

THE MINIMUM NUMBER OF DISJOINT PAIRS IN SET
SYSTEMS AND RELATED PROBLEMS

SHAGNIK DAS, WENYING GAN, BENNY SUDAKOV

Received May 29, 2013
Revised April 29, 2014
Online First February 1, 2016

Let \mathcal{F} be a set system on $[n]$ with all sets having k elements and every pair of sets intersecting. The celebrated theorem of Erdős, Ko and Rado from 1961 says that, provided $n \geq 2k$, any such system has size at most $\binom{n-1}{k-1}$. A natural question, which was asked by Ahlswede in 1980, is how many disjoint pairs must appear in a set system of larger size. Except for the case $k = 2$, solved by Ahlswede and Katona, this problem has remained open for the last three decades.

In this paper, we determine the minimum number of disjoint pairs in small k -uniform families, thus confirming a conjecture of Bollobás and Leader in these cases. Moreover, we obtain similar results for two well-known extensions of the Erdős–Ko–Rado Theorem, determining the minimum number of matchings of size q and the minimum number of t -disjoint pairs that appear in set systems larger than the corresponding extremal bounds. In the latter case, this provides a partial solution to a problem of Kleitman and West.

1. Introduction

A set system \mathcal{F} is said to be *intersecting* if $F_1 \cap F_2 \neq \emptyset$ for all $F_1, F_2 \in \mathcal{F}$. The Erdős–Ko–Rado Theorem is a classic result in extremal set theory, determining how large an intersecting k -uniform set system can be. This gives rise to the natural question of how many disjoint pairs must appear in larger set systems.

We consider this problem, first asked by Ahlswede in 1980. Given a k -uniform set system \mathcal{F} on $[n]$ with s sets, how many disjoint pairs must \mathcal{F}

Mathematics Subject Classification (2010): 05D05, 05C65

Research supported in part by NSF grant DMS-1101185, by AFOSR MURI grant FA9550-10-1-0569 and by a USA-Israel BSF grant.

contain? We denote the minimum by $\text{dp}(n, k, s)$, and determine its value for a range of system sizes s , thus confirming a conjecture of Bollobás and Leader in these cases. This results in a quantitative strengthening of the Erdős–Ko–Rado Theorem. We also provide similar results regarding some well-known extensions of the Erdős–Ko–Rado Theorem, which in particular allow us to partially resolve a problem of Kleitman and West.

We now discuss the Erdős–Ko–Rado Theorem and the history of this problem in greater detail, before presenting our new results.

1.1. Intersecting systems

Extremal set theory is one of the most rapidly developing areas of combinatorics, having enjoyed tremendous growth in recent years. The field is built on the study of very robust structures, which allow for numerous applications to other branches of mathematics and computer science, including discrete geometry, functional analysis, number theory and complexity.

One such structure that has attracted a great deal of attention over the years is the intersecting set system; that is, a collection \mathcal{F} of subsets of $[n]$ that is pairwise-intersecting. The most fundamental question one may ask is how large such a system can be. Observe that we must have $|\mathcal{F}| \leq 2^{n-1}$, since for every set $F \subset [n]$, we can have at most one of F and $[n] \setminus F$ in \mathcal{F} . This bound is easily seen to be tight, and there are in fact numerous extremal systems. For example, one could take the set system consisting of all sets containing some fixed $i \in [n]$. Another construction is to take all sets $F \subset [n]$ of size $|F| > \frac{n}{2}$. If n is odd, this consists of precisely 2^{n-1} sets. If n is even, then we must add an intersecting system of sets of size $\frac{n}{2}$; for instance, $\{F \subset [n] : |F| = \frac{n}{2}, 1 \in F\}$ would suffice.

In some sense, having large sets makes it easier for the system to be intersecting. This leads to the classic theorem of Erdős–Ko–Rado [14], a central result in extremal set theory, which bounds the size of an intersecting set system with all sets restricted to have size k . Here we use $\binom{[n]}{k}$ to denote all subsets of $[n]$ of size k .

Theorem 1.1 (Erdős–Ko–Rado [14], 1961). *If $n \geq 2k$, and $\mathcal{F} \subset \binom{[n]}{k}$ is an intersecting set system, then $|\mathcal{F}| \leq \binom{n-1}{k-1}$.*

This is again tight, as we may take all sets containing some fixed element $i \in [n]$, a system we call a *(full) star with center i* .

As is befitting of such an important theorem, there have been numerous extensions to many different settings, some of which are discussed in Ander-

son's book [6]. We are particularly interested in two, namely t -intersecting systems and q -matching-free systems.

A pair of sets F_1, F_2 is said to be t -intersecting if $|F_1 \cap F_2| \geq t$, and t -disjoint otherwise. A set system \mathcal{F} is t -intersecting if every pair of sets in the system is. When $t=1$, we simply have an intersecting system. A natural construction of a t -intersecting system is to fix some t -set $X \in \binom{[n]}{t}$, and take all k -sets containing X ; we call this a (full) t -star with center X . In their original paper, Erdős–Ko–Rado showed that, provided n was sufficiently large, this was best possible.

Theorem 1.2 (Erdős–Ko–Rado [14], 1961). *If $n \geq n_0(k, t)$, and $\mathcal{F} \subset \binom{[n]}{k}$ is a t -intersecting set system, then $|\mathcal{F}| \leq \binom{n-t}{k-t}$.*

There was much work done on determining the correct value of $n_0(k, t)$, and how large t -intersecting systems can be when n is small. This problem was completely resolved by the celebrated Complete Intersection Theorem of Ahlswede and Khachatrian [5] in 1997.

The second extension we shall consider concerns matchings. A q -matching is a collection of q pairwise-disjoint sets. A set system is therefore intersecting if and only if it does not contain a 2-matching. As an extension of the Erdős–Ko–Rado Theorem, Erdős asked how large a q -matching-free k -uniform set system could be, and in [13] showed that when n is large, the best construction consists of taking all sets meeting $[q-1]$. He further conjectured what the solution should be for small n , and this remains an open problem of great interest. For recent results on this conjecture, see, e.g., [16,17,19,21].

1.2. Beyond the thresholds

The preceding results are all examples of the typical extremal problem, which asks how large a structure can be without containing a forbidden configuration. In this paper, we study their Erdős–Rademacher variants, a name we now explain.

Arguably the most well-known result in extremal combinatorics is a theorem of Mantel [22] from 1907, which states that an n -vertex triangle-free graph can have at most $\lfloor \frac{n^2}{4} \rfloor$ edges. In an unpublished result, Rademacher strengthened this theorem by showing that any graph with $\lfloor \frac{n^2}{4} \rfloor + 1$ edges must contain at least $\lfloor \frac{n}{2} \rfloor$ triangles. In [11] and [12], Erdős extended this first to graphs with a linear number of extra edges, and then to cliques larger than triangles. More generally, for any extremal problem, the corresponding

Erdős–Rademacher problem asks how many copies of the forbidden configuration must appear in a structure larger than the extremal bound.

In the context of intersecting systems, the Erdős–Rademacher question was first investigated by Frankl [15] and, independently, Ahlswede [1] some forty years ago, who showed that the number of disjoint pairs of sets in a set system is minimized by taking the sets to be as large as possible.

Theorem 1.3 (Frankl [15], 1977; Ahlswede [1], 1980). *If $\sum_{i=k+1}^n \binom{n}{i} \leq s \leq \sum_{i=k}^n \binom{n}{i}$, then the minimum number of disjoint pairs in a set system of size s is attained by some system \mathcal{F} with $\cup_{i>k} \binom{[n]}{i} \subseteq \mathcal{F} \subseteq \cup_{i \geq k} \binom{[n]}{i}$.*

Note that while this theorem provides the large-scale structure of extremal systems, it does not determine exactly which systems are optimal. Since we have $\cup_{i>k} \binom{[n]}{i} \subset \mathcal{F}$, each set of size k contributes the same number of disjoint pairs with larger sets. Hence, the total number of disjoint pairs is minimized by minimizing the number of disjoint pairs between the sets of size k , a problem raised by Ahlswede.

Question 1.4 (Ahlswede [1], 1980). *Given $0 \leq s \leq \binom{n}{k}$, which k -uniform set systems $\mathcal{F} \subset \binom{[n]}{k}$ with $|\mathcal{F}| = s$ minimize the number of disjoint pairs?*

By the Erdős–Ko–Rado Theorem, we know that when $s \leq \binom{n-1}{k-1}$, we need not have any disjoint pairs, while for $s > \binom{n-1}{k-1}$, there must be at least one disjoint pair. This question can thus be thought of as the Erdős–Rademacher problem for the Erdős–Ko–Rado Theorem.

This question is also deeply connected to the Kneser graph. The Kneser graph $K(n, k)$ has vertices $V = \binom{[n]}{k}$, with vertices X and Y adjacent if and only if the sets are disjoint. An intersecting set system corresponds to an independent set in the Kneser graph. Question 1.4 is thus asking which s vertices of the Kneser graph induce the smallest number of edges. Since the Kneser graph is regular, this is equivalent to the edge-isoperimetric problem of finding the largest bipartite subgraph of $K(n, r)$ with one part of size s . Kneser graphs have been extensively studied, and the problem of determining their largest bipartite subgraphs was first raised by Poljak and Tuza in [23].

In 2003, Bollobás and Leader [7] presented a new proof of Theorem 1.3, by relaxing the problem to a continuous version and analyzing fractional set systems. They further considered Question 1.4, and conjectured that for small systems, the initial segment of the lexicographical ordering on $\binom{[n]}{k}$ should be optimal. In the lexicographical ordering, we say $A < B$ if $\min(A \triangle B) \in A$; that is, we prefer sets with smaller elements. More generally,

Bollobás and Leader conjectured that all the extremal systems should take the form of what they named ℓ -balls, as defined below. Note that a 1-ball is an initial segment of the lexicographical ordering.

Conjecture 1.5 (Bollobás–Leader [7], 2003). *One of the systems $\mathcal{A}_{r,\ell} = \{A \in \binom{[n]}{k} : |A \cap [r]| \geq \ell\}$ minimizes the number of disjoint pairs.*

When $k=2$, we can think of a set system in $\binom{[n]}{2}$ as a graph on n vertices, and are then asking for which graphs of a given size minimize the number of disjoint pairs of edges. This problem was solved by Ahlswede and Katona [4] in 1978, who showed that the extremal graphs were always either the union of stars (a collection of vertices connected to all the other vertices), or their complement. In that paper, they asked for a different generalization, namely which k -uniform set systems minimize the number of $(k-1)$ -disjoint pairs. This is the Erdős–Rademacher problem for t -intersecting systems when $t = k-1$, and is known as the Kleitman–West problem, which shares some connections to information theory. An exact solution appears difficult to obtain, and a natural conjecture of Kleitman for this problem has been proven to be untrue by Ahlswede and Cai [2]. They later solved a continuous relaxation of the $k=3$ case in [3], marking the furthest progress made on this problem.

1.3. Our results

Our main result verifies Conjecture 1.5 for small systems, showing that initial segments of the lexicographical ordering minimize the number of disjoint pairs. We denote by $\mathcal{L}_{n,k}(s)$ the first s sets in the lexicographical ordering on $\binom{[n]}{k}$. Note that the size of ℓ full stars, say with centers $\{1, 2, \dots, \ell\}$, is $\binom{n}{k} - \binom{n-\ell}{k}$. The following theorem shows that provided n is large enough with respect to k and ℓ , it is optimal to take sets from the first ℓ stars.

Theorem 1.6. *Provided $n > 108k^2\ell(k + \ell)$ and $0 \leq s \leq \binom{n}{k} - \binom{n-\ell}{k}$, $\mathcal{L}_{n,k}(s)$ minimizes the number of disjoint pairs among all systems of s sets in $\binom{[n]}{k}$.*

As a by-product of our proof, we shall obtain a characterization of the extremal systems in this range, which we provide in Proposition 2.2. Corollary 2.4 shows that we can also use Theorem 1.6 to determine which systems are optimal when s is very close to $\binom{n}{k}$.

We further show that $\mathcal{L}_{n,k}(s)$ also minimizes the number of q -matchings.

Theorem 1.7. *Provided $n > n_1(k, q, \ell)$ and $0 \leq s \leq \binom{n}{k} - \binom{n-\ell}{k}$, $\mathcal{L}_{n,k}(s)$ minimizes the number of q -matchings among all systems of s sets in $\binom{[n]}{k}$.*

Finally, we extend our methods to determine which systems minimize the number of t -disjoint pairs. When $t=k-1$, this provides a partial solution to the problem of Kleitman and West. When n is large with respect to k, t and ℓ , an extremal system is contained in the union of ℓ full t -stars. As we shall discuss in Section 4, not all such unions are isomorphic, and once again it is the lexicographical ordering that is optimal.

Theorem 1.8. *Provided $n \geq n_2(k, t, \ell)$ and $0 \leq s \leq \binom{n-t+1}{k-t+1} - \binom{n-t-\ell+1}{k-t+1}$, $\mathcal{L}_{n,k}(s)$ minimizes the number of t -disjoint pairs among all systems of s sets in $\binom{[n]}{k}$.*

We again characterize all extremal systems in Corollary 4.5, while Corollary 4.6 provides the solution when s very close to $\binom{n}{k}$.

1.4. Outline and notation

The remainder of the paper is organized as follows. In Section 2 we study disjoint pairs, and prove Theorem 1.6. In Section 3, we consider the number of q -matchings, and prove Theorem 1.7. In Section 4, we extend our results to t -disjoint pairs, proving Theorem 1.8. In the final section we present some concluding remarks and open problems.

We denote by $[n]$ the set of the first n positive integers, and use this as the ground set for our set systems. Given a set X , $\binom{X}{k}$ is the system of all k -subsets of X . The number of disjoint pairs between two systems \mathcal{F} and \mathcal{G} is given by $\text{dp}(\mathcal{F}, \mathcal{G}) = |\{(F, G) \in \mathcal{F} \times \mathcal{G} : F \cap G = \emptyset\}|$, and the number of disjoint pairs within a system \mathcal{F} is denoted by $\text{dp}(\mathcal{F}) = \frac{1}{2} \text{dp}(\mathcal{F}, \mathcal{F})$.

For given n, k and s , we let $\text{dp}(n, k, s)$ denote the minimum of $\text{dp}(\mathcal{F})$ over all k -uniform set systems on $[n]$ of size s . We define $\text{dp}^{(q)}(\mathcal{F})$ and $\text{dp}^{(q)}(n, k, s)$ similarly for the number of q -matchings in set systems, and $\text{dp}_t(\mathcal{F}), \text{dp}_t(\mathcal{F}, \mathcal{G})$, and $\text{dp}_t(n, k, s)$ for the number of t -disjoint pairs.

