Math. Proc. Camb. Phil. Soc. (2026), **181**, 1001–1020
doi:10.1017/S0305004126101984
First published online 16 April 2026

Set families: restricted distances via restricted intersections

BY ZICHAO DONG[†]
Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea.
e-mail: zichao@ibs.re.kr

JUN GAO[‡]
Mathematics Institute, University of Warwick, Coventry, UK.
e-mail: gj950211@gmail.com

HONG LIU[†]
Extremal Combinatorics and Probability Group (ECOPRO), Institute for Basic Science (IBS), Daejeon, South Korea.
e-mail: hongliu@ibs.re.kr

MINGHUI OUYANG[§]
School of Mathematical Sciences, Peking University, Beijing 100871, China.
e-mail: ouyangminghui1998@gmail.com

AND QIANG ZHOU
Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China, and University of Chinese Academy of Sciences, Beijing, China.
e-mail: zhouqiang2021@amss.ac.cn

(*Received* 27 *April* 2025*; revised* 26 *October* 2025*; accepted* 01 *December* 2025)

Abstract

Denote by $f_D(n)$ the maximum size of a set family $\mathcal{F}$ on $[n] \stackrel{def}{=} \{1, \ldots, n\}$ with distance set D. That is, $|A \triangle B| \in D$ holds for every pair of distinct sets $A, B \in \mathcal{F}$. Kleitman's celebrated discrete isodiametric inequality states that $f_D(n)$ is maximised at Hamming balls of radius $d/2$ when $D = \{1, \ldots, d\}$. We study the generalisation where D is a set of arithmetic progression and determine $f_D(n)$ asymptotically for all homogeneous D. In the special case when D is an interval, our result confirms a conjecture of Huang, Klurman and Pohoata. Moreover, we demonstrate a dichotomy in the growth of $f_D(n)$, showing linear growth in n when D is a non-homogeneous arithmetic progression. Different from previous combinatorial and

[†]Supported by the Institute for Basic Science (IBS-R029-C4).

[‡]Supported by ERC Advanced Grant 101020255 and the Institute for Basic Science (IBS-R029-C4).

[§]Partially supported by a grant from New Cornerstone Science Foundation.

© The Author(s), 2026. Published by Cambridge University Press on behalf of The Cambridge Philosophical Society.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

spectral approaches, we deduce our results by converting the restricted distance problems to restricted intersection problems. Our proof ideas can be adapted to prove upper bounds on t-distance sets in Hamming cubes (also known as binary t-codes), which has been extensively studied by algebraic combinatorialists community, improving previous bounds from polynomial methods and optimisation approaches.

2020 Mathematics Subject Classification: 05D05 (Primary);
94B25, 05C35, 52C10 (Secondary)

1. *Introduction*

The study of set families with restricted intersections is fundamental in extremal set theory. A cornerstone result in this area is the Erdős–Ko–Rado theorem **[21]**, which states that for any pair of positive integers n and k with $n \geq 2k$, if $\mathcal{F} \subseteq \binom{[n]}{k}$ is a k-uniform intersecting family—meaning $A \cap B \neq \varnothing$ $(\forall A, B \in \mathcal{F})$—then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. This result can be generalised to t-intersecting families, where $|A \cap B| \geq t$ $(\forall A, B \in \mathcal{F})$. In this case, it holds that $|\mathcal{F}| \leq \binom{n-t}{k-t}$ for sufficiently large n, and the Ahlswede–Khachatrian theorem **[1]** provides a complete characterisation of the maximum size of a t-intersecting family $\mathcal{F}$ for every n, k and t.

Denote by $A \triangle B$ the symmetric difference between A and B, i.e. $(A \setminus B) \cup (B \setminus A)$, and refer to $|A \triangle B|$ as the distance between them. By replacing the intersection restriction "$|A \cap B| \geq t$" with the distance restriction "$|A \triangle B| \geq t$", we naturally arrive at a problem which is central to coding theory. Intensive research efforts have been dedicated to bounding the size of set families under various distance constraints, including Singleton bound **[48]**, Plotkin bound **[45]**, Hamming bound, Elias–Bassalygo bound **[9]** and Gilbert–Varshamov bound **[26, 30, 50]**. A comprehensive overview of these developments and their significance is available in the standard reference **[49]**.

A closely related problem, originally proposed by Erdős **[20]**, considers families $\mathcal{F} \subseteq 2^{[n]}$ with a forbidden intersection size $|A \cap B| \neq \ell$ $(\forall A, B \in \mathcal{F})$. Erdős conjectured that, with an offer of \$250 for a solution, if the ratio ℓ/n lies between $\big(\varepsilon, 1/2 - \varepsilon\big)$, then the family size $|\mathcal{F}|$ is exponentially smaller than 2^n. This conjecture was first solved by Frankl and Rödl **[23]**. Keevash and Long **[34]** later provided an alternative proof and gave a restricted distance result: if $d/n \in (\varepsilon, 1 - \varepsilon)$ and $\mathcal{F} \subseteq 2^{[n]}$ is a family with $|A \triangle B| \neq d$ $(\forall A, B \in \mathcal{F})$, then $|\mathcal{F}| \leq 2^{(1-\delta)n}$. We also direct readers to **[37]**.

While research on restricted intersections has found substantial applications in areas such as Ramsey theory and discrete geometry (see, e.g., [24, chapter 22]), comparatively little attention has been devoted to the restricted-distance families. Motivated by this gap, in this paper we study set families with restricted distances. For any distance set D, our objective is to investigate $f_D(n)$, the largest possible size of a D-distance family on $[n]$.

1·1. *Generalising Kleitman's isodiametric inequality*

Among all smooth convex bodies of fixed diameter in $\mathbb{R}^n$, the classical Euclidean isodiametric inequality (see **[22]**) asserts that the ball attains the maximum volume. In the discrete setting, solving a conjecture of Erdős, Kleitman **[35]** established an isodiametric result which

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

states that

$$f_D(n) = \begin{cases} \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{t} & \text{if } D = [2t], \\ 2\left(\binom{n-1}{0} + \binom{n-1}{1} + \cdots + \binom{n-1}{t}\right) & \text{if } D = [2t+1]. \end{cases}$$

The bounds above are optimal, as witnessed by Hamming balls of appropriate radius.

Kleitman's original proof relies on sophisticated set operations. Recently, Huang, Klurman and Pohoata **[29]** discovered a smart linear algebraic proof of Kleitman's result. They also applied their technique to general intervals as distance sets and established that

$$\big(1+o(1)\big)\cdot\frac{1}{\binom{t}{s}}\cdot\binom{n}{t-s} \le f_{\{2s+1,2s+2,\ldots,2t\}}(n) \le \big(1+o(1)\big)\cdot\binom{n}{t-s};$$

They conjectured that the lower bound gives the correct asymptotics.

We study the same problem for even more general distance sets that are arithmetic progressions. An integer arithmetic progression is *homogeneous* if its common difference divides its first term, and *non-homogeneous* otherwise. Denote by $D_{\text{even}} \stackrel{\text{def}}{=} \{d \in D\colon d \text{ is even}\}$ the set of all even integers in D.

THEOREM 1·1. *Let d, s, t be positive integers with $1 \le s \le t$ and $D = \big\{sd, (s+1)d, \ldots, td\big\}$ be a homogeneous arithmetic progression. As $n \to \infty$, the following hold*:

(i) *if d is even, then*

$$f_D(n) = \big(1+o(1)\big)\cdot\frac{1}{\binom{t}{|D_{\text{even}}|}}\cdot\binom{2n/d}{|D_{\text{even}}|} = \big(1+o(1)\big)\cdot\frac{1}{\binom{t}{t-s+1}}\cdot\binom{2n/d}{t-s+1};$$

(ii) *if d is odd, then*

$$\begin{aligned} f_D(n) &= \big(1+o(1)\big)\cdot\frac{c}{\binom{\lfloor t/2\rfloor}{|D_{\text{even}}|}}\cdot\binom{n/d}{|D_{\text{even}}|} \\ &= \big(1+o(1)\big)\cdot\frac{c}{\binom{\lfloor t/2\rfloor}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}}\cdot\binom{n/d}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}, \end{aligned}$$

where $c = 1$ if st is even and $c = 2$ if st is odd.

Theorem 1·1 provides asymptotically sharp bounds on $f_D(n)$ for every homogeneous arithmetic progression distance set D with arbitrary common differences, which greatly generalises previous results on interval distance sets. In particular, by setting $d_{1.1} \stackrel{\text{def}}{=} 1$, $s_{1.1} \stackrel{\text{def}}{=} 2s+1$, and $t_{1.1} \stackrel{\text{def}}{=} 2t$, we confirm the above Huang–Klurman–Pohoata conjecture that $f_{\{2s+1,\ldots,2t\}}(n) = \big(1+o(1)\big)\cdot 1/\binom{t}{s}\cdot\binom{n}{t-s}$.

Somewhat surprisingly, the non-homogeneous arithmetic progressions exhibit a completely different behaviour. In fact, we prove that $f_D(n)$ grows only linearly in this case.

THEOREM 1·2. *Given $n, d, s, t, a \in \mathbb{N}_+$ with $1 \le s \le t$ and $1 \le a < d$. For any non-homogeneous arithmetic progression $D = \big\{sd+a, (s+1)d+a, \ldots, td+a\big\}$, the following hold*:

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

(i) *if* $D_{\text{even}} = \varnothing$, *then* $f_D(n) = 2$ *when* $n \geq \min(D)$, *and* $f_D(n) = 1$ *otherwise*;

(ii) *if* $D_{\text{even}} \neq \varnothing$, *then* $\lfloor 2n/\min(D_{\text{even}}) \rfloor \leq f_D(n) \leq n + 2$.

