2}H_2(\frac{\alpha}{2(1-\alpha)}))} \text{ and } c_3 \cdot 2^{n(H_2(\alpha/2)-o(1))}.$$

There is a threshold $\alpha_0 \in (0, 2/3)$ such that alteration method dominates for $0 < \alpha < \alpha_0$ and the two-bit construction dominates for $\alpha_0 < \alpha \leq 2/3$. Numerically, we find this to be at $\alpha_0 \approx 0.16241686$. \square

3 Upper bounds

Before turning to upper bounds, we note that Balogh, Chen, and Li [1] recently proved that for every fixed even $r = 2d$,

$$T(n, 2d) = \Theta\left(\frac{2^n}{n^d}\right).$$

In particular, the lower bound in Lemma 1 is of the correct order in the constant regime.

We start with a uniform upper bound that only uses regularity of the triangle hypergraph. We then prove a stronger estimate when $2r \leq n$.

3.1 A universal $\frac{2}{3}$ upper bound

Let $\mathcal{H}(n, r)$ be the 3-uniform hypergraph whose vertices are \mathbb{F}_2^n and whose hyperedges are the triangles of $H(n, r)$. Let $N = 2^n$. Note that $\mathcal{H}(n, r)$ is regular. Let $D > 0$ be the number of hyperedges that contain a fixed vertex.

Lemma 4. *Every independent set $A \subseteq V(\mathcal{H}(n, r))$ satisfies*

$$|A| \leq \frac{2N}{3}.$$

Proof. Let $C := V(\mathcal{H}(n, r)) \setminus A$. Every hyperedge not contained in A must meet C . Hence the number of hyperedges that meet C is at most

$$\sum_{v \in C} \deg_{\mathcal{H}(n, r)}(v) = |C|D.$$

On the other hand, the total number of hyperedges is

$$|E(\mathcal{H}(n, r))| = \frac{1}{3} \sum_{v \in V(\mathcal{H}(n, r))} \deg_{\mathcal{H}(n, r)}(v) = \frac{ND}{3},$$

since each hyperedge has 3 vertices.

Let X be the number of hyperedges fully contained in A . Then

$$X \geq |E(\mathcal{H}(n, r))| - |C|D = \frac{ND}{3} - |C|D = D\left(\frac{N}{3} - |C|\right).$$

Since A is independent in $\mathcal{H}(n, r)$, we have $X = 0$. Therefore

$$0 \geq D\left(\frac{N}{3} - |C|\right).$$

\square

3.2 Sharper bounds when $2r \leq n$

We begin with the special case $r = 2$, since it introduces the main idea in its simplest form. We then generalize the same method to all even r in the regime $2r \leq n$. In the regime $\frac{n}{2} < r < \frac{2n}{3}$ we do not know an upper bound of order $o(2^n)$.

For the proofs we will work level by level in the hypercube and use a shadow (and cover) counting argument. We begin by fixing the notation for levels, shadows, and covers.

Levels, m -shadows, and m -covers. For $0 \leq k \leq n$, let

$$L_k := \{x \in \mathbb{F}_2^n : |x| = k\}$$

be the k th level of the hypercube, where $|x|$ denotes the Hamming weight of x (the number of 1's in x). Fix an integer $m \geq 1$. Given $x \in L_k$, an m -shadow of x is obtained by changing exactly m of its 1's to 0's. Thus every m -shadow lies in L_{k-m} and x has exactly $\binom{k}{m}$ m -shadows. Dually, an m -cover of x is obtained by changing exactly m of its 0's to 1's. Thus every m -cover lies in L_{k+m} and x has exactly $\binom{n-k}{m}$ m -covers.

Theorem 3. *Any triangle-free subset of $H(n, 2)$ has size at most $\frac{4 \cdot 2^n}{n}$.*

Proof. Let $A \subseteq \mathbb{F}_2^n$ be triangle-free in $H(n, 2)$ and define $S_k := |A \cap L_k|$. We bound S_k for each k and then sum over k .

Fix $k \geq 1$ and consider vertices in $A \cap L_k$. Each vertex in L_k has exactly $\binom{k}{1} = k$ number of 1-shadows in L_{k-1} . We claim that any vertex in L_{k-1} can be a 1-shadow of at most two vertices of $A \cap L_k$. Indeed, if three distinct vertices in L_k had the same 1-shadow, then they pairwise differ only in which one-bit was removed, so they have pairwise Hamming distance 2 and form a triangle in $H(n, 2)$.