Given any set system \mathcal{F} , and a set $X \subset [n]$, we let $\mathcal{F}(X) = \{F \in \mathcal{F} : X \subset F\}$ be those sets in the system containing X . If X is a singleton, we shall drop the set notation, and write $\mathcal{F}(x)$. Finally, we define a *cover* of a system to be a set X with $\cup_{x \in X} \mathcal{F}(x) = \mathcal{F}$; that is, a set of elements that touches every set. A t -*cover* is a collection of t -sets such that every set in the system contains one of the t -sets.

2. Disjoint Pairs

In this section we will show that for small systems, initial segments of the lexicographical ordering, $\mathcal{L}_{n,k}(s)$, minimize the number of disjoint pairs. Note that when $s \leq \binom{n-1}{k-1}$, $\mathcal{L}_{n,k}(s)$ is a star, which is an intersecting system and thus clearly optimal. The following result of Katona et al [20] shows that if we add one set to a full star, the resulting system will also be optimal.

Proposition 2.1. *Suppose $n \geq 2k$. Any system $\mathcal{F} \subset \binom{[n]}{k}$ with $|\mathcal{F}| = \binom{n-1}{k-1} + 1$ contains at least $\binom{n-k-1}{k-1}$ disjoint pairs.*

Our first theorem shows that as we add sets to the system, we should try to cover our system with as few stars as possible, as is achieved by $\mathcal{L}_{n,k}(s)$. Later, in Proposition 2.2, we shall precisely characterize all extremal systems. We begin by recalling the statement of the theorem.

Theorem 1.6. *Provided $n > 108k^2\ell(k + \ell)$ and $0 \leq s \leq \binom{n}{k} - \binom{n-\ell}{k}$, $\mathcal{L}_{n,k}(s)$ minimizes the number of disjoint pairs among all systems of s sets in $\binom{[n]}{k}$.*

In our notation, the above theorem gives $dp(n, k, s) = dp(\mathcal{L}_{n,k}(s))$ for such values of s . Let $1 \leq r \leq \ell$ be such that $\binom{n}{k} - \binom{n-r+1}{k} < s \leq \binom{n}{k} - \binom{n-r}{k}$. Since $\mathcal{L}_{n,k}(s)$ contains all sets meeting $\{1, 2, \dots, r-1\}$, and all the remaining sets contain r , we can provide an explicit formula for the minimum number of disjoint pairs.

$$\begin{aligned} dp(n, k, s) &= dp(\mathcal{L}_{n,k}(s)) \\ &= \sum_{i=1}^r \sum_{\substack{F \in \mathcal{L}_{n,k}(s) \\ \min F = i}} |\{G \in \mathcal{L}_{n,k}(s) : \min G < i, F \cap G = \emptyset\}| \\ &= \sum_{i=2}^{r-1} \binom{n-i}{k-1} \sum_{j=1}^{i-1} \binom{n-j-k}{k-1} \\ &\quad + \left(s - \sum_{i=1}^{r-1} \binom{n-i}{k-1} \right) \sum_{j=1}^{r-1} \binom{n-j-k}{k-1}. \end{aligned}$$

It will be useful to have a simpler upper bound on $dp(n, k, s)$. Note that we can assign each set in $\mathcal{L}_{n,k}(s)$ to an element of $[r]$ it contains. It can then only be disjoint from sets assigned to different elements. In the worst case, an equal number of sets is assigned to each element, giving the bound

$$(1) \quad dp(n, k, s) \leq \binom{r}{2} \left(\frac{s}{r} \right)^2 = \frac{1}{2} \left(1 - \frac{1}{r} \right) s^2.$$

We shall often require bounds on $\binom{n-2}{k-2}$ in terms of s . Since $\mathcal{L}_{n,k}(s)$ contains all sets meeting $[r-1]$ and n is large, the Bonferroni Inequalities give

$$\begin{aligned}
 s = |\mathcal{L}_{n,k}(s)| &\geq (r-1) \binom{n-1}{k-1} - \binom{r-1}{2} \binom{n-2}{k-2} \\
 (2) \qquad &= \left(\frac{(r-1)(n-1)}{k-1} - \binom{r-1}{2} \right) \binom{n-2}{k-2} \geq \frac{rn}{3k} \binom{n-2}{k-2}.
 \end{aligned}$$

Our proof of Theorem 1.6 will proceed according to the following steps. First we shall argue that if a family \mathcal{F} has at most $\frac{1}{2} \left(1 - \frac{1}{r}\right) s^2$ disjoint pairs, then it must contain a popular element; that is, some $x \in [n]$ contained in many sets of \mathcal{F} . The second step consists of a series of arguments to show that \mathcal{F} can be covered by r elements. The final step will then show that among all families that can be covered by r elements, $\mathcal{L}_{n,k}(s)$ minimizes the number of disjoint pairs.

Proof of Theorem 1.6. We prove the theorem by induction on n and s . For the base case, suppose $0 \leq s \leq \binom{n}{k} - \binom{n-1}{k} = \binom{n-1}{k-1}$. In this range, $\mathcal{L}_{n,k}(s)$ is a star, consisting only of sets containing 1, and thus obviously minimizes the number of disjoint pairs.

For the induction step, let \mathcal{F} be an extremal family with $|\mathcal{F}| = s > \binom{n-1}{k-1}$, and so $r \geq 2$. Suppose first that \mathcal{F} contains a full star. Without loss of generality, we may assume \mathcal{F} has all sets containing 1, so $\mathcal{F}(1) = \left\{ F \in \binom{[n]}{k} : 1 \in F \right\}$. Since $\mathcal{F}(1)$ is an intersecting family, we have $\text{dp}(\mathcal{F}) = \text{dp}(\mathcal{F}(1), \mathcal{F} \setminus \mathcal{F}(1)) + \text{dp}(\mathcal{F} \setminus \mathcal{F}(1))$. Now any set $F \in \mathcal{F}$ with $1 \notin F$ is disjoint from exactly $\binom{n-k-1}{k-1}$ sets in $\mathcal{F}(1)$, giving $\text{dp}(\mathcal{F}(1), \mathcal{F} \setminus \mathcal{F}(1)) = |\mathcal{F} \setminus \mathcal{F}(1)| \binom{n-k-1}{k-1} = \left(s - \binom{n-1}{k-1} \right) \binom{n-k-1}{k-1}$, regardless of the structure of $\mathcal{F} \setminus \mathcal{F}(1)$. Since $\mathcal{F} \setminus \mathcal{F}(1)$ is a family of $s - \binom{n-1}{k-1}$ sets in $[n] \setminus \{1\}$, our induction hypothesis implies $\text{dp}(\mathcal{F} \setminus \mathcal{F}(1))$ is minimized by the initial segment of the lexicographical order. Since $\mathcal{L}_{n,k}(s)$ consists of all sets containing 1, and the initial segment of the lexicographical order on $[n] \setminus \{1\}$, it follows that $\mathcal{L}_{n,k}(s)$ is optimal, as claimed.

Hence, we may assume that \mathcal{F} does not contain any full star. Consequently, given any $F \in \mathcal{F}$ and $x \in [n]$, we may replace F by a set containing x .

Step 1: Show there exists some $x \in [n]$ with $|\mathcal{F}(x)| \geq \frac{s}{3r}$.

We begin by showing there cannot be too many moderately popular elements.

Claim 1. $\left| \left\{ x : |\mathcal{F}(x)| \geq \frac{s}{3kr} \right\} \right| < 6kr$.

Proof. Suppose not, and consider $X \subset \{x: |\mathcal{F}(x)| \geq \frac{s}{3kr}\}$ with $|X| = 6kr$. Using (2), we have

$$\begin{aligned} s = |\mathcal{F}| &\geq |\cup_{x \in X} \mathcal{F}(x)| \geq \sum_{x \in X} |\mathcal{F}(x)| - \sum_{x, y \in X} |\mathcal{F}(x) \cap \mathcal{F}(y)| \\ &\geq |X| \cdot \frac{s}{3kr} - \binom{|X|}{2} \binom{n-2}{k-2} \geq 2s - 18k^2r^2 \cdot \frac{3k}{nr} s = \left(2 - \frac{54k^3r}{n}\right) s. \end{aligned}$$

Since $n > 54k^3r$, we reach a contradiction. ■

We now show the existence of a popular element.

Claim 2. *There is some $x \in [n]$ with $|\mathcal{F}(x)| > \frac{s}{3r}$.*

Proof. Since \mathcal{F} is extremal, we must have

$$\frac{1}{2} \left(1 - \frac{1}{r}\right) s^2 \geq \text{dp}(\mathcal{L}_{n,k}(s)) \geq \text{dp}(\mathcal{F}) = \frac{1}{2} \sum_{F \in \mathcal{F}} \text{dp}(F, \mathcal{F}).$$

Now $\text{dp}(F, \mathcal{F}) = s - |\cup_{x \in F} \mathcal{F}(x)| \geq s - \sum_{x \in F} |\mathcal{F}(x)|$, and so we have

$$\begin{aligned} \left(1 - \frac{1}{r}\right) s^2 &\geq \sum_{F \in \mathcal{F}} \left(s - \sum_{x \in F} |\mathcal{F}(x)|\right) = s^2 - \sum_{F \in \mathcal{F}} \sum_{x \in F} |\mathcal{F}(x)| \\ &= s^2 - \sum_x |\mathcal{F}(x)|^2. \end{aligned}$$

Let $X = \{x: |\mathcal{F}(x)| \geq \frac{s}{3kr}\}$, and note that by the previous claim, $|X| < 6kr$. Moreover, without loss of generality, suppose 1 is the most popular element, so $|\mathcal{F}(x)| \leq |\mathcal{F}(1)|$ for all x . We split the above sum into those $x \in X$ and those $x \notin X$, giving

$$(3) \quad \frac{s^2}{r} \leq \sum_{x \in X} |\mathcal{F}(x)|^2 + \sum_{x \notin X} |\mathcal{F}(x)|^2 \leq |\mathcal{F}(1)| \sum_{x \in X} |\mathcal{F}(x)| + \frac{s}{3kr} \sum_{x \notin X} |\mathcal{F}(x)|.$$

We bound the first sum by noting that

$$\begin{aligned} \sum_{x \in X} |\mathcal{F}(x)| &\leq |\cup_{x \in X} \mathcal{F}(x)| + \sum_{\{x, y\} \subset X} |\mathcal{F}(x) \cap \mathcal{F}(y)| \leq s + \binom{|X|}{2} \binom{n-2}{k-2} \\ &\leq \left(1 + \frac{54k^3r}{n}\right) s \leq 2s, \end{aligned}$$

using (2) and our bound on n . The second sum is bounded by $\sum_{x \notin X} |\mathcal{F}(x)| \leq \sum_x |\mathcal{F}(x)| = ks$. Substituting these bounds in (3) gives $\frac{s^2}{r} \leq 2|\mathcal{F}(1)|s + \frac{s^2}{3r}$, and so $|\mathcal{F}(1)| \geq \frac{s}{3r}$, as required. ■

This concludes Step 1.

Step 2: Show there is a cover of size r .

We begin by using the existence of a popular element to argue that there is a reasonably small cover, and then provide a number of claims that together imply an extremal family must in fact be covered by r elements.

Claim 3. $X = \{x : |\mathcal{F}(x)| \geq \frac{s}{3kr}\}$ is a cover for \mathcal{F} .

Proof. Suppose for contradiction X is not a cover. Then there must be some set $F \in \mathcal{F}$ with $F \cap X = \emptyset$, and so $|\mathcal{F}(x)| < \frac{s}{3kr}$ for all $x \in F$. Hence $\text{dp}(F, \mathcal{F}) = s - |\cup_{x \in F} \mathcal{F}(x)| \geq s - \sum_{x \in F} |\mathcal{F}(x)| > s - \frac{s}{3r}$. On the other hand, by Claim 2, we may assume $|\mathcal{F}(1)| \geq \frac{s}{3r}$. Thus if G is any set containing 1, we have $\text{dp}(G, \mathcal{F}) \leq s - |\mathcal{F}(1)| = s - \frac{s}{3r}$. Hence, replacing F with such a set G , which is possible since $\mathcal{F}(1)$ is not a full star, would decrease the number of disjoint pairs in \mathcal{F} , contradicting its optimality.

Hence X must be a cover for \mathcal{F} , as claimed. ■

By Claim 1, we have $|X| \leq 6kr$. Take a minimal subcover of X containing 1; without loss of generality, we may assume this subcover is $[m]$, for some $r \leq m \leq 6kr$. We shall now proceed to show that an extremal family must have $m=r$, giving rise to the smallest possible cover.

Rather than working with the subfamilies $\mathcal{F}(i)$, $i \in [m]$, we shall avoid double-counting by instead considering the subsystems $\mathcal{F}^*(i) = \{F \in \mathcal{F} : \min F = i\}$. Note that the systems $\mathcal{F}^*(i)$ partition \mathcal{F} .

Claim 4. For every $i, j \in [m]$, we have $|\mathcal{F}^*(i)| \geq |\mathcal{F}^*(j)| - \frac{3mk^2}{rn} s$.

Proof. First we claim that there is some $F \in \mathcal{F}$ with $F \cap [m] = \{i\}$, which in particular implies $F \in \mathcal{F}^*(i)$. Indeed, the number of sets in $\mathcal{F}(i)$ containing another element in $[m]$ is less than $m \binom{n-2}{k-2} \leq \frac{3mk}{rn} s \leq \frac{18k^2}{n} s < \frac{s}{3kr}$. However, since $[m]$ is a subcover of X from Claim 3, it follows that $|\mathcal{F}(i)| \geq \frac{s}{3kr}$, and thus we must have our desired set $F \in \mathcal{F}^*(i)$.

For any $j \in [m] \setminus \{i\}$, F can intersect at most $k \binom{n-2}{k-2}$ sets in $\mathcal{F}(j)$, since each of these sets must contain both j and one element from F . Summing over all j and using (2) gives

$$\begin{aligned} \text{dp}(F, \mathcal{F}) &= \sum_{j \neq i} \text{dp}(F, \mathcal{F}^*(j)) \geq \sum_{j \neq i} \left[|\mathcal{F}^*(j)| - k \binom{n-2}{k-2} \right] \\ &= s - |\mathcal{F}^*(i)| - (m-1)k \binom{n-2}{k-2} \geq s - |\mathcal{F}^*(i)| - \frac{3mk^2}{rn} s. \end{aligned}$$

On the other hand, if we were to replace F with a set containing j , it would intersect at least those sets in $\mathcal{F}^*(j)$, and so introduce at most

$s - |\mathcal{F}^*(j)|$ disjoint pairs. Since \mathcal{F} is an extremal family, we must have $s - |\mathcal{F}^*(j)| \geq s - |\mathcal{F}^*(i)| - \frac{3mk^2}{rn}s$, or $|\mathcal{F}^*(i)| \geq |\mathcal{F}^*(j)| - \frac{3mk^2}{rn}s$, as required. ■

Claim 5. $m \leq 6r$.