Our approach differs from both the classical combinatorial shifting in **[35]** and the linear algebraic method in **[29]**. For upper bounds, the main idea is to convert the restricted distance problems to restricted intersection problems. In Theorem 1·1, given a family $\mathcal{F}$ with prescribed distances, the key is to identify within $\mathcal{F}$ a subfamily $\mathcal{F}'$ which constitutes the majority of our family $\mathcal{F}$ (Claim 3·6)and is located on at most two slices of the hypercube (Claim 3·7).The difficult case is when the distance set $D = \{sd, (s+1)d, \ldots, td\}$ is an arithmetic progression with odd dt and even s. To obtain the correct asymptotics for this case, we have to show that the two slices of $\mathcal{F}'$ cannot be both large at the same time. To this end, we show that between the two slices of $\mathcal{F}'$, there is a restricted (cross) intersection pattern and we need to analyse the structure of the common intersection of each slice. In Theorem 1·2, we apply a modular restricted intersection result. It is worth noting that the lower bound constructions towards Theorem 1·1 involve additional difficulty. Inspired by **[29]**, we employ the classical Rödl nibble method. When d, s, t are all odd, we have to build up a "double almost design" to achieve an additional factor 2.

1·2. *Bounding the size of binary t-codes*

Our method is versatile, and it applies to the study of binary t-codes as well. We begin with some terminologies. Let $\mathcal{M}$ be a metric space endowed with distance function $\text{dist}_{\mathcal{M}}$. For finite $X \subseteq \mathcal{M}$, denote by $D(X) \stackrel{\text{def}}{=} \{\text{dist}_{\mathcal{M}}(x, y) \colon x, y \in X,\ x \neq y\}$ the *distance set* of X. For fixed D, we define

$$A(\mathcal{M}, D) \stackrel{def}{=} \max\{|X| : X \subseteq \mathcal{M} \text{ such that } \quad D(X) = D\}.$$

Following the terminology of coding theory, we call X a *t-code* if $|D(X)| = t$. Denote the maximum cardinality of a t-code as $A(\mathcal{M}, t) \stackrel{\text{def}}{=} \max\{A(\mathcal{M}, D) : |D| = t\}$.

It is a fundamental challenge to establish good upper bounds on $A(\mathcal{M}, t)$. Historically, Einhorn and Schoenberg **[18, 19]** initiated the study of t-codes in Euclidean space $\mathcal{M} = \mathbb{R}^n$. As a classical application of the polynomial method, Larman, Rogers, and Seidel **[38]** (later slightly improved by Blokhuis **[11]**) deduced that $n(n+1)/2 \leq A(\mathbb{R}^n, 2) \leq (n+1)(n+2)/2$. The same problem for Euclidean sphere $\mathcal{M} = \mathbb{S}^{n-1} \subseteq \mathbb{R}^n$ has also been intensively studied. Delsarte, Goethals, and Seidel **[16]** established $n(n+1)/2 \leq A(\mathbb{S}^{n-1}, 2) \leq n(n+3)/2$. Related results and improvements can be found in a large series of papers **[7, 8, 27, 31, 39–43, 52]**.

We focus on such a problem concerning Hamming cubes. Let $\mathcal{H}_n$ be the set of binary strings $\mathbb{F}_2^n$ endowed with the Hamming distance. (For any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{F}_2^n$, their *Hamming distance* $\text{dist}_{\mathcal{H}_n}(\boldsymbol{x}, \boldsymbol{y})$ is the number of bits at which $\boldsymbol{x}$ differs from $\mathbf{y}$.) The study of binary codes has long been central to coding theory. Two-distance sets in $\mathcal{H}_n$, binary 2-codes, have received considerable attention as they connect to finite geometry. See **[12]** for an early survey and **[47]** for recent work. Due to the extensive interest on t-codes in $\mathbb{R}^n$ and $\mathbb{S}^{n-1}$, it is natural to bound $A(\mathcal{H}_n, t)$. Earliest such results date back to the influential work of Delsarte **[15]** in 1973, where he established $A(\mathcal{H}_n, t) \leq \sum_{i=0}^{t} \binom{n}{i}$. Most such results are obtained by either the polynomial method **[5, 10]** or the linear programming method **[15, 16]**, which reduce the problem to the study of spherical t-codes in $\mathbb{R}^n$. A folklore conjecture, communicated to us by Yu **[51]**, predicts that the maximum value is achieved by the t-set $\{2, 4, \ldots, 2t\}$.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

CONJECTURE 1·3 (Folklore). *For every sufficiently large n (with respect to t), we have*

$$A(\mathcal{H}_n, t) = A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big) = \begin{cases} \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{t-2} + \binom{n}{t} & \text{if } t \text{ is even}, \\ \binom{n}{1} + \binom{n}{3} + \cdots + \binom{n}{t-2} + \binom{n}{t} & \text{if } t \text{ is odd}. \end{cases}$$

Towards this folklore conjecture, Nozaki and Shinohara **[44]** refined the result of Delsarte **[15]** to $A(\mathcal{H}_n, t) \le \sum_{i:c_i>0} \binom{n}{i}$, where c_i are the coefficients defined by a polynomial related to D in terms of the Krawtchouk polynomial basis. Barg and Musin **[7]** further established an explicit formula $A(\mathcal{H}_n, t) \le \binom{n}{t} + \sum_{i=0}^{t-2} \binom{n}{i}$ under the condition $\sum_{d\in D} d \le tn/2$. For small values of t, partial results were proved in **[7]** for $t = 2, 3, 4$. Later, the $t = 2$ case was resolved by Barg et al. **[6]**.

Our next result establishes $A(\mathcal{H}_n, D) \le A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big)$ for any fixed D and every sufficiently large n, which makes significant progress towards the first equality in Conjecture 1·3. We shall also prove that the second equality holds in Conjecture 1·3.

THEOREM 1·4. *Let $D \subseteq \mathbb{N}_+$ be a t-set. If $D \ne \{2, 4, \ldots, 2t\}$, then as $n \to \infty$ we have*

$$A(\mathcal{H}_n, D) \le \Big(\frac{t}{t+1} + o(1)\Big) \cdot \binom{n}{t}.$$

When $n \ge 2t + 2$, we have

$$A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big) = \begin{cases} \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{t-2} + \binom{n}{t} & \text{if } t \text{ is even}, \\ \binom{n}{1} + \binom{n}{3} + \cdots + \binom{n}{t-2} + \binom{n}{t} & \text{if } t \text{ is odd}. \end{cases}$$

We remark that the bound $n \ge 2t + 2$ above is optimal. Indeed, when $n \le 2t + 1$, the behaviour changes into $A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big) = 2^{n-1}$. Here the lower bound is witnessed by the collection of all binary strings with odd (or even) numbers of 1-s, and the upper bound follows from the observation that $\mathrm{dist}_{\mathcal{H}_n}(\boldsymbol{x}, \boldsymbol{y})$ is odd when the number of 1-s in $\boldsymbol{x}$ and $\mathbf{y}$ are of opposite parities.

Theorem 1·4 can be viewed as a stability result, for it infers that the distance set $\{2, 4, \ldots, 2t\}$ is the unique maximiser, showing that $A(\mathcal{H}_n, D)$ is substantially smaller than $A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big)$ for any other fixed t-set D. In contrast, due to the leading term $\binom{n}{t}$, the previous bounds could not perform well when D differs from $\{2, 4, \ldots, 2t\}$.

Paper organisation. We first prove Theorem 1·1. The lower bounds are given in Section 2 and the upper bounds are deduced in Section 3. We then prove Theorem 1·2 in Section 4. In Section 4, we prove Theorem 1·4. Finally, we include some further discussions and open problems in Section 6.

2. *Proof of the lower bounds in Theorem 1·1*

In this section, we prove the lower bounds in Theorem 1·1. Depending on the parity of *dst*, we divide the constructions for the lower bounds into two cases.

THEOREM 2·1. *Suppose d, s, t are positive integers with $1 \le s \le t$. Let $D = \{sd, (s+1)d, \ldots, td\}$ be a homogeneous arithmetic progression. The following hold for sufficiently large n.*

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

(i) *if dst is even, then there exists a D-distance set family $\mathcal{F} \subseteq 2^{[n]}$ of size*

$$|\mathcal{F}| = \begin{cases} \big(1-o(1)\big)\cdot \dfrac{\binom{2n/d}{t-s+1}}{\binom{t}{t-s+1}} & \text{when } d \text{ is even,} \\ \big(1-o(1)\big)\cdot \dfrac{\binom{n/d}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}}{\binom{\lfloor t/2\rfloor}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}} & \text{when } d \text{ is odd.} \end{cases}$$

(ii) *if dst is odd, then there exists a D-distance set family $\mathcal{F} \subseteq 2^{[n]}$ of size*

$$|\mathcal{F}| = \big(2-o(1)\big)\cdot \frac{\binom{n/d}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}}{\binom{\lfloor t/2\rfloor}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}}.$$

Inspired by [29, theorem 3·2], we are going to apply the following lemma in our constructions to establish Theorem 2·1(i)· This lemma is known as "the Rödl nibble" **[46]**. (See also [3, section 4·7].)

LEMMA 2·2. *Let $s, t \in \mathbb{N}$ with $1 \le s \le t$. As $m \to \infty$, there exists a t-uniform set family $\mathcal{F} \subseteq 2^{[m]}$ of size $|\mathcal{F}| = \big(1-o(1)\big)\cdot\big(\binom{m}{t-s+1}/\binom{t}{t-s+1}\big)$ such that $|A\cap B| \le t-s$ holds for any pair of distinct $A, B \in \mathcal{F}$.*

Proof of Theorem 2·1(i). We first look at the case when d is even. Write $m \overset{\text{def}}{=} \lfloor 2n/d \rfloor$. Thanks to Lemma 2·2, we are able to find a t-uniform family $\mathcal{F}' \subseteq 2^{[m]}$ of size $|\mathcal{F}'| = (1-o(1))\cdot\big(\binom{m}{t-s+1}/\binom{t}{t-s+1}\big)$ such that $|A'\cap B'| \le t-s$ holds for any pair of distinct $A', B' \in \mathcal{F}'$. This implies that

$$|A' \triangle B'| \in \big\{2t, 2t-2, \ldots, 2t-2(t-s)\big\} = \big\{2s, 2s+2, \ldots, 2t\big\}.$$

Since d is even, we can replace each element in $[m]$ by a unique group of $d/2$ distinct elements in $[n]$. This replacement then turns $\mathcal{F}'$ into a D-distance family $\mathcal{F} \subset 2^{[n]}$ of desired size.