Therefore, counting 1-shadows in two ways gives

$$kS_k \leq 2 \binom{n}{k-1}, \quad \text{so} \quad S_k \leq \frac{2}{k} \binom{n}{k-1}.$$

Summing over k ,

$$|A| \leq 1 + \sum_{k=1}^n \frac{2}{k} \binom{n}{k-1} = 1 + \sum_{k=1}^n \frac{2}{n+1} \binom{n+1}{k} \leq 1 + \frac{2}{n+1} \cdot 2^{n+1} \leq \frac{4 \cdot 2^n}{n}.$$

□

In the above summation, we could have used 1-shadows when $k \leq n/2$ and then 1-covers when $k > n/2$. The analysis for covers is symmetric to the analysis for shadows, and gives the same final bound. Although it is not necessary for $r = 2$, we will prefer this viewpoint for the general upper bound.

Before we proceed to the bound for general r , we need a set-theoretic ingredient. Identifying a vertex of the hypercube with its support, we may view vertices on a fixed level as k -subsets of $[n]$. In the argument below, triangle-freeness will force certain such families to contain no three pairwise disjoint sets. This places us in the setting of the Erdős Matching problem [4], which asks for the largest k -uniform family with no s pairwise disjoint sets. We will use the following theorem due to Frankl [6].

Theorem 4. *Suppose that $\mathcal{F} \subseteq \binom{[X]}{k}$, where $|X| \geq ks$, and \mathcal{F} contains no s pairwise disjoint sets. Then,*

$$|\mathcal{F}| \leq (s-1) \binom{|X|-1}{k-1}.$$

We now present the general upper bound.

Theorem 5. *Assume r is even and $2r \leq n$. Then any triangle-free subset of $H(n, r)$ has size $O\left(\frac{r2^n}{n+1}\right)$.*

Proof. Let $A \subseteq \mathbb{F}_2^n$ be triangle-free in $H(n, r)$ and define $S_k := |A \cap L_k|$. We bound S_k for each k and then sum over k .

Fix k and consider vertices in $A \cap L_k$. For $x \in L_k$, consider its $\frac{r}{2}$ -shadows in $L_{k-\frac{r}{2}}$. Thus each $x \in L_k$ has exactly $\binom{k}{r/2}$ number of $\frac{r}{2}$ -shadows.

Fix an $\frac{r}{2}$ -shadow $y \in L_{k-\frac{r}{2}}$. Any $x \in L_k$ having $\frac{r}{2}$ -shadow y is obtained by choosing an $\frac{r}{2}$ -subset of the 0-coordinates of y and turning them into 1's. Equivalently, after fixing y , the possible vertices x correspond to $\frac{r}{2}$ -subsets of a ground set of size

$$n - \left(k - \frac{r}{2}\right).$$

Among vertices of $A \cap L_k$ with the same $\frac{r}{2}$ -shadow y , we cannot have three choices whose added $\frac{r}{2}$ -subsets are pairwise disjoint, because those three vertices would then have pairwise Hamming distance r and would form a triangle in $H(n, r)$. Therefore, for each fixed $\frac{r}{2}$ -shadow y , the family of added $\frac{r}{2}$ -subsets contains no three pairwise disjoint sets.

We now apply Theorem 4 with $s = 3$ and $k = \frac{r}{2}$. This requires the ground set to have size at least $3\frac{r}{2}$, namely

$$n - \left(k - \frac{r}{2}\right) \geq 3\frac{r}{2}.$$

Under this hypothesis, Theorem 4 implies that for each fixed $\frac{r}{2}$ -shadow y , the number of vertices in $A \cap L_k$ having $\frac{r}{2}$ -shadow y is at most

$$2 \cdot \binom{n - \left(k - \frac{r}{2}\right) - 1}{\frac{r}{2} - 1}.$$

Now double count pairs (x, y) where $x \in A \cap L_k$ and y is an $\frac{r}{2}$ -shadow of x . There are $\binom{k}{\frac{r}{2}} S_k$ such pairs. On the other hand, there are $\binom{n}{k-\frac{r}{2}}$ possible $\frac{r}{2}$ -shadows, and for each $\frac{r}{2}$ -shadow at most $2 \cdot \binom{n - \left(k - \frac{r}{2}\right) - 1}{\frac{r}{2} - 1}$ choices of x . Hence

$$\binom{k}{\frac{r}{2}} S_k \leq 2 \binom{n}{k - \frac{r}{2}} \binom{n - \left(k - \frac{r}{2}\right) - 1}{\frac{r}{2} - 1}.$$

Simplifying this inequality gives

$$S_k \leq \binom{n}{k} \cdot \frac{r}{n - (k - r/2)}.$$

When $k \leq n/2$ and $2r \leq n$, we have

$$n - \left(k - \frac{r}{2}\right) = n - k + \frac{r}{2} \geq \frac{n}{2} + \frac{r}{2} \geq 2\frac{r}{2} + \frac{r}{2} = 3\frac{r}{2},$$

so the hypothesis $n - \left(k - \frac{r}{2}\right) \geq 3\frac{r}{2}$ that is needed for Theorem 4 holds in this range. When $k > n/2$, we repeat the same argument using $\frac{r}{2}$ -covers instead of $\frac{r}{2}$ -shadows, which gives the same bound with k replaced by $n - k$. Therefore, when summing over all levels, we sum up to $n/2$ and then double the result.