Proof. We shall now bound $|\mathcal{F}^*(i)|$ by taking $j=1$ in Claim 4. Recall that by Claim 2 we have $|\mathcal{F}(1)| = |\mathcal{F}^*(1)| \geq \frac{s}{3r}$, and from Claim 1 it follows that $m \leq 6kr$. Since $n > 108k^3r$, these bounds give

$$|\mathcal{F}^*(i)| \geq |\mathcal{F}^*(1)| - \frac{3mk^2}{rn}s \geq \frac{s}{3r} - \frac{18k^3}{n}s \geq \frac{s}{6r}.$$

Since $s = |\cup_{i=1}^m \mathcal{F}^*(i)| = \sum_{i=1}^m |\mathcal{F}^*(i)| \geq m \cdot \frac{s}{6r}$, we must have $m \leq 6r$. ■

With this tighter bound on m , we are now able to better estimate the number of disjoint pairs in \mathcal{F} , and in doing so show that we must actually have $m=r$ if \mathcal{F} is extremal.

Claim 6. *If \mathcal{F} minimizes the number of disjoint pairs, then \mathcal{F} can be covered by r elements.*

Proof. Since $\{\mathcal{F}^*(i)\}$ partitions \mathcal{F} into intersecting families, we have $\text{dp}(\mathcal{F}) = \sum_{i < j} \text{dp}(\mathcal{F}^*(i), \mathcal{F}^*(j))$. For $i < j$, note that every set $F \in \mathcal{F}^*(j)$ can intersect at most $k \binom{n-2}{k-2}$ sets in $\mathcal{F}^*(i)$, since those sets would have to contain one element from F as well as i . This shows that $\text{dp}(\mathcal{F}^*(i), \mathcal{F}^*(j)) \geq (|\mathcal{F}^*(i)| - k \binom{n-2}{k-2}) |\mathcal{F}^*(j)|$. Moreover, note that Claim 4 implies the bound $|\mathcal{F}^*(1)| \leq \frac{s}{m} + \frac{3mk^2}{rn}s$, since we must have some $i \in [m]$ with $|\mathcal{F}^*(i)| \leq \frac{1}{m} \sum_{j=1}^m |\mathcal{F}^*(j)| = \frac{s}{m}$, and $|\mathcal{F}^*(i)| \geq |\mathcal{F}^*(1)| - \frac{3mk^2}{rn}s$. Thus

$$\begin{aligned} \text{dp}(\mathcal{F}) &\geq \sum_{i < j} \left(|\mathcal{F}^*(i)| - k \binom{n-2}{k-2} \right) |\mathcal{F}^*(j)| \\ &= \sum_{i < j} |\mathcal{F}^*(i)| |\mathcal{F}^*(j)| - k \binom{n-2}{k-2} \sum_j (j-1) |\mathcal{F}^*(j)| \\ &\geq \frac{1}{2} \left(\left(\sum_i |\mathcal{F}^*(i)| \right)^2 - \sum_i |\mathcal{F}^*(i)|^2 \right) - mk \binom{n-2}{k-2} \sum_j |\mathcal{F}^*(j)| \\ &\geq \frac{1}{2} \left(s^2 - |\mathcal{F}^*(1)| \sum_i |\mathcal{F}^*(i)| \right) - mk \binom{n-2}{k-2} \sum_j |\mathcal{F}^*(j)| \\ &\geq \frac{1}{2} \left(s^2 - \left(\frac{s}{m} + \frac{3mk^2}{rn}s \right) s \right) - \frac{3mk^2}{rn}s^2 = \frac{1}{2} \left(1 - \frac{1}{m} - \frac{9mk^2}{rn} \right) s^2. \end{aligned}$$

On the other hand, since \mathcal{F} is extremal, we have $\text{dp}(\mathcal{F}) \leq \text{dp}(\mathcal{L}_{n,k}(s)) \leq \frac{1}{2} \left(1 - \frac{1}{r}\right) s^2$, and so we must have $\frac{1}{r} \leq \frac{1}{m} + \frac{9mk^2}{rn} \leq \frac{1}{m} + \frac{54k^2}{n}$. Since $n > 54k^2r(k+r) \geq 54k^2r(r+1)$, we have $\frac{54k^2}{n} < \frac{1}{r} - \frac{1}{r+1}$, and hence we require $m \leq r$. Thus \mathcal{F} can be covered by r elements. ■

This completes Step 2.

Step 3: Show that $\mathcal{L}_{n,k}(s)$ is optimal.

We will now complete the induction argument by showing that $\mathcal{L}_{n,k}(s)$ is indeed an extremal family. From the preceding steps we know \mathcal{F} must be covered by r elements, which we may assume to be $[r]$. We shall now use a complementarity argument to deduce the optimality of $\mathcal{L}_{n,k}(s)$.

Let $\mathcal{A} = \left\{A \in \binom{[n]}{k} : A \cap [r] \neq \emptyset\right\}$ be all sets meeting $[r]$, so we have $\mathcal{F} \subset \mathcal{A}$. Let $\mathcal{G} = \mathcal{A} \setminus \mathcal{F}$. We have

$$\text{dp}(\mathcal{F}) = \text{dp}(\mathcal{A}) - \text{dp}(\mathcal{G}, \mathcal{A}) + \text{dp}(\mathcal{G}),$$

since only disjoint pairs contained in \mathcal{F} survive on the right-hand side.

Since $\text{dp}(\mathcal{A})$ is determined solely by r , and hence s , but is independent of the structure of \mathcal{F} , we may treat that term as a constant.

We have $\text{dp}(\mathcal{G}, \mathcal{A}) = \sum_{G \in \mathcal{G}} \text{dp}(G, \mathcal{A})$. For any G , $\text{dp}(G, \mathcal{A})$ is determined by $|G \cap [r]|$, and is maximized when $|G \cap [r]| = 1$. For $\mathcal{F} = \mathcal{L}_{n,k}(s)$, we have $G \cap [r] = \{r\}$ for all $G \in \mathcal{G}$, and so $\mathcal{L}_{n,k}(s)$ maximizes $\text{dp}(\mathcal{G}, \mathcal{A})$.

Finally, we obviously have $\text{dp}(\mathcal{G}) \geq 0$, with equality in the case of $\mathcal{F} = \mathcal{L}_{n,k}(s)$.

Hence, it follows that $\mathcal{L}_{n,k}(s)$ minimizes the number of disjoint pairs, completing the proof. ■

This proof also allows us to characterize all extremal systems.

Proposition 2.2. *Provided $n > 108k^2r(k+r)$ and $\binom{n}{k} - \binom{n-r+1}{k} \leq s \leq \binom{n}{k} - \binom{n-r}{k}$, then a set system $\mathcal{F} \subset \binom{[n]}{k}$ of size s minimizes the number of disjoint pairs if and only if it has one of the two following structures:*

- (i) \mathcal{F} contains $r - 1$ full stars, with the remaining sets forming an intersecting system, or
- (ii) \mathcal{F} has a cover X of size r , and if $\mathcal{G} = \left\{G \in \binom{[n]}{k} \setminus \mathcal{F} : G \cap X \neq \emptyset\right\}$, then \mathcal{G} is intersecting, and $|G \cap X| = 1$ for all $G \in \mathcal{G}$.

Proof. We prove the proposition by induction on n and s . If $0 \leq s \leq \binom{n-1}{k-1}$, then clearly a system is extremal if and only if it is intersecting, as there need not be any disjoint pairs. Since $r = 1$ for this value of s , this is covered by case (i).

For the induction step, note that if \mathcal{F} is extremal and contains a full star, say $\mathcal{F}(1)$, then $\mathcal{F} \setminus \mathcal{F}(1)$ must also be extremal. Applying the induction hypothesis gives the result, since adding a full star to either (i) or (ii) preserves the structure.

Hence, we may assume there is no full star. Claim 6 then shows that \mathcal{F} has a cover of size r , while the complementarity argument from Step 3 gives the above characterization of the system \mathcal{G} . ■

Finally, we use Theorem 1.6 to determine which large systems minimize the number of disjoint pairs. Recall that, as explained in Section 1, we can view this problem as an edge-isoperimetric for the Kneser graph, which is regular. The following lemma links the edge-isoperimetric problem for small and large vertex sets in regular graphs.

Lemma 2.3. *Let $G = (V, E)$ be a regular graph on n vertices. Then $S \subset V$ minimizes the number of edges $e(S)$ over all sets of $|S|$ vertices if and only if $V \setminus S$ minimizes the number of edges over all sets of $n - |S|$ vertices.*

Proof. Let $s = |S|$, and suppose G is d -regular. Summing the degrees of vertices in S , we have

$$2e(S) = ds - e(S, V \setminus S) = ds - (d(n - s) - 2e(V \setminus S)) = d(2s - n) + 2e(V \setminus S),$$

and so $e(S)$ is minimized if and only if $e(V \setminus S)$ is. ■

The following corollary, which is a direct consequence of Theorem 1.6 and Lemma 2.3, shows that the complement of the lexicographical initial segments are optimal when s is close to $\binom{n}{k}$.

Corollary 2.4. *Provided $n > 108k^2\ell(k + \ell)$ and $\binom{n-\ell}{k} \leq s \leq \binom{n}{k}$, $\binom{[n]}{k} \setminus \mathcal{L}_{n,k}(\binom{n}{k} - s)$ minimizes the number of disjoint pairs.*

3. q -matchings

In this section, we determine which set systems minimize the number of q -matchings. This extends Theorem 1.6, which is the case $q = 2$. Note that when $|\mathcal{F}| = s \leq \binom{n}{k} - \binom{n-q+1}{k}$, the lexicographical initial segment does not contain any q -matchings, as all sets meet $[q - 1]$. Indeed, this is known to be the largest such family when $n > (2q - 1)k - q$, as proven by Frankl [16]. We shall show that, provided n is suitably large, $\mathcal{L}_{n,k}(s)$ continues to be optimal for families of size up to $\binom{n}{k} - \binom{n-\ell}{k}$. Unlike for Theorem 1.6, we have made no attempt to optimize the dependence of n on the other

parameters. We provide our calculations in asymptotic notation for ease of presentation, where we fix the parameters k, ℓ and q to be constant and let $n \rightarrow \infty$. However, our result should certainly hold for $n > C\ell^2 k^5 (\ell^2 + k^2) e^{3q}$.

Our proof strategy will be very similar to before: we will first find a popular element, deduce the existence of a smallest possible cover, and then use a complementarity argument to show that the initial segment of the lexicographical order is optimal. The main difference is in the definition of *popular* - rather than considering how many sets contain the element x , we shall be concerned with how many $(q - 1)$ -matchings have a set containing x . To this end, we introduce some new notation. Given a set family \mathcal{F} , and a set F , let $\mathcal{F}^{(q)}(F)$ denote the number of q -matchings $\{F_1, F_2, \dots, F_q\}$ in \mathcal{F} with $\cup_{i=1}^q F_i \cap F \neq \emptyset$. Similarly, for some $x \in [n]$, we let $\mathcal{F}^{(q)}(x) = \mathcal{F}^{(q)}(\{x\})$ be the number of q -matchings with $x \in \cup_{i=1}^q F_i$.

Theorem 1.7. *Provided $n > n_1(k, q, \ell)$ and $0 \leq s \leq \binom{n}{k} - \binom{n-\ell}{k}$, $\mathcal{L}_{n,k}(s)$ minimizes the number of q -matchings among all systems of s sets in $\binom{[n]}{k}$.*

As before, we start with some estimates on $\text{dp}^{(q)}(\mathcal{L}_{n,k}(s))$. Let r be such that $\binom{n}{k} - \binom{n-r+1}{k} < s \leq \binom{n}{k} - \binom{n-r}{k}$. We may assign each set in $\mathcal{L}_{n,k}(s)$ to one of its elements in $[r]$. Note that a q -matching cannot contain two sets assigned to the same element, and so to obtain a q -matching, we must choose sets from different elements in $[r]$. By convexity, the worst case is when the sets are equally distributed over $[r]$, giving the upper bound

$$(4) \quad \text{dp}^{(q)}(\mathcal{L}_{n,k}(s)) \leq \binom{r}{q} \left(\frac{s}{r}\right)^q.$$

In this case we shall also require a lower bound. Note that $\mathcal{L}_{n,k}(s)$ contains all sets meeting $[r - 1]$, with the remaining sets containing $\{r\}$; suppose there are $\alpha \binom{n-1}{k-1}$ such sets. Note that we have $s = \binom{n}{k} - \binom{n-r+1}{k} + \alpha \binom{n-1}{k-1} = (r - 1) \binom{n-1}{k-1} + \alpha \binom{n-1}{k-1}$, so $\binom{n-1}{k-1} \geq \frac{s}{r-1+\alpha}$.

We shall consider two types of q -matchings - those with one of the $\alpha \binom{n-1}{k-1}$ sets that only meet $[r]$ at r , and those without. For the first type, we have $\alpha \binom{n-1}{k-1}$ choices for the set containing r . For the remaining sets in the q -matching, we will avoid any overcounting by restricting ourselves to sets that only contain one element from $[r - 1]$. We can then make one of $\binom{r-1}{q-1}$ choices for how the remaining $q - 1$ sets will meet $[r - 1]$. For each such set, we must avoid all other elements in $[r]$ and all previously used elements, leaving us with at least $\binom{n-kq-r}{k-1} = (1 - o(1)) \binom{n-1}{k-1}$ options.

For the second type of q -matchings, there are $\binom{r-1}{q}$ ways to choose how the sets meet $[r - 1]$, and then at least $\binom{n-kq-r}{k-1}$ choices for each set. Hence,

in total we have

$$\begin{aligned} \text{dp}^{(q)}(\mathcal{L}_{n,k}(s)) &\geq (1 - o(1)) \left(\alpha \binom{r-1}{q-1} + \binom{r-1}{q} \right) \binom{n-1}{k-1}^q \\ &\geq (1 - o(1)) \left(\frac{\alpha \binom{r-1}{q-1} + \binom{r-1}{q}}{(r-1+\alpha)^q} \right) s^q. \end{aligned}$$

For any $s > 0$, this function of α is monotone increasing when $0 \leq \alpha \leq 1$, and so the right-hand side is minimized when $\alpha = 0$. This gives the lower bound

$$(5) \quad \text{dp}^{(q)}(\mathcal{L}_{n,k}(s)) \geq (1 - o(1)) \binom{r-1}{q} \left(\frac{s}{r-1} \right)^q.$$

Having established these bounds, we now prove Theorem 1.7.