If d is odd, we replace D by D_{even} and the same construction as above gives us a D_{even}-distance (hence D-distance) family $\mathcal{F} \subset 2^{[n]}$ of desired size. That is, we begin with a $\lfloor t/2\rfloor$-uniform family of size $(1-o(1))\cdot\big(\binom{m}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}/\binom{\lfloor t/2\rfloor}{\lfloor t/2\rfloor-\lceil s/2\rceil+1}\big)$ and pairwise intersection size at most $\lfloor t/2\rfloor - \lceil s/2\rceil$, and then blowup each element in $[m]$ to a unique group of size d. (We have $\lfloor t/2\rfloor \ge \lceil s/2\rceil$ since st is even.)

To prove Theorem 2·1(ii), we need the following extension of Lemma 2·2.

LEMMA 2·3. *Let $s', t' \in \mathbb{N}_+$ with $1 \le s' \le t'+1$. As $m \to \infty$, there exist two t'-uniform families $\mathcal{F}_1, \mathcal{F}_2 \subseteq 2^{[m]}$, each of size $(1-o(1))\cdot\big(\binom{m}{t'-s'+1}/\binom{t'}{t'-s'+1}\big)$ such that*:

(i) *$|A\cap B| \le t'-s'$ holds for any pair of distinct $A, B \in \mathcal{F}_1$ or $A, B \in \mathcal{F}_2$, and*

(ii) *$|A\cap B| \le t'-s'+1$ holds for any pair of $A \in \mathcal{F}_1$, $B \in \mathcal{F}_2$.*

We remark that one can force $\mathcal{F}_1$ and $\mathcal{F}_2$ to be disjoint in Lemma 2·3. In fact, the second condition directly shows the disjointness, provided that $s' > 1$. However, this is unnecessary for our application.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

Proof of Theorem 2·1(ii) *assuming Lemma* 2·3. Let $s' \stackrel{\text{def}}{=} \lceil s/2 \rceil$, $t' \stackrel{\text{def}}{=} \lfloor t/2 \rfloor$. Conversely, this implies that $s = 2s' - 1$, $t = 2t' + 1$. Define the parameter $m \stackrel{\text{def}}{=} \lfloor n/d \rfloor - 1$.

Let $\mathcal{F}_1, \mathcal{F}_2 \subseteq 2^{[m]}$ be two t'-uniform families obtained from Lemma 2·3. Append a new element $m+1$ to each set in $\mathcal{F}_2$. Then replace each element in $[m+1]$ by a unique group of d distinct elements in $[n]$. Let the resulting families be $\mathcal{F}_1', \mathcal{F}_2'$. For any distinct $A, B \in \mathcal{F}_1'$ or $A, B \in \mathcal{F}_2'$, we have

$$|A \triangle B| \in d \cdot \big\{2t', 2t'-2, \ldots, 2t'-2(t'-s')\big\} = \big\{(s+1)d, (s+2)d, \ldots, (t-1)d\big\} \subseteq D.$$

Moreover, for any $A \in \mathcal{F}_1'$, $B \in \mathcal{F}_2'$, we have

$$|A \triangle B| \in d \cdot \big\{2t'+1, 2t'-1, \ldots, 2t'+1-2(t'-s'+1)\big\} = \big\{sd, (s+1)d, \ldots, td\big\} = D.$$

Since $\mathcal{F}_1'$ and $\mathcal{F}_2'$ are disjoint (due to element $m+1$), we see that $\mathcal{F}_1' \cup \mathcal{F}_2'$ is a D-distance family with

$$|\mathcal{F}_1' \cup \mathcal{F}_2'| = |\mathcal{F}_1'| + |\mathcal{F}_2'| = |\mathcal{F}_1| + |\mathcal{F}_2| = \big(2 - o(1)\big) \cdot \frac{\binom{n/d}{\lfloor t/2 \rfloor - \lceil s/2 \rceil + 1}}{\binom{\lfloor t/2 \rfloor}{\lfloor t/2 \rfloor - \lceil s/2 \rceil + 1}}.$$

We are left to establish Lemma 2·3. The Rödl nibble (Lemma 2·2) already implies the existence of $\mathcal{F}_1$ and $\mathcal{F}_2$ separately. However, it is difficult to control the intersections between sets in $\mathcal{F}_1, \mathcal{F}_2$. To achieve this, we need tools from combinatorial design theory.

An (N, q, r)-*design* is a family $\mathcal{S}$ of q-subsets of $[N]$ in which every r-subset of $[N]$ appears as a subset of exactly one set in $\mathcal{S}$. By considering the codegree of an i-subset for $i = 0, 1, \ldots, r-1$, an obvious necessary condition for the existence of an (N, q, r)-design is the *divisibility condition*:

$$\binom{q-i}{r-i} \text{divides} \binom{N-i}{r-i} \quad \text{for } i = 0, 1, \ldots, r-1.$$

A profound result due to Keevash **[32]** states that this divisibility condition is sufficient for large N.

THEOREM 2·4. *Let q, r be positive integers with $q \geq r$. For every sufficiently large positive integer N satisfying the divisibility condition, an (N, q, r)-design exists.*

We remark that alternative and simpler proofs of Theorem 2·4 can be found in **[14, 28, 33]**.

To prove Lemma 2·3, the idea is to begin with an $(N, t', t'-s'+2)$-design, and then extract $\mathcal{F}_1, \mathcal{F}_2$ successively from it. Recall that a hypergraph is *linear* if any pair of its edges intersects in at most one vertex, and a *matching* in a hypergraph is a set of vertex-disjoint hyperedges.

Informally, the following technical result from **[2, 36]** says that an almost perfect matching always exists in a uniform almost regular linear hypergraph.

THEOREM 2·5. *For any positive integers $\ell \geq 3$ and K, there exists some $D_0 = D_0(\ell, K)$ such that the following property holds: Suppose $D \geq D_0$ and let $\mathcal{H} = (V, \mathcal{E})$ be an ℓ-uniform linear hypergraph with $D - K\sqrt{D \log D} \leq \deg_{\mathcal{H}}(v) \leq D$ for all vertices $v \in V$. Then as $|V| \to \infty$, there exists a matching covering all but at most $O\big(|V| D^{-1/(\ell-1)} \log^{3/2} D\big)$ vertices.*

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

Due to the almost regularity restriction $D - K\sqrt{D\log D} \le \deg_{\mathcal{H}}(x) \le D$ in Theorem 2·5, we have to find a design (using the deep result Theorem 2·4) instead of an almost design (via Lemma 2·2).

Proof of Lemma 2·3. When $s' = 1$, upon setting $\mathcal{F}_1 = \mathcal{F}_2 \stackrel{\text{def}}{=} \binom{[m]}{t'}$ we are done. If $s' = t' + 1$, then we can choose two single-set families supported on distinct elements $\mathcal{F}_1 = [t']$, $\mathcal{F}_2 = \{t'+1, \dots, 2t'\}$.

Assume $t' \ge s' \ge 2$ then. Pick the largest integer $N \le m$ satisfying $N \equiv q \pmod{Q}$, where

$$Q \stackrel{\text{def}}{=} r! \cdot \prod_{i=0}^{r-1} \binom{q-i}{r-i}, \quad q \stackrel{\text{def}}{=} t', \quad r \stackrel{\text{def}}{=} t' - s' + 2.$$

Since s', t' (hence q, r) are fixed, we have $m - N = O(1)$. For $i = 0, 1, \dots, r-1$, we compute

$$\begin{aligned}(r-i)! \cdot \binom{N-i}{r-i} = (N-i)\cdots(N-r+1) &\equiv (q-i)\cdots(q-r+1) \\ &= (r-i)! \cdot \binom{q-i}{r-i} \pmod{Q}.\end{aligned}$$

Since $(r-i)! \cdot \binom{q-i}{r-i}$ divides Q, we have $\binom{q-i}{r-i}$ divides $\binom{N-i}{r-i}$. So, N satisfies the divisibility condition. Due to Theorem 2·4, we can find an $(N, t', t'-s'+2)$-design $\mathcal{S}$ on $[N]$ when m is sufficiently large.

Construct an ℓ-uniform linear hypergraph $\mathcal{H} = (V, \mathcal{E})$ of uniformity $\ell = \binom{t'}{t'-s'+1}$ as follows. Let V be the set of $(t'-s'+1)$-subsets of $[N]$. For each $A \in \mathcal{S}$, include a hyperedge $\{B \in V\colon B \subseteq A\}$ into $\mathcal{E}$. Since each $(t'-s'+2)$-subset of $[N]$ is contained in exactly one t'-subset in $\mathcal{S}$, the hypergraph $\mathcal{H}$ is linear and D-regular, where $D = (N - (t'-s'+1))/(t' - (t'-s'+1))$. By Theorem 2·5, since $D = \Omega(N)$ and $D_0 = O(1)$, there exists a matching $\mathcal{M}$ in the hypergraph $\mathcal{H}$ covering $\big(1 - o(1)\big)$-proportion of its vertices. Phrasing differently, this matching $\mathcal{M}$ misses $o\big(\binom{N}{t'-s'+1}\big)$ vertices. Take $\mathcal{F}_1$ to be the set of t'-sets in $\mathcal{S}$ which corresponds to the hyperedges in $\mathcal{M}$.

Consider the ℓ-uniform linear hypergraph $\mathcal{H}' = (V, \mathcal{E}')$ for $\mathcal{E}' \stackrel{\text{def}}{=} \mathcal{E} \setminus \mathcal{M}$. This is the hypergraph obtained by stripping a matching from the D-regular hypergraph $\mathcal{H}$. It follows that the degree of each vertex in $\mathcal{H}'$ is either $D-1$ or D. From Theorem 2·5 we deduce that $\mathcal{H}'$ also admits an almost perfect matching. Take $\mathcal{F}_2$ to be the set of t'-sets in $\mathcal{S}$ corresponding to this matching.

For any pair of distinct sets $A, B \in \mathcal{F}_1$ or $A, B \in \mathcal{F}_2$, since their corresponding hyperedges in $\mathcal{H}$ are vertex-disjoint, we have $|A \cap B| \le t' - s'$. Also, for any pair of $A \in \mathcal{F}_1$ and $B \in \mathcal{F}_2$, we have $|A \cap B| \le t' - s' + 1$ as $\mathcal{H}$ is linear. Since $m - N = O(1)$, the families $\mathcal{F}_1, \mathcal{F}_2$ meet all requirements in lem:rodlvariant, and hence the proof is complete.

3. *Proof of the upper bounds in Theorem* 1·1

In this section, we prove the following result which implies the upper bounds in Theorem 1·1.