Using the trivial bound $S_k \leq \binom{n}{k}$ for $k \leq \frac{r}{2}$, we obtain

$$\begin{aligned}
|A| &= \sum_{k=0}^n S_k \leq 2 \left(\sum_{k=0}^{\frac{r}{2}} \binom{n}{k} + \sum_{k=\frac{r}{2}+1}^{n/2} \binom{n}{k} \cdot \frac{r}{n - (k - r/2)} \right) \\
&\leq 2 \left(\sum_{k=0}^{\frac{r}{2}} \binom{n}{k} + \sum_{k=\frac{r}{2}+1}^{n/2} \binom{n}{k} \cdot \frac{r}{n - k + 1} \right) \\
&\leq 2 \left(\sum_{k=0}^{\frac{r}{2}} \binom{n}{k} + \sum_{k=\frac{r}{2}+1}^{n/2} \binom{n+1}{k} \cdot \frac{r}{n+1} \right) = O\left(\frac{r2^n}{n+1}\right).
\end{aligned}$$

□

4 Remarks and open problems

For the regime where r is fixed and $n \rightarrow \infty$, Balogh, Chen, Li in [1] give the exact order of growth of the largest triangle-free set. They show that $T(n, r) = \Theta\left(\frac{2^n}{n^{r/2}}\right)$. Thus we can identify the following open problem in the regime where r is not a constant.

When $r = o(n)$, our upper bound gives

$$T(n, r) = O\left(\frac{r2^n}{n+1}\right) = o(2^n).$$

By contrast, when $r = \alpha n$ with fixed $0 < \alpha \leq 1/2$, our upper bound only gives

$$T(n, r) = O(2^n),$$

This leads to the following questions.

Problem 1. Determine the correct asymptotic behavior of $T(n, r)$ when $r = o(n)$ and $r \rightarrow \infty$.

Problem 2. Determine the asymptotic behavior of $T(n, r)$ in the linear regime $r = \alpha n$. In particular, for fixed $\alpha \in (0, \frac{1}{2})$, is it true that

$$T(n, \alpha n) \leq 2^{(1-\varepsilon_\alpha)n}$$

for some $\varepsilon_\alpha > 0$?

Acknowledgments. The first author is grateful to Dániel Varga for introducing her to the problem and for his many helpful suggestions.

References

- [1] J. Balogh, C. Chen, B. Li, Forbidding Exactly One Hamming Distance, *arXiv:2604.05607*, 2026.
- [2] L. A. Bassalygo, G. Cohen, and G. Zémor, Codes with forbidden distances, *Discrete Mathematics* **213** (2000), no. 1–3, 3–11.

- [3] H. Enomoto, P. Frankl, N. Ito, and K. Nomura, Codes with given distances, *Graphs and Combinatorics* **3** (1987), 25–38.
- [4] P. Erdős, A problem on independent r -tuples, *Ann. Univ. Sci. Budapest.* 8, 93–95, 1965.
- [5] P. Frankl and V. Rödl, Forbidden intersections, *Transactions of the American Mathematical Society* **300** (1987), no. 1, 259–286.
- [6] P. Frankl, The shifting technique in extremal set theory, *Surveys in Combinatorics* **123** (1987), 81–110.
- [7] P. Frankl, Improved bounds for Erdős’ matching conjecture, *Journal of Combinatorial Theory, Series A* **120** (2013), no. 5, 1068–1072.
- [8] M. Kauers, R. O’Donnell, L.-Y. Tan, and Y. Zhou, Hypercontractive inequalities via SOS, and the Frankl–Rödl graph, *arXiv:1212.5324*, 2012.
- [9] D. Kleitman, On a combinatorial conjecture of Erdős, *Journal of Combinatorial Theory, Series A* **43** (1986), 85–90.
- [10] M. Grötschel, L. Lovász, and A. Schrijver, Relaxations of vertex packing, *Journal of Combinatorial Theory, Series B* **40** (1986), 330–343.
- [11] L. Lovász, On the Shannon capacity of a graph, *IEEE Transactions on Information Theory* **25** (1979), no. 1, 1–7.
- [12] P. Mukkamala, Triangle-free subsets of the Hypercube, *arXiv:2506.18782v1*, 2025.
- [13] D. Castro-Silva, F. M. de Oliveira Filho, L. Slot, and F. Vallentin, A recursive theta body for hypergraphs, *Combinatorica* **43** (2023), 909–938.