Proof of Theorem 1.7. Our proof is by induction, on n, q and s . The base case for $q = 2$ is given by Theorem 1.6¹. As noted earlier, if $s \leq \binom{n}{k} - \binom{n-q+1}{k}$, then $\mathcal{L}_{n,k}(s)$ does not contain any q -matchings, and hence is clearly optimal. Hence, we may proceed to the induction step, with $q \geq 3$ and $\binom{n}{k} - \binom{n-q+1}{k} < s \leq \binom{n}{k} - \binom{n-\ell}{k}$. In particular, we have $q \leq r \leq \ell$ and $s = \Omega(n^{k-1})$.

Let \mathcal{F} be an extremal system of size s . We again first consider the case where \mathcal{F} contains a full star, which we shall assume to be all sets containing 1. We split our q -matchings based on whether or not they meet 1, giving $\text{dp}^{(q)}(\mathcal{F}) = |\mathcal{F}^{(q)}(1)| + \text{dp}^{(q)}(\mathcal{F} \setminus \mathcal{F}(1))$.

Note that every $(q-1)$ -matching not meeting 1 can be extended to a q -matching by exactly $\binom{n-k(q-1)-1}{k-1}$ sets containing 1, so $|\mathcal{F}^{(q)}(1)| = \text{dp}^{(q-1)}(\mathcal{F} \setminus \mathcal{F}(1)) \binom{n-k(q-1)-1}{k-1}$. By the induction hypothesis, $\text{dp}^{(q-1)}(\mathcal{F} \setminus \mathcal{F}(1))$ is minimized by the lexicographical order. Similarly, $\text{dp}^{(q)}(\mathcal{F} \setminus \mathcal{F}(1))$ is also minimized by the lexicographical order, and hence we deduce that $\text{dp}^{(q)}(\mathcal{F}) \geq \text{dp}^{(q)}(\mathcal{L}_{n,k}(s))$.

Thus we may assume that \mathcal{F} does not contain any full stars. Hence, for any $x \in [n]$ and any $F \in \mathcal{F}$, we may replace F by a set containing x .

Step 1: Show there is a popular element $x \in [n]$, with $|\mathcal{F}^{(q-1)}(x)| = \Omega(s^{q-1})$.

A $(q-1)$ -matching in \mathcal{F} can be extended to a q -matching by a set $F \in \mathcal{F}$ precisely when the other $q-1$ sets do not meet F . Thus F is in $\text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(F)|$ q -matchings. Summing over all F gives

$$q \cdot \text{dp}^{(q)}(\mathcal{F}) = \sum_{F \in \mathcal{F}} \left(\text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(F)| \right)$$

¹ Alternatively, we may use the trivial base case of $q = 1$, where we merely count the number of sets.

$$= s \cdot \text{dp}^{(q-1)}(\mathcal{F}) - \sum_{F \in \mathcal{F}} |\mathcal{F}^{(q-1)}(F)|.$$

By the induction hypothesis, $\text{dp}^{(q-1)}(\mathcal{F}) \geq \text{dp}^{(q-1)}(\mathcal{L}_{n,k}(s))$, and since \mathcal{F} is extremal, we must have $\text{dp}^{(q)}(\mathcal{F}) \leq \text{dp}^{(q)}(\mathcal{L}_{n,k}(s))$. Combining these facts with the bounds from (4) and (5), we get

$$\begin{aligned} \sum_{F \in \mathcal{F}} |\mathcal{F}^{(q-1)}(F)| &= s \cdot \text{dp}^{(q-1)}(\mathcal{F}) - q \cdot \text{dp}^{(q)}(\mathcal{F}) \\ &\geq s \cdot \text{dp}^{(q-1)}(\mathcal{L}_{n,k}(s)) - q \cdot \text{dp}^{(q)}(\mathcal{L}_{n,k}(s)) \\ &\geq (1 - o(1)) \binom{r-1}{q-1} \frac{s^q}{(r-1)^{q-1}} - q \binom{r}{q} \frac{s^q}{r^q} = \Omega(s^q). \end{aligned}$$

Averaging over the s sets in \mathcal{F} , we must have $|\mathcal{F}^{(q-1)}(F)| = \Omega(s^{q-1})$ for some $F \in \mathcal{F}$. Since $\mathcal{F}^{(q-1)}(F) = \cup_{x \in F} \mathcal{F}^{(q-1)}(x)$, by averaging over the k elements in F we have $|\mathcal{F}^{(q-1)}(x)| = \Omega(s^{q-1})$ for some $x \in F$.

This completes Step 1.

Step 2: Show there is a cover of size r .

From Step 1, we know there is some popular element, which we may assume to be 1. We start by showing the existence of a reasonably small cover.

Claim 7. $X = \{x : |\mathcal{F}^{(q-1)}(x)| \geq \frac{1}{k} |\mathcal{F}^{(q-1)}(1)|\}$ is a cover for \mathcal{F} .

Proof. Suppose for contradiction that X was not a cover for \mathcal{F} . Then there is some set $F \in \mathcal{F}$ such that $F \cap X = \emptyset$, and so $|\mathcal{F}^{(q-1)}(x)| < \frac{1}{k} |\mathcal{F}^{(q-1)}(1)|$ for all $x \in F$. Since $\mathcal{F}^{(q-1)}(F) = \cup_{x \in F} \mathcal{F}^{(q-1)}(x)$, the number of q -matchings F is contained in is given by

$$\begin{aligned} \text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(F)| &\geq \text{dp}^{(q-1)}(\mathcal{F}) - \sum_{x \in F} |\mathcal{F}^{(q-1)}(x)| \\ &> \text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(1)|. \end{aligned}$$

On the other hand, a set containing 1 can be in at most $\text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(1)|$ q -matchings. Since $\mathcal{F}(1)$ is not a full star, we may replace F with a set containing 1, which would decrease the number of q -matchings in \mathcal{F} . This contradicts the optimality of \mathcal{F} , and it follows that X is a cover. ■

Having shown that this set X is a cover, we now show that X is not too big; its size is bounded by a function of k , q and ℓ .

Claim 8. $|X| = O(1)$.

Proof. As there can be at most s^{q-1} $(q-1)$ -matchings in \mathcal{F} , we have

$$\begin{aligned} \frac{1}{k} |\mathcal{F}^{(q-1)}(1)||X| &\leq \sum_{x \in X} |\mathcal{F}^{(q-1)}(x)| \leq \sum_{x \in [n]} |\mathcal{F}^{(q-1)}(x)| \\ &= k(q-1) \text{dp}^{(q-1)}(\mathcal{F}) \leq k(q-1)s^{q-1}. \end{aligned}$$

Since $|\mathcal{F}^{(q-1)}(1)| = \Omega(s^{q-1})$, this gives $|X| = O(1)$, as required. ■

Now take a minimal subcover of X , which we may assume to be $[m]$, where $m = O(1)$. We shall shift our focus from $(q-1)$ -matchings to the individual sets themselves. For each $i \in [m]$, we shall let $\mathcal{F}^-(i) = \{F \in \mathcal{F} : F \cap [m] = \{i\}\}$ be those sets in \mathcal{F} that meet $[m]$ precisely at i ; by the minimality of the cover, these subsystems are non-empty. Since any set in $\mathcal{F}(i) \setminus \mathcal{F}^-(i)$ must contain not just i but also some other element in $[m]$, we have $|\mathcal{F}^-(i)| \geq |\mathcal{F}(i)| - m \binom{n-2}{k-2} = |\mathcal{F}(i)| - o(s)$.

We will now show that for an extremal system, we must have $m = r$. We first require the following claim.

Claim 9. For any $i, j \in [m]$, we have $|\mathcal{F}(i)| = |\mathcal{F}(j)| + o(s)$.

Proof. Recall that set $F \in \mathcal{F}$ contributes $\text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(F)|$ q -matchings to \mathcal{F} . By estimating $|\mathcal{F}^{(q-1)}(F)|$ for sets containing i or j , we shall show that if $|\mathcal{F}(i)|$ and $|\mathcal{F}(j)|$ are very different, then we can decrease the number of q -matchings by shifting sets.

Consider a set $F \in \mathcal{F}^-(i)$. We wish to bound $|\mathcal{F}^{(q-1)}(F)|$.

For every $(q-1)$ -matching in $\mathcal{F}^{(q-1)}(F)$, we must have at least one of the sets in the $(q-1)$ -matching meeting F . Either this set can contain i , in which case there are $|\mathcal{F}(i)|$ possibilities, or it contains some element in $F \setminus \{i\}$, as well as some element in $[m]$. However, the number of options in the latter case is at most $mk \binom{n-2}{k-2} = o(s)$. We can then count the number of possibilities for the other sets in the matching just as we did when establishing the inequalities (4) and (5). First we choose representatives $A \subset [m] \setminus \{i\}$ for the other $q-2$ sets, and then we choose sets corresponding to the given elements; that is, $H \in \mathcal{F}(a)$ for all $a \in A$. This provides an overestimate for $|\mathcal{F}^{(q-1)}(F)|$, as some of these collections of $q-1$ sets may not be disjoint, while some are counted multiple times. However, we do obtain the upper bound

$$(6) \quad |\mathcal{F}^{(q-1)}(F)| \leq (1 + o(1)) |\mathcal{F}(i)| \sum_{A \in \binom{[m] \setminus \{i\}}{q-2}} \prod_{a \in A} |\mathcal{F}(a)|.$$

We now consider replacing F by some set G containing j , and determine how many new q -matchings would be formed. The number of q -matchings

G contributes is $\text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(G)| \leq \text{dp}^{(q-1)}(\mathcal{F}) - |\mathcal{F}^{(q-1)}(j)|$, since $j \in G$.

To bound $|\mathcal{F}^{(q-1)}(j)|$, note that we can form $(q-1)$ -matchings containing j by first choosing a set from $\mathcal{F}^-(j)$, then choosing a set of $q-2$ other representatives $A \subset [m] \setminus \{j\}$, and choosing disjoint sets $H \in \mathcal{F}^-(a)$, $a \in A$. To ensure the sets we choose are disjoint, we must avoid any elements we have already used. There can be at most $k(q-1)$ such elements, and so we have to avoid at most $\binom{n-1}{k-1} - \binom{n-k(q-1)-1}{k-1} \leq k(q-1) \binom{n-2}{k-2} = o(s)$ sets each time. By choosing the sets from $\mathcal{F}^-(a)$, and not $\mathcal{F}(a)$, we ensure there is no overcounting, as each such $(q-1)$ -matching has a unique set of representatives in $[m]$. Thus we have the bound

$$\begin{aligned}
 (7) \quad & |\mathcal{F}^{(q-1)}(j)| \geq |\mathcal{F}^-(j)| \sum_{A \in \binom{[m] \setminus \{j\}}{q-2}} \prod_{a \in A} (|\mathcal{F}^-(a)| - o(s)) \\
 & = (1 - o(1)) |\mathcal{F}(j)| \sum_{A \in \binom{[m] \setminus \{j\}}{q-2}} \prod_{a \in A} |\mathcal{F}(a)|,
 \end{aligned}$$

since $|\mathcal{F}^-(a)| = |\mathcal{F}(a)| - o(s)$ for all $a \in [m]$.

Since \mathcal{F} is optimal, we must have $|\mathcal{F}^{(q-1)}(F)| \geq |\mathcal{F}^{(q-1)}(G)|$. Comparing (6) and (7), we find

$$\begin{aligned}
 (1 + o(1)) |\mathcal{F}(i)| & \sum_{A \in \binom{[m] \setminus \{i\}}{q-2}} \prod_{a \in A} |\mathcal{F}(a)| \\
 & \geq (1 - o(1)) |\mathcal{F}(j)| \sum_{A \in \binom{[m] \setminus \{j\}}{q-2}} \prod_{a \in A} |\mathcal{F}(a)|.
 \end{aligned}$$

Some terms appear on both sides of the inequality, and so taking the difference gives

$$(|\mathcal{F}(i)| - |\mathcal{F}(j)|) \sum_{A \in \binom{[m] \setminus \{i,j\}}{q-2}} \prod_{a \in A} |\mathcal{F}(a)| \geq o(s^{q-1}).$$

This implies $|\mathcal{F}(i)| \geq |\mathcal{F}(j)| + o(s)$. By symmetry, the reverse inequality also holds, and thus $|\mathcal{F}(i)| = |\mathcal{F}(j)| + o(s)$, as required. ■

Note that we have

$$s = |\mathcal{F}| = \left| \bigcup_{i \in [m]} \mathcal{F}(i) \right| \geq \sum_{i=1}^m |\mathcal{F}(i)| - \sum_{i < j} |\mathcal{F}(i) \cap \mathcal{F}(j)|.$$

Since $|\mathcal{F}(i) \cap \mathcal{F}(j)| \leq \binom{n-2}{k-2} = o(s)$ for all i, j , it follows that $\sum_{i=1}^m |\mathcal{F}(i)| = s + o(s)$. Claim 9 shows that all the stars have approximately the same size, and so $|\mathcal{F}(i)| = \frac{s}{m} + o(s)$ for each $1 \leq i \leq m$. We can now show that we have a smallest possible cover.

Claim 10. *If \mathcal{F} is extremal, then \mathcal{F} can be covered by r elements.*

Proof. Now that we have control over the sizes of the subsystems $\mathcal{F}(i)$, we can estimate the number of q -matchings the system contains. As in our calculations for Claim 9, we can obtain a q -matching by choosing a collection A of q elements in $[m]$, and then choosing sets from the corresponding subsystems $\mathcal{F}(a)$, $a \in A$. In order for this choice of sets to form a q -matching, each set we choose should avoid the elements of the previously chosen sets, of which there can be at most $k(q-1)$. Moreover, to avoid overcounting, we shall choose sets from $\mathcal{F}^-(a)$, and so shall avoid the other $m-1$ elements of $[m]$. Thus, for a given $a \in A$, the forbidden sets are those containing a , and one of at most $k(q-1) + m - 1$ other elements, and so we forbid at most $(k(q-1) + m - 1) \binom{n-2}{k-2} = o(s)$ sets. Thus we have

$$\text{dp}^{(q)}(\mathcal{F}) \geq \sum_{A \in \binom{[m]}{q}} \prod_{a \in A} (|\mathcal{F}(a)| - o(s)) = (1 - o(1)) \binom{m}{q} \left(\frac{s}{m}\right)^q.$$

On the other hand, since \mathcal{F} is extremal, we must have $\text{dp}^{(q)}(\mathcal{F}) \leq \text{dp}^{(q)}(\mathcal{L}_{n,k}(s)) \leq \binom{r}{q} \left(\frac{s}{r}\right)^q$. As $\binom{m}{q} \left(\frac{s}{m}\right)^q$ is increasing in m , these bounds imply we must have $m = r$. ■

This concludes Step 2.