THEOREM 3·1. *Suppose d, s, t are positive integers with $1 \le s \le t$. Let $D = \{sd, (s+1)d, \dots, td\}$ be a homogeneous arithmetic progression, and $\mathcal{F} \subseteq 2^{[n]}$ be a D-distance family.*

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

As $n \to \infty$, we have

$$|\mathcal{F}| \le \begin{cases} \big(1+o(1)\big) \cdot \prod\limits_{\ell \in D_{\text{even}}} \frac{2n}{\ell} & \text{if } dst \text{ is even,} \\ \big(2+o(1)\big) \cdot \prod\limits_{\ell \in D_{\text{even}}} \frac{2n}{\ell} & \text{if } dst \text{ is odd.} \end{cases}$$

We remark that if $D_{\text{even}} = \varnothing$ (*dst* is odd and $s = t$), conventionally the empty product is 1.

Indeed, when *dst* is odd, $\lfloor t/2 \rfloor = (t-1)/2$, $\lceil s/2 \rceil = (s+1)/2$, $|D_{\text{even}}| = \lfloor t/2 \rfloor - \lceil s/2 \rceil + 1 = (t-s)/2$, and so

$$\big(2+o(1)\big) \cdot \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell} = \big(2+o(1)\big) \cdot \frac{\left(\frac{n}{d}\right)^{\frac{t-s}{2}}}{\frac{s+1}{2} \cdot \frac{s+3}{2} \cdot \ldots \cdot \frac{t-1}{2}} = \big(2+o(1)\big) \cdot \frac{\binom{n/d}{(t-s)/2}}{\binom{(t-1)/2}{(t-s)/2}},$$

implying the upper bound in Theorem 1·1. Other cases are similar.

3·1. *Distance families versus intersecting families*

A set family $\mathcal{F}$ is *k-uniform* if $|A| = k$ $(\forall A \in \mathcal{F})$. For a set of non-negative integers L, we say that $\mathcal{F}$ is an *L-intersecting* family if $|A \cap B| \in L$ holds for each pair of distinct $A, B \in \mathcal{F}$. A classical result concerning intersecting families is the Frankl–Wilson theorem below.

THEOREM 3·2 **[25]**. *Suppose $0 \le \ell_1 < \ldots < \ell_r < k$ are integers and write $L \overset{def}{=} \{\ell_1, \ldots, \ell_r\}$. If $\mathcal{F} \subseteq 2^{[n]}$ is a k-uniform L-intersecting family, then $|\mathcal{F}| \le \binom{n}{r}$.*

Let $\mathcal{F}$ be a D-distance family. By flipping the belonging relation of a single number in $[n]$ to each set from $\mathcal{F}$ simultaneously, we may assume without loss of generality that $\varnothing \in \mathcal{F}$. It follows that $|A| = |A \triangle \varnothing| \in D$ holds for all $A \in \mathcal{F} \setminus \{\varnothing\}$. Partition $\mathcal{F} \setminus \varnothing$ into a disjoint union of $\mathcal{F}_d$ for $d \in D$. From the fact $2|A \cap B| = |A| + |B| - |A \triangle B|$ we deduce that $\mathcal{F}_d$ is an L_d-intersecting family, where the set of intersection sizes is $L_d \overset{\text{def}}{=} d - D_{\text{even}}/2 = \big\{d - k/2 \colon k \in D_{\text{even}}\big\}$. After dropping negative integers in L_d, the Frankl–Wilson theorem (Theorem 3·2) directly shows the following crude upper bound.

COROLLARY 3·3. *If $\mathcal{F} \subseteq 2^{[n]}$ is a D-distance family, then $|\mathcal{F}| \le |D| \cdot n^{|D_{\text{even}}|}$.*

One can see that the order of magnitude in Corollary 3·3 already matches that from the upper bounds in Theorem 1·1, and the most interesting part in our main result is to find the sharp leading coefficient. The following result proved by Deza, Erdős and Frankl **[17]** can be viewed as a refinement of the Frankl–Wilson theorem. This is crucial for us to establish the upper bounds in Theorem 1·1. Instead of the original statement, here we state a slightly weaker version.

THEOREM 3·4. **[24**, theorem 18·1]. *Let $0 \le \ell_1 < \ldots < \ell_r < k$ and $n \ge 2^k k^3$ be integers and write $L \overset{def}{=} \{\ell_1, \ldots, \ell_r\}$. If $\mathcal{F} \subseteq 2^{[n]}$ is a k-uniform L-intersecting family, then $|\mathcal{F}| \le \prod_{i=1}^{r} (n - \ell_i)/(k - \ell_i)$. Moreover, if $|\mathcal{F}| \ge 2^k k^2 n^{r-1}$, then there exists $C \subseteq [n]$ with $|C| = \ell_1$ such that C is contained in every set in $\mathcal{F}$.*

We also need the following corollary of Theorem 3·4.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

LEMMA 3·5. *If $\mathcal{G} \subseteq 2^{[n]}$ is a k-uniform D-distance set family with $n \geq 2^k k^3$, then $|\mathcal{G}| \leq \prod_{\ell \in D_{\text{even}}} 2n/\ell$.*

Proof. For distinct $A, B \in \mathcal{G}$, since $|A \triangle B| = |A| + |B| - 2|A \cap B|$, we see that $\mathcal{G}$ is an L-intersecting family, where $L \stackrel{\text{def}}{=} \big(k - D_{\text{even}}/2\big) \cap \mathbb{Z}_{\geq 0}$. It then follows from Theorem 3·4 that

$$|\mathcal{G}| \leq \prod_{\ell \in L} \frac{n-\ell}{k-\ell} \leq \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell}.$$

3.2. Proof of Theorem 3·1

The $D_{\text{even}} = \varnothing$ case is easy. We shall discuss it separately at the beginning of Section 4.

Here and after we assume that $D_{\text{even}} \neq \varnothing$. By passing to a subprogression of D if necessary, we may assume that there is a pair of sets $\mathcal{F}$ realising the maximum distance td in D. By flipping the belonging relation of a single element in $[n]$ to each set from $\mathcal{F}$ simultaneously and then relabeling, we may further assume that $\varnothing \in \mathcal{F}$ and $[td] \in \mathcal{F}$. Introduce the auxiliary subfamily

$$\mathcal{F}' \stackrel{\text{def}}{=} \big\{A \in \mathcal{F} \colon \big|A \setminus [td]\big| \geq d^*/2\big\}, \text{ where } d^* \stackrel{\text{def}}{=} \max(D_{\text{even}}).$$

To upper bound $|\mathcal{F}|$, we are to show that $|\mathcal{F} \setminus \mathcal{F}'|$ is small; and to upper bound $|\mathcal{F}'|$ by analyzing its intersection pattern and applying the Frankl–Wilson and the Deza–Erdős–Frankl theorems.

CLAIM 3.6. *As $n \to \infty$, we have $|\mathcal{F} \setminus \mathcal{F}'| = O\big(n^{|D_{\text{even}}|-1}\big)$.*

Proof. For each $S \subseteq [td]$, define $\mathcal{F}_S \stackrel{\text{def}}{=} \big\{A \in \mathcal{F} \setminus \mathcal{F}' \colon A \cap [td] = S\big\}$. It suffices to show that $|\mathcal{F}_S| = O\big(n^{|D_{\text{even}}|-1}\big)$. For distinct $A, B \in \mathcal{F}_S$, note that $|A \triangle B| \leq \big|A \setminus [td]\big| + \big|B \setminus [td]\big| < d^*$. Thus, $\mathcal{F}_S$ is a $D \setminus \{d^*\}$-distance family. The desired bound on $|\mathcal{F}_S|$ then follows from Corollary 3·3.

It remains to upper bound the size of $\mathcal{F}'$.

CLAIM 3.7. *For any $A \in \mathcal{F}'$, we have*

$$|A| = \begin{cases} td & \text{if } td \text{ is even,} \\ td \text{ or } (t-1)d & \text{if } td \text{ is odd.} \end{cases}$$

Proof. Write $x \stackrel{\text{def}}{=} \big|A \cap [td]\big|$ and $y \stackrel{\text{def}}{=} \big|A \setminus [td]\big| \geq d^*/2$. Since $\mathcal{F}'$ is a D-distance family,

$$y + (td - x) = \big|A \setminus [td]\big| + \big|[td] \setminus A\big| = \big|A \triangle [td]\big| \leq td \implies y \leq x.$$

It follows that $|A| = x + y \geq 2y \geq d^*$, and so $|A| \geq td$ if td is even while $|A| \geq (t-1)d$ if td is odd. On the other hand, $|A| \leq td$ as $\varnothing \in \mathcal{F}$, concluding the claim.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

If td is even, then $d^* = td$, and Claim 3.7 says that $\mathcal{F}'$ is a td-uniform D-distance set family. So, from Lemma 3·5 and Claim 3·6 we deduce that

$$|\mathcal{F}| = |\mathcal{F}'| + |\mathcal{F} \setminus \mathcal{F}'| \le \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell} + O\big(n^{|D_{\text{even}}|-1}\big) = \big(1 + o(1)\big) \cdot \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell}. \tag{3·1}$$

If td is odd, then $d^* = (t-1)d$. Due to Claim 3·7, we are able to partition $\mathcal{F}'$ into a td-uniform family $\mathcal{F}_1' \stackrel{\text{def}}{=} \{A \in \mathcal{F}' : |A| = td\}$ and a $(t-1)d$-uniform family $\mathcal{F}_2' \stackrel{\text{def}}{=} \{A \in \mathcal{F}' : |A| = (t-1)d\}$. It then follows from Lemma 3·5 and Claim 3·6 that

$$|\mathcal{F}| = |\mathcal{F}_1'| + |\mathcal{F}_2'| + |\mathcal{F} \setminus \mathcal{F}'| \le \big(2 + o(1)\big) \cdot \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell}. \tag{3·2}$$

By combining (3·1) and (3·2), Theorem 3·1 is proved in all but one case. We are left with improving the upper bound on $|\mathcal{F}|$ by a factor of 2 when d, t are odd and s is even. In this case,

$$D_{\text{even}} = \big\{sd, (s+2)d, \ldots, (t-1)d\big\}.$$

To achieve this, we need to analyse $\mathcal{F}_1', \mathcal{F}_2'$ carefully and utilise the structural part of Theorem 3·4.