Step 3: Show that $\mathcal{L}_{n,k}(s)$ is optimal.

We complete the induction by showing that $\mathcal{L}_{n,k}(s)$ does indeed minimize the number of q -matchings. From the previous steps, we may assume that an extremal system \mathcal{F} is covered by $[r]$. As before, we shall let $\mathcal{A} = \left\{ A \in \binom{[n]}{k} : A \cap [r] \neq \emptyset \right\}$, so $\mathcal{F} \subset \mathcal{A}$, and we let $\mathcal{G} = \mathcal{A} \setminus \mathcal{F}$. Note that for every $G \in \mathcal{G}$, $\text{dp}^{(q-1)}(\mathcal{A}) - |\mathcal{A}^{(q-1)}(G)|$ counts the number of q -matchings in \mathcal{A} containing G . Hence

$$\begin{aligned} \text{dp}^{(q)}(\mathcal{F}) &\geq \text{dp}^{(q)}(\mathcal{A}) - \sum_{G \in \mathcal{G}} \left(\text{dp}^{(q-1)}(\mathcal{A}) - |\mathcal{A}^{(q-1)}(G)| \right) \\ &= \text{dp}^{(q)}(\mathcal{A}) - |\mathcal{G}| \text{dp}^{(q-1)}(\mathcal{A}) + \sum_{G \in \mathcal{G}} |\mathcal{A}^{(q-1)}(G)|. \end{aligned}$$

Now the first two terms are independent of the structure of \mathcal{F} . We claim that $|\mathcal{A}^{(q-1)}(G)|$ is minimized when $|G \cap [r]| = 1$. Indeed, fix some $G \in \mathcal{G}$. Note that the number of $(q-1)$ -matchings in \mathcal{A} that only meet G outside $[r]$ is at most $kr \binom{n-2}{k-2} s^{q-2} = o(s^{q-1})$, since we must choose one of k elements of G and one of r elements of $[r]$ for the set to contain, and then there are at most s^{q-2} choices for the remaining $q-2$ sets. Hence almost all the $(q-1)$ -matchings in $\mathcal{A}^{(q-1)}(G)$ meet G in $G \cap [r]$, and thus $|\mathcal{A}^{(q-1)}(G)|$ is obviously minimized when $|G \cap [r]| = 1$.

When $\mathcal{F} = \mathcal{L}_{n,k}(s)$, we have $G \cap [r] = \{r\}$ for all $G \in \mathcal{G}$, and so the right-hand side is minimized. Moreover, because \mathcal{G} is an intersecting system, it follows that every $(q-1)$ -matching in \mathcal{A} can contain at most 1 set from \mathcal{G} , and so the above inequality is in fact an equality. This shows that $\mathcal{L}_{n,k}(s)$ minimizes the number of q -matchings.

This completes the induction step, and thus the proof of Theorem 1.7. ■

4. t -disjoint pairs

We now seek a different extension of Theorem 1.6. Recall that we call a pair of sets F_1, F_2 t -intersecting if $|F_1 \cap F_2| \geq t$, and t -disjoint otherwise. As shown by Wilson [24], provided $n \geq (k-t+1)(t+1)$, the largest t -intersecting system consists of $\binom{n-t}{k-t}$ sets that share a common t -set $X \in \binom{[n]}{t}$; we call such a system a (full) t -star with center X . Note that $\mathcal{L}_{n,k}(\binom{n-t}{k-t})$ is itself a t -star with center $[t]$. In the following theorem, we show that when n is sufficiently large, the minimum number of t -disjoint pairs is attained by taking full t -stars. In this setting, not all unions of t -stars are isomorphic, as the structure depends on how the centers intersect. We show that it is optimal to have the centers be the first few sets in the lexicographical ordering on $\binom{[n]}{t}$, which is the case for $\mathcal{L}_{n,k}(s)$.

Theorem 1.8. *Provided $n \geq n_2(k, t, \ell)$ and $0 \leq s \leq \binom{n-t+1}{k-t+1} - \binom{n-t-\ell+1}{k-t+1}$, $\mathcal{L}_{n,k}(s)$ minimizes the number of t -disjoint pairs among all systems of s sets in $\binom{[n]}{k}$.*

It shall sometimes be helpful to count the number of t -intersecting pairs instead of t -disjoint pairs. Thus we introduce the notation $\text{int}_t(\mathcal{F})$ to represent the number of t -intersecting pairs of sets in \mathcal{F} , and $\text{int}_t(\mathcal{F}, \mathcal{G}) = |\{(F, G) \in \mathcal{F} \times \mathcal{G} : |F \cap G| \geq t\}|$ to count the number of cross- t -intersections between \mathcal{F} and \mathcal{G} . Note that a set F is t -intersecting with itself, since $|F \cap F| = k > t$. Since $\sum_{F \in \mathcal{F}} \text{int}_t(F, \mathcal{F})$ counts the t -intersecting pairs between distinct sets twice, and those with the same set only once, we obtain the identity $\sum_{F \in \mathcal{F}} \text{int}_t(F, \mathcal{F}) = 2 \text{int}_t(\mathcal{F}) - |\mathcal{F}|$.

We begin with a heuristic calculation that suggests why it is optimal to have full t -stars. Let \mathcal{F} be a full t -star, say with center $X \in \binom{[n]}{t}$, and let F be a set not containing X . For a set G in \mathcal{F} to be t -intersecting with F , G must contain the t elements of X , as well as some $t - |F \cap X|$ elements from F . The number of such sets G is maximized when $|F \cap X| = t - 1$, giving

$$(8) \quad \text{int}_t(F, \mathcal{F}) \leq (k - t + 1) \binom{n - t - 1}{k - t - 1} = O(n^{k-t-1}) = o(|\mathcal{F}|).$$

Hence, if a t -star does not contain a set F , F is t -disjoint from almost all its members. It should thus be optimal to take full t -stars, as that is where the t -intersections come from. Indeed, this turns out to be the case. As we shall see, for a set system \mathcal{F} , the leading term in $\text{dp}_t(\mathcal{F})$ is determined by the number of t -stars in \mathcal{F} . While unions of t -stars may be non-isomorphic, the differences only affect the lower order terms of $\text{dp}_t(\mathcal{F})$.

In order to prove Theorem 1.8, we shall require a few preliminary results. Proposition 4.1 can be thought of as a rough characterization of extremal systems, as it shows that the extremal systems should be supported on the right number of t -stars. To this end, it will be useful to define an *almost full t -star* to be a t -star in \mathcal{F} containing $(1 - o(1)) \binom{n-t}{k-t}$ sets. Formally, this means that for all fixed k, t and ℓ , there is some sufficiently small $\varepsilon = \varepsilon(k, t, \ell) > 0$, and a t -star is almost full if it contains $(1 - \varepsilon) \binom{n-t}{k-t}$ sets.

Proposition 4.1. *Suppose $n \geq n_2(k, \ell, t)$, and $\binom{n-t+1}{k-t+1} - \binom{n-t-r+2}{k-t+1} < s \leq \binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1}$. If $\mathcal{F} \subset \binom{[n]}{k}$ has the minimum number of t -disjoint pairs over all systems of s sets, then either:*

- (i) \mathcal{F} contains $r - 1$ full t -stars,
- (ii) \mathcal{F} consists of r almost full t -stars, or
- (iii) \mathcal{F} consists of $r - 1$ almost full t -stars.

Once we have determined the large-scale structure of the extremal systems, the following lemmas allow us to analyze the lower-order terms and determine that the lexicographical ordering is indeed optimal.

Lemma 4.2 shows that of all unions of r full t -stars, the lexicographical ordering contains the fewest sets. This may seem to contradict the lexicographical ordering being optimal, given that the heuristic given by (8) suggests that it is optimal to take as few t -stars as possible, and hence we might try to make the union of these stars accommodate as many sets as possible. However, it is because there is more overlap between the lexicographical t -stars that there are fewer t -disjoint pairs between stars.

Lemma 4.2. *Suppose $n \geq n_2(k, t, r)$, and let \mathcal{F} be the union of r full t -stars in $\binom{[n]}{k}$. Then $|\mathcal{F}| \geq \binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1}$, with equality if and only if \mathcal{F} is isomorphic to $\mathcal{L}_{n,k} \left(\binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1} \right)$.*

The next lemma shows that if we have r full t -stars, and add a new set to the system, we minimize the number of new t -disjoint pairs created when we have the lexicographical initial segment.

Lemma 4.3. *Suppose $n \geq n_2(k, t, r)$, let $\mathcal{L} = \mathcal{L}_{n,k} \left(\binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1} \right)$ be the first r full t -stars in the lexicographical order, and let L be a set containing $\{1, 2, \dots, t-1\}$ that is not in \mathcal{L} . Let \mathcal{F} be the union of r full t -stars with centers $\{X_1, X_2, \dots, X_r\}$, and let F be any k -set not in \mathcal{F} . Then $\text{dp}_t(F, \mathcal{F}) \geq \text{dp}_t(L, \mathcal{L})$, with equality if and only if $\mathcal{F} \cup \{F\}$ is isomorphic to $\mathcal{L} \cup \{L\}$.*

However, the comparison in Lemma 4.3 is not entirely fair, as Lemma 4.2 shows that \mathcal{L} will have fewer sets than \mathcal{F} , while we ought to be comparing systems of the same size. We do this in our final lemma, in the cleanest case when the system \mathcal{F} is a union of full t -stars.

Lemma 4.4. *Suppose $n \geq n_2(k, t, r)$, let \mathcal{F} be the union of r full t -stars with centers $X_i, 1 \leq i \leq r$, and let $\mathcal{L} = \mathcal{L}_{n,k}(|\mathcal{F}|)$. Then $\text{dp}_t(\mathcal{F}) \geq \text{dp}_t(\mathcal{L})$, with equality if and only if \mathcal{F} is isomorphic to \mathcal{L} .*

Armed with Proposition 4.1 and these three lemmas, whose proofs we defer until later in this section, we now show how to deduce Theorem 1.8.

Proof of Theorem 1.8. Let r be such that $\binom{n-t+1}{k-t+1} - \binom{n-t-r+2}{k-t+1} < s \leq \binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1}$. In this range $\mathcal{L}_{n,k}(s)$ consists of $r-1$ full t -stars, with the remaining sets forming a partial r th t -star. If $r=1$, then $\mathcal{L}_{n,k}(s)$ is t -intersecting, and therefore clearly optimal. Hence, we may assume $r \geq 2$, and in particular this implies $s = \Omega(n^{k-t})$.

Suppose \mathcal{F} is an optimal system of size s . By analyzing the three cases in Proposition 4.1 in turn, we shall show that $\text{dp}_t(\mathcal{F}) \geq \text{dp}_t(\mathcal{L}_{n,k}(s))$, thus completing the proof of Theorem 1.8.

Case (i): Suppose \mathcal{F} contains $r-1$ full t -stars, whose union we shall denote by \mathcal{F}_1 , and $s_2 = s - |\mathcal{F}_1|$ other sets, denoted by \mathcal{F}_2 . We then have

$$\begin{aligned} \text{dp}_t(\mathcal{F}) &= \text{dp}_t(\mathcal{F}_1) + \text{dp}_t(\mathcal{F}_1, \mathcal{F}_2) + \text{dp}_t(\mathcal{F}_2) \geq \text{dp}_t(\mathcal{F}_1) + \text{dp}_t(\mathcal{F}_1, \mathcal{F}_2) \\ &= \text{dp}_t(\mathcal{F}_1) + \sum_{F \in \mathcal{F}_2} \text{dp}_t(F, \mathcal{F}_1) \geq \text{dp}_t(\mathcal{F}_1) + s_2 \cdot \text{dp}_t(F_0, \mathcal{F}_1), \end{aligned}$$

where $F_0 \in \mathcal{F}_2$ minimizes $\text{dp}_t(F, \mathcal{F}_1)$.

Let $\mathcal{L} = \mathcal{L}_{n,k}(s)$ be the corresponding lexicographical initial segment, $\mathcal{L}_1 = \mathcal{L}_{n,k}(|\mathcal{F}_1|)$ be the first $|\mathcal{F}_1|$ sets in the lexicographical ordering, and let $\mathcal{L}_2 = \mathcal{L} \setminus \mathcal{L}_1$ be the next s_2 sets. By Lemma 4.2, it follows that \mathcal{L}_1 consists of at least $r-1$ full t -stars, and so \mathcal{L}_2 lies entirely within the r th lexicographical t -star, and is thus t -intersecting. Hence

$$\begin{aligned} \text{dp}_t(\mathcal{L}) &= \text{dp}_t(\mathcal{L}_1) + \text{dp}_t(\mathcal{L}_1, \mathcal{L}_2) + \text{dp}_t(\mathcal{L}_2) = \text{dp}_t(\mathcal{L}_1) + \text{dp}_t(\mathcal{L}_1, \mathcal{L}_2) \\ &= \text{dp}_t(\mathcal{L}_1) + \sum_{L \in \mathcal{L}_2} \text{dp}_t(L, \mathcal{L}_1) \leq \text{dp}_t(\mathcal{L}_1) + s_2 \cdot \text{dp}_t(L_0, \mathcal{L}_1), \end{aligned}$$

where $L_0 \in \mathcal{L}_2$ maximizes $\text{dp}_t(L, \mathcal{L}_1)$ (in fact, by symmetry, this is equal for all $L \in \mathcal{L}_2$).

Note that L_0 will belong to the r th t -star of \mathcal{L} , and hence, $\text{dp}_t(L_0, \mathcal{L}_1)$ will only count t -disjoint pairs between L_0 and the union of the first $r-1$ t -stars of \mathcal{L}_1 . By Lemma 4.3, we have $\text{dp}_t(F_0, \mathcal{F}_1) \geq \text{dp}_t(L_0, \mathcal{L}_1)$, and by Lemma 4.4, we have $\text{dp}_t(\mathcal{F}_1) \geq \text{dp}_t(\mathcal{L}_1)$, from which we deduce $\text{dp}_t(\mathcal{F}) \geq \text{dp}_t(\mathcal{L})$, as required.

Case (ii): In this case we have r almost full t -stars. Using a complementarity argument, we shall reduce this to case (i).

Suppose \mathcal{F} is the union of r almost full t -stars with centers $\{X_1, X_2, \dots, X_r\}$, let $\mathcal{A} = \cup_{i=1}^r \{A \in \binom{[n]}{k} : X_i \subset A\}$ be the system of all sets containing some X_i , and let $\mathcal{G} = \mathcal{A} \setminus \mathcal{F}$. On account of the t -stars being almost full, we have $|\mathcal{G}| = o(n^{k-t})$.

Running the same complementarity argument as in the proof of Theorem 1.6, we have

$$\begin{aligned} \text{dp}_t(\mathcal{F}) &= \text{dp}_t(\mathcal{A}) - \text{dp}_t(\mathcal{G}, \mathcal{A}) + \text{dp}_t(\mathcal{G}) \\ (9) \qquad &= \text{dp}_t(\mathcal{A}) - \sum_{G \in \mathcal{G}} \text{dp}_t(G, \mathcal{A}) + \text{dp}_t(\mathcal{G}). \end{aligned}$$

To minimize $\text{dp}_t(\mathcal{F})$, we seek to maximize $\sum_{G \in \mathcal{G}} \text{dp}_t(G, \mathcal{A})$ while minimizing $\text{dp}_t(\mathcal{G})$. We shall obtain these extrema by shifting the system so that the missing sets, \mathcal{G} , will all belong to one of the t -stars $\mathcal{A}(X_i)$. In this case, the shifted system, \mathcal{F}' , will contain $r-1$ full t -stars. Hence, we will have reduced the problem to case (i), and so $\text{dp}_t(\mathcal{F}) \geq \text{dp}_t(\mathcal{F}') \geq \text{dp}_t(\mathcal{L}_{n,k}(s))$, as desired.

Note that when \mathcal{G} is a subset of one of the t -stars, \mathcal{G} is t -intersecting, and so $\text{dp}_t(\mathcal{G}) = 0$ is minimized. We now show how to choose which t -star \mathcal{G} should belong to in order to maximize $\sum_{G \in \mathcal{G}} \text{dp}_t(G, \mathcal{A})$.

Since \mathcal{A} is of fixed size, maximizing $\text{dp}_t(G, \mathcal{A})$ is equivalent to minimizing $\text{int}_t(G, \mathcal{A})$. For $G \in \mathcal{G}$, $\text{int}_t(G, \mathcal{A})$ is determined by the intersections

$\{G \cap X_i : 1 \leq i \leq r\}$. There are only a bounded number of possibilities for these intersections, and so we may choose one which minimizes $\text{int}_t(G, \mathcal{A})$, under the restriction that $X_i \subset G$ for some i , since $G \in \mathcal{A}$. By (8), the number of t -intersecting pairs between G and a t -star not containing G is $o(s)$, and so this minimum occurs when G contains some X_i and no other elements from $\cup_j X_j \setminus X_i$. The number of choices for the set G is then at least $\binom{n-rt}{k-t}$, since after choosing the t elements of X_i , we wish to avoid the remaining elements in $\cup_j X_j$, of which there are at most $(r-1)t$. Since $\binom{n-rt}{k-t} \geq |G| = o(n^{k-t})$, we may choose all $G \in \mathcal{G}$ to come from the t -star with center X_i in order to minimize the right hand side of (9). We have thus resolved case (ii).

Case (iii): In this case we have $r-1$ almost full t -stars. Since the size of this system is at most $(r-1)\binom{n-t}{k-t}$, while the size of the first $r-1$ t -stars in $\mathcal{L}_{n,k}(s)$ is $\binom{n-t+1}{k-t+1} - \binom{n-t-r+2}{k-t+1} = (r-1)\binom{n-t}{k-t} + o(n^{k-t})$, we can conclude that r th partial t -star in $\mathcal{L}_{n,k}(s)$ has only $o(n^{k-t})$ sets.

Given the system \mathcal{F} , we shall construct a larger system \mathcal{F}' by filling the $r-1$ almost full t -stars. Suppose we have to add s_1 sets in order to do so. Note that since the t -stars were almost full, we have $s_1 = o(n^{k-t})$. Since each of the s_1 sets is added to an almost full t -star, it contributes at least $(1-o(1))\binom{n-t}{k-t}$ t -intersecting pairs. Hence $\text{int}_t(\mathcal{F}') \geq \text{int}_t(\mathcal{F}) + (1-o(1))s_1\binom{n-t}{k-t}$.

On the other hand, consider adding the same number of sets to the lexicographical initial segment. The sets in $\mathcal{L}_{n,k}(s+s_1) \setminus \mathcal{L}_{n,k}(s)$ all belong only to the r th t -star, which has only $o(n^{k-t})$ sets. Our calculation in (8) shows that each such set also only gains $o(n^{k-t})$ t -intersections from the other stars, and so we have $\text{int}_t(\mathcal{L}_{n,k}(s+s_1)) \leq \text{int}_t(\mathcal{L}_{n,k}(s)) + s_1 \cdot o(n^{k-t})$.

Now \mathcal{F}' consists of $r-1$ full t -stars, and so by Lemma 4.4, we have $\text{dp}_t(\mathcal{F}') \geq \text{dp}_t(\mathcal{L}_{n,k}(s+s_1))$, or, equivalently, $\text{int}_t(\mathcal{F}') \leq \text{int}_t(\mathcal{L}_{n,k}(s+s_1))$. Thus $\text{int}_t(\mathcal{F}) + (1-o(1))s_1\binom{n-t}{k-t} \leq \text{int}_t(\mathcal{L}_{n,k}(s)) + s_1 \cdot o(n^{k-t})$, and so $\text{int}_t(\mathcal{F}) \leq \text{int}_t(\mathcal{L}_{n,k}(s))$, with a strict inequality unless $s_1 = 0$. This implies $\text{dp}_t(\mathcal{F}) \geq \text{dp}_t(\mathcal{L}_{n,k}(s))$, as required.

Hence, we may conclude that for any system \mathcal{F} with s sets, we have $\text{dp}_t(\mathcal{F}) \geq \text{dp}_t(\mathcal{L}_{n,k}(s))$, proving Theorem 1.8. ■

By analyzing the cases when we have equality, and using the fact that in Lemmas 4.2, 4.3 and 4.4 we only have equality when the systems are isomorphic to the lexicographical ordering, we can characterize all extremal systems.

Corollary 4.5. *Suppose $n \geq n_2(k, \ell, t)$, and $0 \leq s \leq \binom{n-t+1}{k-t+1} - \binom{n-t-\ell+1}{k-t+1}$, and $\mathcal{F} \subset \binom{[n]}{k}$ minimizes the number of t -disjoint pairs over all systems of s sets.*

Then all sets $F \in \mathcal{F}$ share some common $(t-1)$ -set X , and $\mathcal{F}' = \{F \setminus X : F \in \mathcal{F}\}$ minimizes the number of disjoint pairs over all systems of s sets in $\binom{[n] \setminus X}{k-t+1}$.

Moreover, note that this problem can again be thought of as an isoperimetric inequality in an appropriate graph. Consider the generalized Kneser graph $K(n, k, t)$, with $V = \binom{[n]}{k}$ and an edge between the sets X and Y if $|X \cap Y| < t$. This is a regular graph, with every vertex having degree $\sum_{i=0}^{t-1} \binom{k}{i} \binom{n-k}{k-i}$. Thus, combining Theorem 1.8 with Lemma 2.3, we can determine which large systems minimize the number of t -disjoint pairs.

Corollary 4.6. *Provided $n \geq n_2(k, t, \ell)$ and $0 \leq s \leq \binom{n-t+1}{k-t+1} - \binom{n-t-\ell+1}{k-t+1}$, $\binom{[n]}{k} \setminus \mathcal{L}_{n,k}(s)$ minimizes the number of t -disjoint pairs among all systems of $\binom{n}{k} - s$ sets in $\binom{[n]}{k}$.*

It remains to prove the proposition and lemmas. We begin with a proof of Proposition 4.1. The strategy will be very similar to that of Theorem 1.6; assuming the extremal system \mathcal{F} does not have $r-1$ full t -stars, we shall show there is some popular element (that is, an element contained in many sets of \mathcal{F}). From this we will deduce the existence of a small cover, and shall show that either case (ii) or case (iii) must hold.

Proof of Proposition 4.1. We may assume that $r \geq 2$, since if $r = 1$, then case (i) is trivially satisfied. We first estimate the number of t -intersecting pairs in $\mathcal{L}_{n,k}(s)$, so that we have a lower bound on $\text{int}_t(\mathcal{F})$ for any extremal system \mathcal{F} .

Note that $\mathcal{L}_{n,k}(s)$ consists of $r-1$ full t -stars, with the remaining sets forming a partial t -star; suppose there are $\alpha \binom{n-t}{k-t}$ such sets. Since there are $\binom{n-t-1}{k-t-1} = o(n^{k-t})$ sets common to any two t -stars, it follows that $s = (r-1 + \alpha) \binom{n-t}{k-t} + o(n^{k-t})$.

Now any two sets in the same t -star are t -intersecting, while (8) shows that a set is t -intersecting with $o(n^{k-t})$ sets from the other t -stars. Hence, for any extremal system \mathcal{F} we have the bound

$$\begin{aligned} \text{int}_t(\mathcal{F}) &\geq \text{int}_t(\mathcal{L}_{n,k}(s)) = (r-1) \binom{\binom{n-t}{k-t}}{2} + \left(\alpha \binom{n-t}{k-t} \right) + o(n^{2(k-t)}) \\ &= \frac{r-1 + \alpha^2}{2} \binom{n-t}{k-t}^2 + o(n^{2(k-t)}). \end{aligned}$$

Suppose \mathcal{F} contains p full t -stars. If $p = r-1$, then case (i) holds, and we are done. Hence, we may assume $0 \leq p \leq r-2$. Let \mathcal{F}_1 be the union of the p full t -stars, and let $\mathcal{F}_2 = \mathcal{F} \setminus \mathcal{F}_1$ be the remaining sets.

By the same reasoning as above, we must have $|\mathcal{F}_1| = p\binom{n-t}{k-t} + o(n^{k-t})$, and $\text{int}_t(\mathcal{F}_1) = \frac{1}{2}p\binom{n-t}{k-t}^2 + o(n^{2(k-t)})$. No set $F \in \mathcal{F}_2$ is in any of the t -stars of \mathcal{F}_1 , and so (8) gives $\text{int}_t(\mathcal{F}_1, \mathcal{F}_2) = |\mathcal{F}_2| \cdot o(n^{k-t}) = o(n^{2(k-t)})$. Thus $\text{int}_t(\mathcal{F}) = \text{int}_t(\mathcal{F}_1) + \text{int}_t(\mathcal{F}_1, \mathcal{F}_2) + \text{int}_t(\mathcal{F}_2) = \frac{1}{2}p\binom{n-t}{k-t}^2 + \text{int}_t(\mathcal{F}_2) + o(n^{2(k-t)})$, and hence we must have

$$\text{int}_t(\mathcal{F}_2) \geq \frac{r-p-1+\alpha^2}{2} \binom{n-t}{k-t}^2 + o(n^{2(k-t)}) = \Omega(n^{2(k-t)}).$$

We shall now deduce the existence of a t -cover of size $r-p-1$ or $r-p$ for \mathcal{F}_2 , and then show that we must fall into case (ii) or (iii). The first step is to find a t -set that is in many members of \mathcal{F}_2 . Note that none of the t -stars in \mathcal{F}_2 are full, and hence we may shift sets in \mathcal{F}_2 .

Claim 11. *There is some set $X_1 \in \binom{[n]}{t}$ with $|\mathcal{F}_2(X_1)| = \Omega(n^{k-t})$.*

Proof. Let $X_1 \in \binom{[n]}{t}$ be the set maximizing $|\mathcal{F}_2(X)|$. We have

$$\begin{aligned} \text{int}_t(\mathcal{F}_2) - \frac{1}{2}|\mathcal{F}| &= \frac{1}{2} \sum_{F \in \mathcal{F}_2} \text{int}_t(F, \mathcal{F}_2) \\ &= \frac{1}{2} \sum_{F \in \mathcal{F}_2} \left| \cup_{X \in \binom{F}{t}} \mathcal{F}_2(X) \right| \leq \frac{1}{2} \sum_{F \in \mathcal{F}_2} \sum_{X \in \binom{F}{t}} |\mathcal{F}_2(X)| \\ &\leq \frac{1}{2} \sum_{F \in \mathcal{F}_2} \binom{k}{t} |\mathcal{F}_2(X_1)| = \frac{1}{2} \binom{k}{t} |\mathcal{F}_2| |\mathcal{F}_2(X_1)|. \end{aligned}$$

Since $|\mathcal{F}_2| = (r-p-1+\alpha)\binom{n-t}{k-t} = O(n^{k-t})$, and $\text{int}_t(\mathcal{F}_2) = \Omega(n^{2(k-t)})$, it follows that $|\mathcal{F}_2(X_1)| = \Omega(n^{k-t})$, as desired. ■

This allows us to find a small t -cover.

Claim 12. $\mathcal{X} = \left\{ X \in \binom{[n]}{t} : |\mathcal{F}_2(X)| \geq \frac{1}{2} \binom{k}{t} |\mathcal{F}_2(X_1)| \right\}$ is a t -cover for \mathcal{F}_2 .

Proof. Suppose not. Then there is some $F \in \mathcal{F}$ such that for all $X \in \binom{F}{t}$, $|\mathcal{F}_2(X)| < \frac{1}{2} \binom{k}{t} |\mathcal{F}_2(X_1)|$. Thus $\text{int}_t(F, \mathcal{F}_2) \leq \sum_{X \in \binom{F}{t}} |\mathcal{F}_2(X)| < \frac{1}{2} |\mathcal{F}_2(X_1)|$.

Since F has $o(n^{k-t})$ t -intersecting pairs in \mathcal{F}_1 , it follows that $\text{int}_t(F, \mathcal{F}) \leq \frac{1}{2} |\mathcal{F}_2(X_1)| + o(n^{k-t})$.

If we were to replace F with some set G containing X_1 , which is possible as $\mathcal{F}(X_1)$ is not a full t -star, then we would create at least $|\mathcal{F}_2(X_1)|$ t -intersecting pairs. Since $|\mathcal{F}_2(X_1)| = \Omega(n^{k-t})$, it follows that $\text{int}_t(G, \mathcal{F}) > \text{int}_t(F, \mathcal{F})$, which contradicts \mathcal{F} being optimal.

Hence, \mathcal{X} must be a t -cover for \mathcal{F}_2 , as claimed. ■

Claim 13. $|\mathcal{X}| = O(1)$.