For distinct $A, B \in \mathcal{F}'$, we compute the possible sizes of their intersection as follows:

$$|A \cap B| = \frac{|A| + |B| - |A \triangle B|}{2} \in \begin{cases} td - \frac{D_{\text{even}}}{2} & \text{if } A, B \in \mathcal{F}_1'; \\ (t-1)d - \frac{D_{\text{even}}}{2} & \text{if } A, B \in \mathcal{F}_2'; \\ (t-1)d - \frac{D_{\text{even}}}{2} & \text{if } A \in \mathcal{F}_1' \text{ and } B \in \mathcal{F}_2'. \end{cases}$$

Denote

$$L_1 \stackrel{\text{def}}{=} td - \tfrac{D_{\text{even}}}{2} = \Big\{\tfrac{(t+1)d}{2}, \tfrac{(t+3)d}{2}, \ldots, \tfrac{(2t-s)d}{2}\Big\},$$
$$L_2 \stackrel{\text{def}}{=} (t-1)d - \tfrac{D_{\text{even}}}{2} = \Big\{\tfrac{(t-1)d}{2}, \tfrac{(t+1)d}{2}, \ldots, \tfrac{(2t-s-2)d}{2}\Big\}.$$

It follows that $\mathcal{F}_1'$ is an L_1-intersecting family and $\mathcal{F}_2'$ is an L_2-intersecting family.

For any set family $\mathcal{G}$, we denote its *center* as $\bigcap \mathcal{G} \stackrel{\text{def}}{=} \bigcap_{A \in \mathcal{G}} A$. If the center of $\mathcal{F}_1'$ is of size smaller than $\min\big(td - D_{\text{even}}/2\big) = (t+1)d/2$, then Theorem 3·4 implies that $|\mathcal{F}_1'| = O\big(n^{|D_{\text{even}}|-1}\big)$, and hence

$$|\mathcal{F}| = |\mathcal{F}_1'| + |\mathcal{F}_2'| + |\mathcal{F} \setminus \mathcal{F}'| \le \big(1 + o(1)\big) \cdot \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell}.$$

We may then assume $\big|\bigcap \mathcal{F}_1'\big| \ge (t+1)d/2$ and similarly $\big|\bigcap \mathcal{F}_2'\big| \ge \min\big((t-1)d - D_{\text{even}}/2\big) = (t-1)d/2$.

Arbitrarily fix a $(t+1)d/2$-element subset $X_1 \subseteq \bigcap \mathcal{F}_1'$ and a $(t-1)d/2$-element subset $X_2 \subseteq \bigcap \mathcal{F}_2'$.

Case 1. $X_2 \subseteq X_1$.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

Denote $Y \overset{\text{def}}{=} X_1 \setminus X_2$ and observe that $|Y| = (t+1)d/2 - (t-1)d/2 = d$. For each $S \subseteq Y$, we write

$$\mathcal{F}_{2,S} \overset{\text{def}}{=} \{A \in \mathcal{F}_2' \colon A \cap Y = S\}$$

and thus partition $\mathcal{F}_2' = \bigcup_{S \subseteq Y} \mathcal{F}_{2,S}$ into $2^{|Y|} = 2^d$ subfamilies. For each non-empty $S \subseteq Y$ and any distinct $A, B \in \mathcal{F}_{2,S}$, we have $|A \cap B| \geq |X_2| + |S| > (t-1)d/2$. Then $\mathcal{F}_{2,S}$ is a $(t-1)d$-uniform and $L_2 \setminus \big\{(t-1)d/2\big\}$-intersecting family, and so from Theorem 3·2 we deduce that $|\mathcal{F}_{2,S}| = O\big(n^{|D_{\text{even}}|-1}\big)$.

Define $\mathcal{F}_{2,\varnothing}^{+} \overset{\text{def}}{=} \{A \cup Y \colon A \in \mathcal{F}_{2,\varnothing}\}$. Then, crucially, we observe that the union $\mathcal{F}_1' \cup \mathcal{F}_{2,\varnothing}^{+}$ forms a td-uniform L_1-intersecting family, since sets between $\mathcal{F}_1'$, $\mathcal{F}_2'$ are L_2-intersecting and $L_1 = L_2 + |Y|$. This implies $|\mathcal{F}_1' \cup \mathcal{F}_{2,\varnothing}^{+}| \leq \prod_{\ell \in D_{\text{even}}} 2n/\ell$ by Theorem 3·4.

We claim that $\mathcal{F}_1'$ and $\mathcal{F}_{2,\varnothing}^{+}$ are disjoint. Suppose there is some $A \cup Y \in \mathcal{F}_1' \cap \mathcal{F}_{2,\varnothing}^{+}$. Again, sets between $\mathcal{F}_1'$ and $\mathcal{F}_2'$ have intersection size at most $\max(L_2) = (2t-s-2)d/2 \leq (t-2)d$ as $s \geq 2$ is even. However, $A \cup Y \in \mathcal{F}_1'$ and $A \in \mathcal{F}_{2,\varnothing} \subseteq \mathcal{F}_2'$ have intersection size $|A| = (t-1)d$, a contradiction. So,

$$\begin{aligned} |\mathcal{F}| &= |\mathcal{F}_1'| + |\mathcal{F}_2'| + |\mathcal{F} \setminus \mathcal{F}'| \\ &= |\mathcal{F}_1' \cup \mathcal{F}_{2,\varnothing}^{+}| + \sum_{\varnothing \neq S \subseteq Y} |\mathcal{F}_{2,S}| + |\mathcal{F} \setminus \mathcal{F}'| \leq \big(1 + o(1)\big) \cdot \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell}. \end{aligned}$$

Case 2. $X_2 \nsubseteq X_1$.

Pick some element $x \in X_2$ and set $Z \in \mathcal{F}_1'$ such that $x \notin Z$. For each $T \subseteq Z$, we write

$$\mathcal{F}_{2,T} \overset{\text{def}}{=} \{A \in \mathcal{F}_2' \colon A \cap Z = T\}$$

and partition $\mathcal{F}_2'$ into $2^{|Z|} = 2^{td}$ subfamilies. Recall that for any $A \in \mathcal{F}_2'$, we have $|A \cap Z| \in L_2$, and so $\mathcal{F}_{2,T} = \varnothing$ whenever $|T| < \min(L_2) = (t-1)d/2$. Fix an arbitrary $T \subseteq Z$ with $|T| \geq (t-1)d/2$. For every distinct $A, B \in \mathcal{F}_{2,T}$, observe that $T \cup \{x\} \subseteq A \cap B$, hence $|A \cap B| \geq (t+1)d/2$. It follows that $\mathcal{F}_{2,T}$ is $(t-1)d$-uniform and $L_2 \setminus \{(t-1)d/2\}$-intersecting, and so $|\mathcal{F}_{2,T}| = O(n^{|D_{\text{even}}|-1})$ by Theorem 3·2. So,

$$|\mathcal{F}| = |\mathcal{F}_1'| + |\mathcal{F}_2'| + |\mathcal{F} \setminus \mathcal{F}'| = |\mathcal{F}_1'| + \sum_{T \subseteq Z} |\mathcal{F}_{2,T}| + |\mathcal{F} \setminus \mathcal{F}'| \leq \big(1 + o(1)\big) \cdot \prod_{\ell \in D_{\text{even}}} \frac{2n}{\ell}.$$

We thus conclude the case when s is even and d, t are odd. The proof of Theorem 3·1 is complete·

4. *Proof of Theorem* 1·2

The $D_{\text{even}} = \varnothing$ case is easy, and is already included in the previous work due to Huang, Klurman and Pohoata [29, section 3]. In fact, the D-distance family $\big\{\varnothing, [sd + a]\big\}$ shows $f_D(n) \geq 2$, and from

$$|A \triangle B| + |B \triangle C| + |C \triangle A| = 2\big(|A \cup B \cup C| - |A \cap B \cap C|\big)$$

we deduce that $f_D(n) \leq 2$. It follows that $f_D(n) = 2$ whenever $D_{\text{even}} = \varnothing$ and $n \geq \min(D)$.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

Hereafter we focus on the $D_{\text{even}} \neq \varnothing$ case. Set $u \stackrel{\text{def}}{=} \min(D_{\text{even}})/2$. Then any collection of $\lfloor n/u \rfloor$ disjoint u-subsets of $[n]$ is a D-distance family, and so $f_D(n) \geq \lfloor n/u \rfloor = \lfloor 2n/\min(D_{\text{even}}) \rfloor$. The lower bound follows.

To deduce the upper bound, we need the following result concerning modular intersecting families due to Babai, Frankl, Kutin and Štefankovičč.

THEOREM 4·1. [4, theorem 1.1]. *Suppose p is a prime, $q \stackrel{def}{=} p^k$ $(k \in \mathbb{N}_+)$, and $L \subseteq \{0, 1, \ldots, q-1\}$. Assume X is an n-element ground set and $A_1, \ldots, A_m$ are subsets of X with the following property*:

(1) *for any $i, j \in [m]$, the modular cardinality $|A_i \cap A_j| \pmod q \in L$ if and only if $i \neq j$.*

Set $D \stackrel{def}{=} 2^{|L|-1}$. Then $m \leq \binom{n}{D} + \binom{n}{D-1} + \cdots + \binom{n}{0}$.

Let $\mathcal{F}$ be a D-difference family and assume without loss of generality that $\varnothing \in \mathcal{F}$. Consider the family $\mathcal{F}' \stackrel{def}{=} \mathcal{F} \setminus \{\varnothing\}$. Then $|A| = |A \triangle \varnothing| \in D$ holds for any $A \in \mathcal{F}'$. For any sets A, B, note that

$$2|A \cap B| = |A| + |B| - |A \triangle B|. \tag{4·1}$$

Since $a < d$, we can find a prime power $q = p^k$ that divides d but not a. Here p is a prime.

(1) *If $p = 2$, then the fact $D_{\text{even}} \neq \varnothing$ implies that a is even and $k \geq 2$. For each $A \in \mathcal{F}'$, we have $|A| \equiv a \pmod{2^{k-1}}$. For distinct $A, B \in \mathcal{F}'$, from (4·1) we deduce that $|A \cap B| \equiv (a/2) \pmod{2^{k-1}}$. Notice that $a \not\equiv (a/2) \pmod{2^{k-1}}$ as q does not divide a. Thus, Theorem 4·1 applied to 2^{k-1} and $L = \{a/2\}$ implies that $|\mathcal{F}'| \leq n+1$, and so $|\mathcal{F}| \leq n+2$.*

(2) *If $p \geq 3$, then q is odd. For every $A \in \mathcal{F}'$, we have $|A| \equiv a \pmod q$. For distinct $A, B \in \mathcal{F}'$, from (4·1) we deduce that $2|A \cap B| \equiv a \pmod q$ hence $|A \cap B| \equiv a(q+1)/2 \pmod q$. Again, notice that $a \not\equiv a(q+1)/2 \pmod q$ since q does not divide a. Therefore, Theorem 4·1 applied to q and $L = \{a(q+1)/2\}$ implies that $|\mathcal{F}'| \leq n+1$, and so $|\mathcal{F}| \leq n+2$.*

By combining the cases above, the proof of Theorem 1·2 is complete.

5. *Proof of Theorem 1·4*

Fix a t-set $D \subseteq \mathbb{N}_+$ with $D \neq \{2, 4, \ldots, 2t\}$ and take n be sufficiently large in terms of t. Thanks to Corollary 3·3, it suffices to prove the special case that D consists of even numbers. Let $\mathcal{F}$ be a D-distance family. Write $d \stackrel{\text{def}}{=} \max(D)$ and assume without loss of generality that $\varnothing \in \mathcal{F}$, $[d] \in \mathcal{F}$. Then for any set $A \in \mathcal{F} \setminus \{\varnothing\}$, we have $|A| \in D$.

5·1. *Bounding $A(\mathcal{H}_n, D)$ away from $A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big)$*

If $d \geq 4t^2$, then we write $\mathcal{F}(z) \stackrel{\text{def}}{=} \{A \in \mathcal{F} \colon |A| = z\}$ for each $z \in D$. Due to Lemma 3·5,

$$|\mathcal{F}| = |\{\varnothing\}| + \sum_{z \in D} |\mathcal{F}(z)| \leq 1 + \sum_{z \in D} \left(\prod_{\ell \in D} \frac{2n}{\ell} \right) \leq 1 + t \cdot \frac{n^{t-1}}{(t-1)!} \cdot \frac{n}{2t^2} \leq \left(\frac{1}{2} + o(1) \right) \cdot \binom{n}{t}.$$

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

We may then assume $d < 4t^2$. Similar to the proof of Theorem 3·1, we denote

$$\mathcal{F}' \stackrel{\text{def}}{=} \big\{A \in \mathcal{F} \colon \big|A \setminus [d]\big| \geq d/2\big\}$$

and bound from above the sizes of $\mathcal{F}'$ and $\mathcal{F} \setminus \mathcal{F}'$ separately.

(1) For any $A \in \mathcal{F}'$, write $x \stackrel{\text{def}}{=} \big|A \cap [d]\big|$ and $y \stackrel{\text{def}}{=} \big|A \setminus [d]\big| \geq d/2$. Then $\big|A \triangle [d]\big| = d - x + y \leq d$, and so $x \geq y \geq d/2$, which implies that $|A| = x + y \geq d$. Combining this with $|A| = |A \triangle \varnothing| \leq d$, we see that $|A| = d$. So, $\mathcal{F}'$ is d-uniform. Thus, Lemma 3·5, along with the facts that D is a t-set consisting of even numbers and $D \neq \{2, 4, \ldots, 2t\}$, implies $|\mathcal{F}'| \leq \prod_{\ell \in D} 2n/\ell \leq n^t/(t-1)!(t+1)$.

(2) For any $S \subseteq [d]$, denote $\mathcal{F}_S \stackrel{\text{def}}{=} \big\{A \in \mathcal{F} \setminus \mathcal{F}' \colon A \cap [d] = S\big\}$. For any pair of distinct $A, B \in \mathcal{F}_S$, from $|A \triangle B| \leq \big|A \setminus [d]\big| + \big|B \setminus [d]\big| < d$ we deduce that $\mathcal{F}_S$ is a $D \setminus \{d\}$-distance family. Then Corollary 3·3 implies $|\mathcal{F}_S| \leq (t-1) \cdot n^{t-1}$. It follows that $|\mathcal{F} \setminus \mathcal{F}'| < 2^{4t^2}(t-1) \cdot n^{t-1}$.

Combining the estimates above, we conclude the first part of Theorem 1·4 by estimating

$$|\mathcal{F}| = |\mathcal{F}'| + |\mathcal{F} \setminus \mathcal{F}'| \leq \frac{n^t}{(t-1)!(t+1)} + 2^{4t^2}(t-1) \cdot n^{t-1} \leq \Big(\frac{t}{t+1} + o(1)\Big) \cdot \binom{n}{t}.$$

5·2. *The unique maximiser $D_0 = \{2, 4, \ldots, 2t\}$*

Let $n \geq 2t+2$ and $D_0 = \{2, 4, \ldots, 2t\}$. Then $A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big) = f_{D_0}(n)$. We need to prove that

$$f_{D_0}(n) = \begin{cases} \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{t-2} + \binom{n}{t} & \text{if } t \text{ is even,} \\ \binom{n}{1} + \binom{n}{3} + \cdots + \binom{n}{t-2} + \binom{n}{t} & \text{if } t \text{ is odd.} \end{cases}$$

When t is even (resp. odd), the bound is achieved by subsets of $[n]$ of even (resp. odd) size up to t.

Let $\mathcal{F} \subset 2^{[n]}$ be a D_0-distance set family. We are going to construct an auxiliary graph in which we may view $\mathcal{F}$ as an independent set. The following corollary of Cauchy's Interlacing Theorem was discovered earlier by Cvetković. This provides a useful technique to upper bound the independence number $\alpha(G)$ of a graph G. (A proof can be found after [29, corollary 2·5].)

PROPOSITION 5·1 ([13]). *Let $G = ([m], E)$ be an m-vertex graph and M be a pseudo-adjacency matrix of G, i.e., a symmetric $m \times m$ matrix with $M_{ij} = 0$ whenever $ij \notin E$. Let $n_{\leq 0}(M)$, $n_{\geq 0}(M)$ be the number of non-positive, non-negative eigenvalues of M. Then $\alpha(G) \leq \min\big\{n_{\leq 0}(M), n_{\geq 0}(M)\big\}$.*

If the distance $|X \triangle Y|$ is even, then we observe that $|X|$ and $|Y|$ are of the same parity. So, $\mathcal{F}$ consists solely of even-sized subsets or odd-sized subsets. By taking the symmetric difference with $\{1\}$ altogether, we may assume without loss of generality that $\mathcal{F} \subseteq \mathcal{X} \stackrel{\text{def}}{=} \big\{X \subseteq 2^{[n]} \colon |X| \text{ is even}\big\}$.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

For each integer $k \geq 0$, construct an $|\mathcal{X}| \times |\mathcal{X}|$ matrix M_k whose entries are either 0 or 1 with

$$(M_k)_{X,Y} = 1 \text{ if and only if } X, Y \in \mathcal{X} \text{ and } |X \triangle Y| = 2k.$$

Let $V \stackrel{\text{def}}{=} \mathbb{R}^{\mathcal{X}}$ be the vector space indexed by $\mathcal{X}$. For any $S \subseteq [n]$, set $\boldsymbol{v}_S \in V$ with $(\boldsymbol{v}_S)_X \stackrel{\text{def}}{=} (-1)^{|X \cap S|}$.

PROPOSITION 5·2. *We have $\boldsymbol{v}_S = \mathbf{v}_{[n]\setminus S}$ $(\forall S \subseteq [n])$ and the vectors $\{\boldsymbol{v}_S \colon S \subseteq [n]\}$ form an orthogonal basis of V. Moreover, the matrix M_k has an eigenvector $\boldsymbol{v}_S$ of eigenvalue $\sum_{i=0}^{|S|} (-1)^i \binom{|S|}{i}\binom{n-|S|}{2k-i}$.*

Proof. From $X \in \mathcal{X}$ we deduce that

$$|X \cap S| \equiv |X| - |X \cap S| = \big|X \cap \big([n] \setminus S\big)\big| \pmod 2,$$

and so $\boldsymbol{v}_S = \mathbf{v}_{[n]\setminus S}$. For any pair of $S, S' \subseteq [n]$ with $S' \neq S$ and $S' \neq [n] \setminus S$, we have

$$\begin{aligned}
\langle \boldsymbol{v}_S, \mathbf{v}_{S'} \rangle &= \sum_{X \in \mathcal{X}} (-1)^{|X \cap S|} \cdot (-1)^{|X \cap S'|} = \sum_{X \in \mathcal{X}} (-1)^{|X \cap (S \triangle S')|} \\
&= \sum_{k=0}^{\infty} \Bigg(\sum_{X \subseteq [n] \colon |X| = 2k} (-1)^{|X \cap (S \triangle S')|} \Bigg) \\
&= \sum_{k=0}^{\infty} \Bigg(\sum_{i=0}^{2k} (-1)^i \binom{|S \triangle S'|}{i} \binom{n - |S \triangle S'|}{2k - i} \Bigg) \\
&= \sum_{i=0}^{\infty} \Bigg((-1)^i \binom{|S \triangle S'|}{i} \cdot \sum_{k=\lceil i/2 \rceil}^{\infty} \binom{n - |S \triangle S'|}{2k - i} \Bigg) \\
&\stackrel{\text{(a)}}{=} \sum_{i=0}^{\infty} \Bigg((-1)^i \binom{|S \triangle S'|}{i} \cdot \frac{2^{n - |S \triangle S'|}}{2} \Bigg) \stackrel{\text{(b)}}{=} 0,
\end{aligned}$$

where (a) follows from $S' \neq [n] \setminus S$ together with $\sum_{i=0}^{\infty} (-1)^i \binom{m}{i} = 0$, and (b) follows from $S \neq S'$. Since $\big|\{\boldsymbol{v}_S \colon S \subseteq [n]\}\big| = 2^{n-1} = |\mathcal{X}|$, the set $\{\boldsymbol{v}_S \colon S \subseteq [n]\}$ forms an orthogonal basis of V.