Proof. We have

$$\begin{aligned} \binom{k}{t} |\mathcal{F}_2| &= \sum_{F \in \mathcal{F}_2} \left| \binom{F}{t} \right| = \sum_{X \in \binom{[n]}{t}} |\mathcal{F}_2(X)| \geq \sum_{X \in \mathcal{X}} |\mathcal{F}_2(X)| \\ &\geq \frac{1}{2 \binom{k}{t}} |\mathcal{F}_2(X_1)| |\mathcal{X}|. \end{aligned}$$

Since $|\mathcal{F}_2| = O(n^{k-t})$ and $|\mathcal{F}_2(X_1)| = \Omega(n^{k-t})$, it follows that $|\mathcal{X}| = O(1)$, as claimed. ■

Hence, we can write $\mathcal{X} = \{X_1, X_2, \dots, X_m\}$, where $m = O(1)$. Note that there are at most $\binom{n-t-1}{k-t-1} = o(n^{k-t})$ sets in common between any two stars, while the number of sets each t -star contains is at least $\frac{1}{2 \binom{k}{t}} |\mathcal{F}_2(X_1)| = \Omega(n^{k-t})$. Thus in what follows, we consider only those sets in exactly one t -star $\mathcal{F}_2(X_i)$, and shall only lose $o(n^{2(k-t)})$ t -intersecting pairs.

Claim 14. For all $1 \leq i < j \leq m$, $|\mathcal{F}_2(X_i)| = |\mathcal{F}_2(X_j)| + o(n^{k-t})$.

Proof. Consider a set $F \in \mathcal{F}_2(X_i)$. F is t -intersecting with all sets in $\mathcal{F}_2(X_i)$, and, by (8), t -disjoint from almost all other sets. Thus $\text{int}_t(F, \mathcal{F}) = |\mathcal{F}_2(X_i)| + o(n^{k-t})$. If we were instead to replace F with a set G containing X_j , which is possible as $\mathcal{F}(X_j)$ is not a full t -star, then we would create at least $|\mathcal{F}_2(X_j)|$ new t -intersecting pairs. Since \mathcal{F} is optimal, we must have $|\mathcal{F}_2(X_i)| + o(n^{k-t}) \geq |\mathcal{F}_2(X_j)|$.

By symmetry, it follows that $|\mathcal{F}_2(X_i)| = |\mathcal{F}_2(X_j)| + o(n^{k-t})$. ■

Recall that we had $|\mathcal{F}_2| = (r - p - 1 + \alpha) \binom{n-t}{k-t} + o(n^{k-t})$. By Claim 14, it follows that these sets are almost equally distributed between the m t -stars in the t -cover \mathcal{X} , and so $|\mathcal{F}_2(X_i)| = \frac{r-p-1+\alpha}{m} \binom{n-t}{k-t} + o(n^{k-t})$ for each $1 \leq i \leq m$. Moreover, since $|\mathcal{F}_2(X_i)| \leq \binom{n-t}{k-t}$, we must have $m \geq r - p - 1$ if $\alpha = o(1)$, or $m \geq r - p$ if $\alpha = \Omega(1)$.

We can now estimate $\text{int}_t(\mathcal{F}_2)$. We know every set belonging only to the t -star $\mathcal{F}_2(X_i)$ contributes $|\mathcal{F}_2(X_i)| + o(n^{k-t})$ t -intersecting pairs, while there are only $o(n^{2(k-t)})$ t -intersecting pairs from sets in multiple t -stars. Thus

$$\begin{aligned} \text{int}_t(\mathcal{F}_2) &= \frac{1}{2} \sum_{F \in \mathcal{F}_2} \text{int}_t(F, \mathcal{F}_2) + \frac{1}{2} |\mathcal{F}| \\ &= \frac{1}{2} \sum_{i=1}^m \sum_{F \in \mathcal{F}_2(X_i)} \text{int}_t(F, \mathcal{F}_2) + o(n^{2(k-t)}) \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{2} \sum_{i=1}^m |\mathcal{F}_2(X_i)| \left(|\mathcal{F}_2(X_i)| + o(n^{k-t}) \right) + o(n^{2(k-t)}) \\
 &= \frac{1}{2} \sum_{i=1}^m |\mathcal{F}_2(X_i)|^2 + o(n^{2(k-t)}) \\
 &= \frac{(r-p-1+\alpha)^2}{2m} \binom{n-t}{k-t}^2 + o(n^{2(k-t)}).
 \end{aligned}$$

On the other hand, we had the bound

$$\text{int}_t(\mathcal{F}_2) \geq \frac{r-p-1+\alpha^2}{2} \binom{n-t}{k-t}^2 + o(n^{2(k-t)}).$$

Comparing the two, we must have

$$(10) \quad \frac{(r-p-1+\alpha)^2}{2m} \geq \frac{r-p-1+\alpha^2}{2} + o(1).$$

Note that we can write $\frac{r-p-1+\alpha^2}{2} = \frac{1}{2} \sum_{i=1}^m x_i^2$, where

$$x_i = \begin{cases} 1 & 1 \leq i \leq r-p-1 \\ \alpha & i = r-p \\ 0 & r-p+1 \leq i \leq m \end{cases}.$$

Let $\bar{x} = \frac{1}{m} \sum_{i=1}^m x_i = \frac{r-p-1+\alpha}{m}$. With this definition, we then have $\frac{(r-p-1+\alpha)^2}{2m} = \frac{1}{2} m \bar{x}^2$. Since

$$\sum_{i=1}^m x_i^2 = m \bar{x}^2 + \sum_{i=1}^m (x_i - \bar{x})^2,$$

for (10) to hold, we must have $\sum_{i=1}^m (x_i - \bar{x})^2 = o(1)$, and thus $x_i = \bar{x} + o(1)$ for all $1 \leq i \leq m$.

Since $x_1 = 1$, $x_{r-p} = \alpha$, and $x_{r-p+1} = 0$, we must have $m \leq r-p$. Recalling our earlier bound $m \geq r-p-1$, there are only two possibilities. We could have $m = r-p$ and $\alpha = 1 - o(1)$. In this case, each of the $r-p$ t -stars in \mathcal{F}_2 has size $\frac{r-1-p+\alpha}{m} \binom{n-t}{k-t} + o(n^{k-t}) = (1 - o(1)) \binom{n-t}{k-t}$. Combined with the p full t -stars in \mathcal{F}_1 , we see that \mathcal{F} consists of r almost full t -stars, and so we are in case (ii).

The other possible solution is to have $m = r-p-1$, with $\alpha = o(1)$. This implies \mathcal{F}_2 consists of $r-1-p$ almost full t -stars, which, combined with the p full t -stars of \mathcal{F}_1 , means \mathcal{F} falls under case (iii). This completes the proof of Proposition 4.1. ■

We complete this section by proving the three lemmas. First we show that unions of lexicographical stars contain the fewest sets.

Proof of Lemma 4.2. Note that the first r t -stars in the lexicographical ordering have centers $Y_i = \{1, 2, \dots, t-1, t+i-1\}$, $1 \leq i \leq r$, and their union has size $s = \binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1}$. Letting $\mathcal{L} = \mathcal{L}_{n,k} \left(\binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1} \right)$, note that for any set $I \subset [r]$, since $|\cup_{i \in I} Y_i| = t + |I| - 1$, we have $|\cap_{i \in I} \mathcal{L}(Y_i)| = \binom{n-t-|I|+1}{k-t-|I|+1}$. Thus, by Inclusion-Exclusion,

$$\begin{aligned} |\mathcal{L}| &= |\cup_{i=1}^r \mathcal{L}(Y_i)| = \sum_i |\mathcal{L}(Y_i)| - \sum_{i_1 < i_2} |\mathcal{L}(Y_{i_1}) \cap \mathcal{L}(Y_{i_2})| + O(n^{k-t-2}) \\ &= r \binom{n-t}{k-t} - \binom{r}{2} \binom{n-t-1}{k-t-1} + O(n^{k-t-2}). \end{aligned}$$

Now we consider the size of \mathcal{F} . Suppose \mathcal{F} is the union of the r full t -stars with centers $\{X_1, \dots, X_r\}$. We have

$$\begin{aligned} |\mathcal{F}| &= |\cup_{i=1}^r \mathcal{F}(X_i)| \geq \sum_{i=1}^r |\mathcal{F}(X_i)| - \sum_{i_1 < i_2} |\mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})| \\ &= r \binom{n-t}{k-t} - \sum_{i_1 < i_2} |\mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})|. \end{aligned}$$

For every $i_1 < i_2$ we have $|\mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})| = \binom{n-|X_{i_1} \cup X_{i_2}|}{k-|X_{i_1} \cup X_{i_2}|}$. If $|X_{i_1} \cap X_{i_2}| \leq t-2$, then $|X_{i_1} \cup X_{i_2}| \geq t+2$. Hence $|\mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})| = O(n^{k-t-2})$, and so

$$|\mathcal{F}| \geq r \binom{n-t}{k-t} - \left(\binom{r}{2} - 1 \right) \binom{n-t-1}{k-t-1} + O(n^{k-t-2}) > |\mathcal{L}|.$$

Hence we must have $|X_{i_1} \cap X_{i_2}| = t-1$ for all $i_1 < i_2$.

Now, by Inclusion-Exclusion, we have

$$|\mathcal{F}| - r \binom{n-t}{k-t} + \binom{r}{2} \binom{n-t-1}{k-t-1} = \sum_{\substack{I \subset [r] \\ |I| \geq 3}} (-1)^{|I|+1} |\cap_{i \in I} \mathcal{F}(X_i)|.$$

For any set F containing $a \geq 3$ sets X_i , the contribution to the right-hand side is

$$\sum_{b=3}^a (-1)^{b+1} \binom{a}{b} = (1-1)^a + 1 - a + \binom{a}{2} = 1 - a + \binom{a}{2} \geq 1.$$

If we have some $i_1 < i_2 < i_3$ with $|X_{i_1} \cup X_{i_2} \cup X_{i_3}| = t + 1$, then we would have $\binom{n-t-1}{k-t-1}$ sets containing X_{i_1} , X_{i_2} and X_{i_3} . By the preceding equation, we then have

$$|\mathcal{F}| \geq r \binom{n-t}{k-t} - \binom{r}{2} \binom{n-t-1}{k-t-1} + \binom{n-t-1}{k-t-1} > |\mathcal{L}|.$$

Hence, we may assume $|X_{i_1} \cup X_{i_2} \cup X_{i_3}| \geq t + 2$ for all $i_1 < i_2 < i_3$. Since we must have $|X_{i_1} \cap X_{i_2}| = t - 1$ for all $i_1 < i_2$, this implies all of the sets X_i share a common $(t - 1)$ -set, and hence \mathcal{F} is isomorphic to \mathcal{L} , as desired. \blacksquare

The next lemma showed that when adding a set to r full t -stars, the lexicographical stars minimize the number of new t -disjoint pairs.

Proof of Lemma 4.3. \mathcal{L} is the union of the t -stars with centers $\{Y_1, Y_2, \dots, Y_r\}$, as in Lemma 4.2. Since all these sets, and L , contain $[t - 1]$, it is easy to see that

$$\begin{aligned} \text{dp}_t(L, \mathcal{L}) &\leq \sum_{i=1}^r \text{dp}_t(L, \mathcal{L}(Y_i)) - \sum_{i_1 < i_2} \text{dp}_t(L, \mathcal{L}(Y_{i_1}) \cap \mathcal{L}(Y_{i_2})) \\ &\quad + \sum_{i_1 < i_2 < i_3} \text{dp}_t(L, \mathcal{L}(Y_{i_1}) \cap \mathcal{L}(Y_{i_2}) \cap \mathcal{L}(Y_{i_3})) \\ &= r \binom{n-k-1}{k-t} - \binom{r}{2} \binom{n-k-2}{k-t-1} + O(n^{k-t-2}). \end{aligned}$$

On the other hand, we have

$$\text{dp}_t(F, \mathcal{F}) \geq \sum_{i=1}^r \text{dp}_t(F, \mathcal{F}(X_i)) - \sum_{i_1 < i_2} \text{dp}_t(F, \mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})).$$

The first term can be evaluated as follows. Since

$$\text{dp}_t(F, \mathcal{F}(X_i)) = \sum_{a=0}^{t-1-|F \cap X_i|} \binom{k-|F \cap X_i|}{a} \binom{n-k-t+|F \cap X_i|}{k-t-a},$$

if $|F \cap X_i| = t - 1$ we have $\text{dp}_t(F, \mathcal{F}(X_i)) = \binom{n-k-1}{k-t}$, while $\text{dp}_t(F, \mathcal{F}(X_i)) \geq \binom{n-k-2}{k-t} + (k-t+2) \binom{n-k-2}{k-t-1} = \binom{n-k-1}{k-t} + (k-t+1) \binom{n-k-2}{k-t-1}$ otherwise. Moreover, for every $i_1 < i_2$ we have the bound $\text{dp}_t(F, \mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})) \leq |\mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})| \leq \binom{n-t-1}{k-t-1}$. Hence, if $|F \cap X_i| \leq t - 2$ for some i ,

$$\text{dp}_t(F, \mathcal{F}) \geq r \binom{n-k-1}{k-t} + (k-t+1) \binom{n-k-2}{k-t-1} - \binom{r}{2} \binom{n-t-1}{k-t-1}$$

$$\begin{aligned}
 &= r \binom{n-k-1}{k-t} - \left(\binom{r}{2} - (k-t+1) \right) \binom{n-k-2}{k-t-1} + O(n^{k-t-2}) \\
 &> \text{dp}_t(L, \mathcal{L}).
 \end{aligned}$$

Thus we may assume $|F \cap X_i| = t - 1$ for all i . Given this condition, it follows that

$$\begin{aligned}
 &\text{dp}_t(F, \mathcal{F}(X_{i_1}) \cap \mathcal{F}(X_{i_2})) \\
 &= \begin{cases} \binom{n-k-2}{k-t-1} & \text{if } F \cap X_{i_1} = F \cap X_{i_2} \\ 0 & \text{otherwise, since } |F \cap (X_{i_1} \cup X_{i_2})| \geq t \end{cases} .
 \end{aligned}$$

Hence, in order to have $\text{dp}_t(F, \mathcal{F}) \leq \text{dp}_t(L, \mathcal{L}) = r \binom{n-k-1}{k-t} - \binom{r}{2} \binom{n-k-2}{k-t-1} + O(n^{k-t-2})$, we must have $F \cap X_{i_1} = F \cap X_{i_2}$ for all $i_1 < i_2$. This implies that F shares a common $(t - 1)$ -set with all the sets X_i , and thus $\mathcal{F} \cup \{F\}$ is isomorphic to $\mathcal{L} \cup \{L\}$, as required. ■

The final lemma showed that the union of any r full t -stars contains at least as many disjoint pairs as the initial segment of the lexicographical ordering with the same number of sets.