Abbreviate $[n] \setminus Z$ as Z^c for any $Z \subseteq [n]$. For $X, Y \subseteq [n]$, observe that

$$|Y \cap S| = |X \cap S| + |X^c \cap Y \cap S| - |X \cap Y^c \cap S|.$$

Think about $\boldsymbol{v}_S$ as a column vector. It then follows that

$$\begin{aligned}
(M_k \cdot \boldsymbol{v}_S)_X &= \sum_{Y \in \mathcal{X}} (M_k)_{X,Y} \cdot (\boldsymbol{v}_S)_Y = \sum_{Y \subseteq [n] \colon |X \triangle Y| = 2k} (-1)^{|Y \cap S|} \\
&= (-1)^{|X \cap S|} \sum_{Y \subseteq [n] \colon |X \triangle Y| = 2k} (-1)^{|X^c \cap Y \cap S| + |X \cap Y^c \cap S|} \\
&\stackrel{\text{(c)}}{=} (-1)^{|X \cap S|} \sum_{Y \subseteq [n] \colon |Y| = 2k} (-1)^{|Y \cap S| + |\varnothing|} \\
&= (\boldsymbol{v}_S)_X \cdot \sum_{i=0}^{|S|} (-1)^i \binom{|S|}{i} \binom{n - |S|}{2k - i},
\end{aligned}$$

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

where at (c) we flip for every element of X the belonging between it and X, Y altogether, hence we assume $X = \varnothing$ in the summation. This is valid because such a flipping does not affect the summand $(-1)^{|X^c \cap Y \cap S| + |X \cap Y^c \cap S|}$. So, $\boldsymbol{v}_S$ is an eigenvector of M_k whose eigenvalue is exactly

$$\sum_{i=0}^{|S|} (-1)^i \binom{|S|}{i} \binom{n-|S|}{2k-i}.$$

Construct a graph G on vertex set $\mathcal{X}$, where we put an edge between distinct $S, T \in \mathcal{X}$ if and only if $|S \triangle T| \geq 2t+2$. The D_0-distance family $\mathcal{F}$ corresponds to an independent set in G. Define a matrix $M \stackrel{\text{def}}{=} \sum_{k \geq t+1} \binom{k-1}{t} M_k$. (The infinite range of k is unproblematic because $M_k = 0$ whenever $k > n/2$.) Then M is a pseudo-adjacency matrix of G if we identify $\mathcal{X}$ and $[|\mathcal{X}|]$.

For every integer k, denote by $[x^k] f(x)$ the coefficient of the x^k term in the Laurent expansion of a meromorphic function $f(x)$ in the $|x| > 1$ region on the complex plane. For instance, we have the identity $\binom{k-1}{t} = (-1)^{t+1} [x^{-2k}] \left(1/(1-x^2)^{t+1}\right)$ because of $\binom{-(t+1)}{k} = (-1)^k \binom{k+t}{t}$ and the expansion

$$\frac{1}{(1-x^2)^{t+1}} = (-1)^{t+1} x^{-2(t+1)} (1-x^{-2})^{-(t+1)} = (-1)^{t+1} x^{-2(t+1)} \sum_{k=0}^{\infty} \binom{k+t}{t} x^{-2k}.$$

By Proposition 5·2, $\boldsymbol{v}_S$ is an eigenvector of each M_k, so the eigenvalue of M with respect to it is

$$\begin{aligned}
\lambda_S &= \sum_{k \geq t+1} \left(\binom{k-1}{t} \sum_{i=0}^{|S|} (-1)^i \binom{|S|}{i} \binom{n-|S|}{2k-i} \right) \\
&= (-1)^{t+1} \sum_{k \geq t+1} [x^{-2k}] \left(\frac{1}{(1-x^2)^{t+1}} \right) \cdot [x^{2k}] \left((1-x)^{|S|} (1+x)^{n-|S|} \right) \\
&= (-1)^{t+1} [x^0] \left(\frac{(1-x)^{|S|} (1+x)^{n-|S|}}{(1-x^2)^{t+1}} \right).
\end{aligned}$$

(1) If $|S| \in \{t+1, \ldots, n-t-1\}$, then the denominator of $(1-x)^{|S|}(1+x)^{n-|S|} / (1-x^2)^{t+1}$ gets cancelled out and hence the quotient is a polynomial in x. So, the constant term is exactly the evaluation of the function at $x = 0$, which is 1. It follows that $\lambda_S = (-1)^{t+1}$.

(2) If $|S| \in \{1, \ldots, t\}$, then we compute

$$\begin{aligned}
[x^0] \frac{(1-x)^{|S|}(1+x)^{n-|S|}}{(1-x^2)^{t+1}} &= [x^0] \frac{(1+x)^{n-|S|-t-1}}{(1-x)^{t+1-|S|}} \\
&= [x^0] \left((-1)^{t+1-|S|} \cdot \frac{1}{x^{t+1-|S|}} \cdot \frac{(1+x)^{n-|S|-t-1}}{(1-\frac{1}{x})^{t+1-|S|}} \right) \\
&= (-1)^{t+1-|S|} [x^{t+1-|S|}] \\
&\quad \times \left((1+x)^{n-|S|-t-1} \left(1 + \frac{1}{x} + \cdots \right)^{t+1-|S|} \right).
\end{aligned}$$

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

It follows from $n \geq 2t+2$ that $n-|S|-t-1 \geq t+1-|S|$. So, every coefficient is non-zero, and hence the sign of λ_S is $(-1)^{t+1} \cdot (-1)^{t+1-|S|} = (-1)^{|S|}$.

(3) If $|S| \in \{n-t, \ldots, n\}$, then $\boldsymbol{\nu}_S = \mathbf{v}_{[n]\setminus S}$ (part of Proposition 5·2) implies that $\lambda_S = \lambda_{[n]\setminus S}$. It follows from the previous case that the sign of λ_S is $(-1)^{n-|S|}$.

According to Proposition 5·1,

(1) if t is even, then $\alpha(G) \leq n_{\geq 0}(M) = \binom{n}{0} + \binom{n}{2} + \cdots + \binom{n}{t}$;
(2) if t is odd, then $\alpha(G) \leq n_{\leq 0}(M) = \binom{n}{1} + \binom{n}{3} + \cdots + \binom{n}{t}$.

We thus conclude the proof of Theorem 1·4 by noticing that $|\mathcal{F}| \leq \alpha(G)$.

6. *Concluding remarks*

In this paper, we determine $f_D(n)$ asymptotically for all homogeneous arithmetic progress D. When D is non-homogeneous, Theorem 1·2 shows that $f_D(n)$ grows linearly in n. We are unable to derive an asymptotically tight bound on $f_D(n)$. It remains an interesting problem to determine the leading coefficient here. We propose the following easier problem.

Question 6·1. Suppose D is a non-homogeneous arithmetic progression with $\min(D) \geq 3$. Is it true that for every sufficiently large n, we have $f_D(n) \leq (1-\varepsilon)n$ for some $\varepsilon > 0$?

One natural relaxation of the above question is to impose additional uniformity constraints on the family $\mathcal{F}$ and make it an L-intersecting family for some L, where many well-developed methods and theories can be applied. A prototypical example of such a relaxed question is the following.

Question 6·2. Let $p \geq 3$ be a prime with $p \nmid r$. Suppose $\mathcal{F} \subseteq 2^{[n]}$ is a set family satisfying:

(i) $|A| \equiv r \not\equiv 0 \pmod{p}$ holds for any $A \in \mathcal{F}$, and
(ii) $|A \cap B| \equiv 0 \pmod{p}$ holds for any pair of distinct $A, B \in \mathcal{F}$.

Is it true that there exists some $\varepsilon > 0$ such that $|\mathcal{F}| \leq (1-\varepsilon)n$ all sufficiently large n?

In Question 6·2, the family $\mathcal{F}$ is an $\{2r, 2r+p, 2r+2p, \ldots\}$-distance family. When $r = 1$, this is evidently false as witnessed by all singletons in $[n]$. For a general r, Shengtong Zhang provided the following counterexample: let $q \equiv r \pmod{p}$ be a sufficiently large prime (power) and consider the finite projective plane $\mathrm{PG}(2, \mathbb{F}_q)$. It contains q^2+q+1 points, q^2+q+1 lines (each of size $q+1$), where any two lines intersect in exactly one point. By adding $p-1$ dummy points simultaneously into each line, one obtains an $\{r \bmod p\}$-intersecting family of size q^2+q+1 on $[q^2+q+p]$.

Nevertheless, this example does not refute Question 6·1, for the distance set $D = \{2q\}$ therein grows with n. One may further hope that, if D is a fixed non-homogeneous arithmetic progression with $\min D \geq 3$, then $f_D(n) \leq n/2$. However, the following shows that this is false when $p = 3$.

Example. There exists a 5-uniform $\{0, 3\}$-intersecting family on [n] of size $5\lfloor n/9 \rfloor$.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

Proof. For $i = 1, 2, \ldots, \lfloor n/9 \rfloor$, write $\widetilde{i} \stackrel{\text{def}}{=} 9(i-1)$ and consider

$$\widetilde{i} + \{1,2,3,4,5\},\ \widetilde{i} + \{1,2,3,6,7\},\ \widetilde{i} + \{1,2,3,8,9\},\ \widetilde{i} + \{1,2,4,6,8\},\ \widetilde{i} + \{1,3,4,6,9\}.$$

Then the $5\lfloor n/9 \rfloor$ many 5-element subsets from above form a $\{0,3\}$-intersection family.

This construction shows that the lower bounds in Theorem 1·2 are not asymptotically correct for some specific D, notably $D = \{4, 10\}$. To be specific, it implies that $f_{\{4,10\}}(n) \geq 5\lfloor n/9 \rfloor$.

Regarding Conjecture 1·3, our approach fails to deduce the first equality for all D. This is because in one of our tools, the Deza–Erdős–Frankl theorem (Theorem 3·4), one needs n to be at least exponential in the uniformity k. sec:exact in fact establishes a slightly stronger result than Theorem 1·4 exact that $A(\mathcal{H}_n, D) \leq A\big(\mathcal{H}_n, \{2, 4, \ldots, 2t\}\big)$ holds for all D and large n with $|D| = t$ and $\max D \leq \log n/2$. Towards a complete resolution of Conjecture 1·3, we suspect that essential new ideas are required. In particular, when $t = 1$, the existence of an n-by-n Hadamard matrix implies that $A\big(\mathcal{H}_n, \{2\}\big) = A\big(\mathcal{H}_n, \{n/2\}\big) = n$. This shows that, in constrast to the case when n is sufficiently large compared to D, the set $D_0 = \{2, 4, \ldots, 2t\}$ is not necessarily a unique maximiser of $A(\mathcal{H}_n, D)$ in general, and hence our proof cannot be easily adapted to resolve the conjecture.

Acknowledgments. We are grateful to Boris Bukh for bringing the Deza–Erdős–Frankl theorem to our attention. We also benefited from discussions with Andrey Kupavskii, Cosmin Pohoata, Wei-Hsuan Yu, and Shengtong Zhang. This work was initiated at the 2nd ECOPRO Student Research Program in the summer of 2024. MO would like to thank ECOPRO for hosting. Part of this work was done during a visit of ZD and MO to Shandong University. They are thankful to Guanghui Wang for his support.

REFERENCES

[1] R. AHLSWEDE and L. H. KHACHATRIAN. The complete intersection theorem for systems of finite sets. *European J. Combin.* **18**(2) (1997), 125–136.

[2] N. ALON, J. KIM and J. SPENCER. Nearly perfect matchings in regular simple hypergraphs. *Israel J. Math.* **100** (1997), 171–187.

[3] N. ALON and J. SPENCER. *The probabilistic method*. Wiley Series in Discrete Mathematics and Optimisation (John Wiley & Sons, Inc., Hoboken, NJ, fourth edition, 2016).

[4] L. BABAI, P. FRANKL, S. KUTIN and D. Štefankovicč. Set systems with restricted intersections modulo prime powers. *J. Combin. Theory Ser. A* **95**(1) (2001), 39–73.

[5] E. BANNAI, E. BANNAI and D. STANTON. An upper bound for the cardinality of an s-distance subset in real Euclidean space. II. *Combinatorica* **3**(2) (1983), 147–152.

[6] A. BARG, A. GLAZYRIN, W.-J. KAO, C.-Y. LAI, P.-C. TSENG and W.-H. YU. On the size of maximal binary codes with 2, 3 and 4 distances. *Comb. Theory* **4**(1) (2024), Paper No. 7, 28.

[7] A. BARG and O. R. MUSIN. Bounds on sets with few distances. *J. Combin. Theory Ser. A* **118**(4) (2011), 1465–1474.

[8] A. BARG and W.-H. YU. New bounds for spherical two-distance sets. *Exp. Math.* **22**(2) (2013), 187–194.

[9] L. A. BASSALYGO. New upper bounds for error-correcting codes. *Problemy PeredacŠi Informacii* **1**(4) (1965), 41–44.

[10] A. BLOKHUIS. *Few-distance sets*, volume 7 of *CWI Tract*. Stichting Mathematisch Centrum, Centrum voor Wiskunde en Informatica (Amsterdam, 1984).

[11] A. BLOKHUIS. A new upper bound for the cardinality of 2-distance sets in Euclidean space. In *Convexity and graph theory (Jerusalem,* 1981*)*. North-Holland Math. Stud. vol. 87 (North-Holland, Amsterdam, 1984), pp 65–66.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

[12] R. CALDERBANK and W. M. KANTOR. The geometry of two-weight codes. *Bull. London Math. Soc.* **18**(2) (1986), 97–122.
[13] D. M. CVETKOVIĆ. Chromatic number and the spectrum of a graph. *Publ. Inst. Math. (Beograd) (N.S.)* **14**(28) (1972), 25–38.
[14] M. DELCOURT and L. POSTLE. Refined absorption: a new proof of the existence conjecture. *Preprint*: arXiv:2402.17855 (2024).
[15] P. DELSARTE. Four fundamental parameters of a code and their combinatorial significance. *Information and Control* **23** (1973), 407–438.
[16] P. DELSARTE, J. M. GOETHALS and J. J. SEIDEL. Spherical codes and designs. *Geom. Dedicata* **6**(3) (1977), 363–388.
[17] M. DEZA, P. ERDŐS and P. FRANKL. Intersection properties of systems of finite sets. *Proc. London Math. Soc. (3)* **36**(2) (1978), 369–384.
[18] S. J. EINHORN and I. J. SCHOENBERG. On Euclidean sets having only two distances between points. I. *Indag. Math.* **28** (1966), 479–488. Nederl. Akad. Wetensch. Proc. Ser. A **69**.
[19] S. J. EINHORN and I. J. SCHOENBERG. On Euclidean sets having only two distances between points. II. *Indag. Math.* **28** (1966), 489–504. Nederl. Akad. Wetensch. Proc. Ser. A **69**.
[20] P. ERDŐS. Problems and results in graph theory and combinatorial analysis. In *Proceedings of the Fifth British Combinatorial Conference (Univ. Aberdeen, Aberdeen,* 1975), Congress. Numer. vol. XV of (Utilitas Math. Winnipeg, MB, 1976), pp. 169-192.
[21] P. ERDŐS, C. KO and R. RADO. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser.* (2) **12** (1961), 313–320.
[22] L. C. EVANS and R. F. GARIEPY. *Measure theory and fine properties of functions*. Textbooks in Mathematics (CRC Press, Boca Raton, FL, revised edition, 2015).
[23] P. FRANKL and V. RÖDL. Forbidden intersections. *Trans. Amer. Math. Soc.* **300**(1) (1987), 259–286.
[24] P. FRANKL and N. TOKUSHIGE. *Extremal problems for finite sets*. Stud. Math. Libr. vol. 86 (American Mathematical Society, Providence, RI, 2018).
[25] P. FRANKL and R. M. WILSON. Intersection theorems with geometric consequences. *Combinatorica* **1**(4) (1981), 357–368.
[26] E. N. GILBERT. A comparison of signalling alphabets. *The Bell System Technical Journal* **31**(3) (1952), 504–522.
[27] A. GLAZYRIN and W.-H. YU. Upper bounds for *s*-distance sets and equiangular lines. *Adv. Math.* **330** (2018), 810–833.
[28] S. GLOCK, D. KÜHN, A. LO and D. OSTHUS. The existence of designs via iterative absorption: hypergraph *F*-designs for arbitrary *F*. *Mem. Amer. Math. Soc.* **284**(1406) (2023), v+131.
[29] H. HUANG, O. KLURMAN and C. POHOATA. On subsets of the hypercube with prescribed hamming distances. *J. Combin. Theory Ser. A* **171** (2020), 105–156.
[30] T. JIANG and A. VARDY. Asymptotic improvement of the Gilbert-Varshamov bound on the size of binary codes. *IEEE Trans. Inform. Theory* **50**(8) (2004), 1655–1664.
[31] Z. JIANG, J. TIDOR, Y. YAO, S. ZHANG and Y. ZHAO. Spherical two-distance sets and eigenvalues of signed graphs. *Combinatorica* **43**(2) (2023), 203–232.
[32] P. KEEVASH. The existence of designs. *Preprint:* arXiv:1401.3665.
[33] P. KEEVASH. A short proof of the existence of designs. *Preprint:* arXiv:2411.18291 (2024).
[34] P. KEEVASH and E. LONG. Frankl-Rödl-type theorems for codes and permutations. *Trans. Amer. Math. Soc.* **369**(2) (2017), 1147–1162.
[35] D. J. KLEITMAN. On a combinatorial conjecture of Erdős. *J Combin. Theory* **1** (1966), 209–214.
[36] A. V. KOSTOCHKA and V. RÖDL. Partial Steiner systems and matchings in hypergraphs. *Random Structures Algorithms* **13**(3-4) (1998), 335–347.
[37] A. KUPAVSKII, A. SAGDEEV and D. ZAKHAROV. Cutting corners, *Preprint:* arXiv:2211.17150. (2022). Available at: https://www.sciencedirect.com/science/article/pii/S0095895625000930.
[38] D. G. LARMAN, C. A. ROGERS and J. J. SEIDEL. On two-distance sets in Euclidean space. *Bull. London Math. Soc.* **9**(3) (1977), 261–267.
[39] F.-Y. LIU and W.-H. YU. Semidefinite programming bounds for spherical three-distance sets. *Electron. J. Combin.* **31**(4) (2024), Paper No. 4.11, 23.
[40] O. R. MUSIN. Spherical two-distance sets. *J. Combin. Theory Ser. A* **116**(4) (2009), 988–995.
[41] O. R. MUSIN. Graphs and spherical two-distance sets. *European J. Combin.* **80** (2019), 311–325.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press

[42] O. R. MUSIN and H. NOZAKI. Bounds on three- and higher-distance sets. *European J. Combin.* **32**(8) (2011), 1182–1190.
[43] H. NOZAKI. A generalisation of Larman-Rogers-Seidel's theorem. *Discrete Math.* **311**(10-11) (2011), 792–799.
[44] H. NOZAKI and M. SHINOHARA. On a generalisation of distance sets. *J. Combin. Theory Ser. A* **117**(7) (2010), 810–826.
[45] M. PLOTKIN. Binary codes with specified minimum distance. *IRE Trans.*, IT-6 (1960), 445–450.
[46] V. RÖDL. On a packing and covering problem. *European J. Combin.* **6**(1) (1985), 69–78.
[47] M. SHI, T. HONOLD, P. SOLÉ, Y. QIU, R. WU and Z. SEPASDAR. The geometry of two-weight codes over $\mathbb{Z}_{p^m}$. *IEEE Trans. Inform. Theory* **67**(12) (2021), 7769–7781.
[48] R. C. SINGLETON. Maximum distance q-nary codes. *IEEE Trans. Inform. Theory*, IT-10 (1964), 116–118.
[49] J. H. VAN LINT. *Introduction to coding theory*. Grad. Texts Math. vol 86 (Springer-Verlag, Berlin, third edition, (1999).
[50] R. R. VARSHAMOV. Estimate of the number of signals in error correcting codes. *Doklady Akad. Nauk, S.S.S.R.* **117** (1957), 739–741.
[51] W.-H. YU. Private communication. To appear in *Combin. Theory.*
[52] W.-H. YU. New bounds for equiangular lines and spherical two-distance sets. *SIAM J. Discrete Math.* **31**(2) (2017), 908–917.

https://doi.org/10.1017/S0305004126101984 Published online by Cambridge University Press