Proof of Lemma 4.4. We shall find it more convenient to count the number of t -intersecting pairs. Suppose \mathcal{F} is the union of the full t -stars with centers $\{X_1, X_2, \dots, X_r\} \subset \binom{[n]}{t}$. By Lemma 4.2, it follows that $\mathcal{L} = \mathcal{L}_{n,k}(|\mathcal{F}|)$ consists of the full t -stars with centers $\{Y_1, Y_2, \dots, Y_r\}$, possibly with some additional sets in an $(r + 1)$ st t -star with center Y_{r+1} , where $Y_i = \{1, 2, \dots, t - 1, t - 1 + i\}$. Note that in this setting we have $|\mathcal{F}| = |\mathcal{L}|$.

We first show that if $|X_i \cap X_j| \leq t - 2$ for some $1 \leq i < j \leq r$, then the r full t -stars of \mathcal{L} alone contain more t -intersecting pairs than \mathcal{F} . We have

$$\begin{aligned}
 (11) \quad \text{int}_t(\mathcal{F}) &\leq \sum_{i=1}^r \text{int}_t(\mathcal{F}(X_i)) + \sum_{i < j} \text{int}_t(\mathcal{F}(X_i) \setminus \mathcal{F}(X_j), \mathcal{F}(X_j) \setminus \mathcal{F}(X_i)) \\
 &= r \binom{\binom{n-t}{k-t}}{2} + r \binom{n-t}{k-t} + \sum_{i < j} \text{int}_t(\mathcal{F}(X_i) \setminus \mathcal{F}(X_j), \mathcal{F}(X_j) \setminus \mathcal{F}(X_i)),
 \end{aligned}$$

where the inequality is due to the fact that t -intersecting pairs involving sets in multiple t -stars are overcounted.

First suppose $|X_i \cap X_j| = t - 1$. Given a set $F \in \mathcal{F}(X_i) \setminus \mathcal{F}(X_j)$, we wish to bound how many sets $G \in \mathcal{F}(X_j) \setminus \mathcal{F}(X_i)$ can be t -intersecting with F . Since $X_i \cap X_j \subset F \cap G$, we require G to contain one additional element of F . However, this element cannot be from X_i , as then we would have $G \in \mathcal{F}(X_i)$. Thus there are $k - t$ choices for this additional element. Given that G already

contains X_j , there are $\binom{n-t-1}{k-t-1}$ ways to choose the remaining elements of G . Hence, there can be at most $(k-t)\binom{n-t-1}{k-t-1}$ such sets G , giving

$$\begin{aligned} \text{int}_t(\mathcal{F}(X_i) \setminus \mathcal{F}(X_j), \mathcal{F}(X_j) \setminus \mathcal{F}(X_i)) &\leq (k-t) \binom{n-t-1}{k-t-1} |\mathcal{F}(X_i) \setminus \mathcal{F}(X_j)| \\ &\leq (k-t) \binom{n-t}{k-t} \binom{n-t-1}{k-t-1}. \end{aligned}$$

Now suppose $|X_i \cap X_j| \leq t-2$. There are two types of $F \in \mathcal{F}(X_i) \setminus \mathcal{F}(X_j)$: those with $F \cap X_j = X_i \cap X_j$, and those with $(F \setminus X_i) \cap X_j \neq \emptyset$. In the first case, note that for $G \in \mathcal{F}(X_j) \setminus \mathcal{F}(X_i)$ to be t -intersecting with F , G must contain at least 2 elements from F in addition to X_j . Hence, there are at most $\binom{k}{2} \binom{n-t-2}{k-t-2}$ such sets. In the second case, note that there are at most $t \binom{n-t-1}{k-t-1}$ such sets F , as we can choose at most t elements from $X_j \setminus X_i$ for F to contain, and then there are $\binom{n-t-1}{k-t-1}$ ways to choose the remaining elements for F . For each such F , in order for $G \in \mathcal{F}(X_j) \setminus \mathcal{F}(X_i)$ to be t -intersecting with F , G must contain some element of F in addition to $F \cap X_j$. There are at most k choices for this element, with $\binom{n-t-1}{k-t-1}$ ways to complete G . Thus there are at most $kt \binom{n-t-1}{k-t-1}^2$ t -intersections of this type. Hence, we have

$$\begin{aligned} &\text{int}_t(\mathcal{F}(X_i) \setminus \mathcal{F}(X_j), \mathcal{F}(X_j) \setminus \mathcal{F}(X_i)) \\ &\leq \binom{k}{2} \binom{n-t}{k-t} \binom{n-t-2}{k-t-2} + kt \binom{n-t-1}{k-t-1}^2 = O(n^{2k-2t-2}). \end{aligned}$$

Substituting these bounds into (11), if we have d pairs $\{i, j\}$ with $|X_i \cap X_j| \leq t-2$, we have

$$(12) \quad \text{int}_t(\mathcal{F}) \leq r \binom{n-t}{2} + \left(\binom{r}{2} - d \right) (k-t) \binom{n-t}{k-t} \binom{n-t-1}{k-t-1} + O(n^{2k-2t-2}).$$

We now provide a lower bound for $\text{int}_t(\mathcal{L})$, considering only the r full t -stars. There are two types of t -intersecting pairs: those from within a single t -star, and those between two t -stars. Within each of the r t -stars, every pair of sets is t -intersecting, and hence there are at least $r \binom{n-t}{2}$ pairs of the first type.

Now we count the second type of pairs, and consider two t -stars with centers Y_i and Y_j . Since $Y_i \cap Y_j = [t-1]$, any sets $G_i \in \mathcal{L}(Y_i)$ and $G_j \in \mathcal{L}(Y_j)$ must have at least $t-1$ elements in common. They will therefore be t -intersecting if they have one more common element. Moreover, this common element should not be from $Y_i \cup Y_j$, as otherwise the two sets are in fact in

the same t -star, making this pair of the first type. For each fixed G_i , there are therefore $k - t$ choices for the final intersection, and $\binom{n-t-1}{k-t-1}$ choices for corresponding sets G_j .

There is some slight error in this calculation, as, for example, pairs involving sets contained in multiple t -stars are overcounted. However, it is easy to check that the error is of order $n^{2k-2t-2}$, and thus a lower-order term. This gives

$$\text{int}_t(\mathcal{L}) \geq r \binom{\binom{n-t}{k-t}}{2} + (k-t) \binom{r}{2} \binom{n-t}{k-t} \binom{n-t-1}{k-t-1} + O(n^{2k-2t-2}).$$

Comparing this to (12), we find that unless $d=0$, we must have $\text{int}_t(\mathcal{L}) > \text{int}_t(\mathcal{F})$, as desired. It remains to consider the case when $|X_i \cap X_j| = t - 1$ for all $1 \leq i < j \leq r$.

There are only two possibilities. In the first, all the sets X_i share $t - 1$ elements in common, in which case \mathcal{F} is isomorphic to \mathcal{L} . The second case, up to isomorphism, is when $r \leq t + 1$, and $X_i \in \binom{[t+1]}{t}$. Note that if $1 \leq r \leq 2$, the two constructions are isomorphic, so we may assume $r \geq 3$.

In this case, as we know the exact structure of both constructions, we are able to compute the number of intersecting pairs rather more precisely. We begin with \mathcal{F} , the union of r full t -stars with centers from $\binom{[t+1]}{t}$.

\mathcal{F} contains all $\binom{n-t-1}{k-t-1}$ sets containing $[t+1]$, and then $r \binom{n-t-1}{k-t}$ sets that meet $[t+1]$ in t elements. The sets containing $[t+1]$ are t -intersecting with all other sets in \mathcal{F} .

On the other hand, if $F \in \mathcal{F}(X_i)$ is such that $F \cap [t+1] = X_i$, then there are three types of sets in \mathcal{F} that can be t -intersecting with F :

- (i) a set containing $[t+1]$,
- (ii) a set whose intersection with $[t+1]$ is precisely X_i , or
- (iii) a set whose intersection with $[t+1]$ is X_j for some $j \neq i$.

There are $\binom{n-t-1}{k-t-1}$ sets of type (i) and $\binom{n-t-1}{k-t}$ sets of type (ii). For a set to be of type (iii), it must contain some X_j , not contain X_i , and then meet F in some element of $F \setminus X_i$. For each choice of j , the set should contain the t elements of X_j , not the single element in $X_i \setminus X_j$, and should not avoid the remaining $k - t$ elements of F . Hence there are $\binom{n-t-1}{k-t} - \binom{n-k-1}{k-t}$ such sets.

Putting this all together, we find

$$(13) \quad 2 \text{int}_t(\mathcal{F}) - |\mathcal{F}| = \sum_{F \in \mathcal{F}} \text{int}_t(F, \mathcal{F}) = I_1 + I_2,$$

where

$$I_1 = \binom{n-t-1}{k-t-1} \left[\binom{n-t-1}{k-t-1} + r \binom{n-t-1}{k-t} \right],$$

and

$$I_2 = r \binom{n-t-1}{k-t} \left[\binom{n-t-1}{k-t-1} + \binom{n-t-1}{k-t} + (r-1) \left(\binom{n-t-1}{k-t} - \binom{n-k-1}{k-t} \right) \right].$$

We now turn our attention to \mathcal{L} .

First observe that we have r full stars, with centers $\{Y_1, Y_2, \dots, Y_r\}$. The remaining sets fall into an $(r+1)$ st star with center Y_{r+1} . To avoid overcounting, we shall partition \mathcal{L} into the subsystems $\mathcal{L}^*(i) = \{L \in \mathcal{L} : \min(L \setminus [t-1]) = t-1+i\}$, $1 \leq i \leq r+1$; that is, $L \in \mathcal{L}^*(i)$ if $\mathcal{L}(Y_i)$ is the first t -star containing L .

For $1 \leq i \leq r$, $\mathcal{L}^*(i)$ consists of all sets containing $[t-1] \cup \{t-1+i\}$, but disjoint from the interval $[t, t-2+i]$. Hence we have $|\mathcal{L}^*(i)| = \binom{n-t-i+1}{k-t}$. Summing up the telescoping binomial coefficients, we find the first r t -stars contain $\binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1}$ sets. $\mathcal{L}^*(r+1)$ then contains enough sets to make \mathcal{L} the right size, and so

$$|\mathcal{L}^*(r+1)| = |\mathcal{F}| - |\cup_{i=1}^r \mathcal{L}^*(i)| = \left[\binom{n-t-1}{k-t-1} + r \binom{n-t-1}{k-t} \right] - \left[\binom{n-t+1}{k-t+1} - \binom{n-t-r+1}{k-t+1} \right].$$

Note that all the subsystems $\mathcal{L}^*(i)$ are t -intersecting. Moreover, if $j < i$, and $L \in \mathcal{L}^*(i)$, then for a set $K \in \mathcal{L}^*(j)$ to be t -intersecting with L , it must contain $[t-1] \cup \{t-1+j\}$, be disjoint from the interval $[t, t-2+j]$, and contain one of the $k-t+1$ elements in $L \setminus [t-1+j]$. Hence we have $\text{int}_t(L, \mathcal{L}^*(j)) = \binom{n-t-j+1}{k-t} - \binom{n-k-j}{k-t}$. We can now count the number of t -intersecting pairs in \mathcal{L} :

$$\begin{aligned} 2 \text{int}_t(\mathcal{L}) - |\mathcal{L}| &= \sum_{L \in \mathcal{L}} \text{int}_t(L, \mathcal{L}) \\ (14) \quad &= \sum_{i=1}^{r+1} \left(\text{int}_t(\mathcal{L}^*(i), \mathcal{L}^*(i)) + 2 \sum_{j < i} \text{int}_t(\mathcal{L}^*(i), \mathcal{L}^*(j)) \right) \\ &= \sum_{i=1}^{r+1} |\mathcal{L}^*(i)|^2 + 2 \sum_{j < i} |\mathcal{L}^*(i)| \left[\binom{n-t-j+1}{k-t} - \binom{n-k-j}{k-t} \right]. \end{aligned}$$

We now wish to show $\text{int}_t(\mathcal{L}) \geq \text{int}_t(\mathcal{F})$; that is, to show the quantity in (14) is greater than that in (13). To make this task easier, we shall

rewrite all products of binomial coefficients in the form $\binom{n-t}{k-t}^2$, $\binom{n-t}{k-t}\binom{n-t}{k-t-1}$, or $\binom{n-t}{k-t-1}^2$, using the identities

$$\binom{m-a}{r} = \binom{m}{r} - a\binom{m}{r-1} + \binom{a+1}{2}\binom{m}{r-2} + O(m^{r-3})$$

and

$$\begin{aligned} \binom{n-t}{k-t}\binom{n-t}{k-t-2} &= \frac{n-k+1}{n-k+2} \cdot \frac{k-t-1}{k-t} \cdot \binom{n-t}{k-t-1}^2 \\ &= \frac{k-t-1}{k-t} \binom{n-t}{k-t-1}^2 + O(n^{2k-2t-3}). \end{aligned}$$

After performing the routine but tedious calculations, we find

$$\begin{aligned} 2\text{int}_t(\mathcal{F}) - |\mathcal{F}| &= r\binom{n-t}{k-t}^2 + r(r-1)(k-t)\binom{n-t}{k-t}\binom{n-t}{k-t-1} \\ &- \left[\frac{1}{2}r(r-1)(k-t)^2 + 2r(r-1)(k-t) - \left(\frac{3r}{2} - 1\right)(r-1) \right] \binom{n-t}{k-t-1}^2 \\ &\quad + O(n^{2k-2t-3}), \end{aligned}$$

and

$$\begin{aligned} 2\text{int}_t(\mathcal{L}) - |\mathcal{L}| &= r\binom{n-t}{k-t}^2 + r(r-1)(k-t)\binom{n-t}{k-t}\binom{n-t}{k-t-1} \\ &- \left[\frac{1}{2}r(r-1)(k-t)^2 + 2r(r-1)(k-t) - \frac{1}{4}(r-1)^2(r^2+4) \right] \binom{n-t}{k-t-1}^2 \\ &\quad + O(n^{2k-2t-3}). \end{aligned}$$

The coefficient of the leading term of the difference between the two constructions is thus

$$\begin{aligned} \frac{2\text{int}_t(\mathcal{L}) - 2\text{int}_t(\mathcal{F})}{\binom{n-t}{k-t-1}^2} - O\left(\frac{1}{n}\right) &= \frac{1}{4}(r-1)^2(r^2+4) - \left(\frac{3r}{2} - 1\right)(r-1) \\ &= \frac{1}{4}(r+1)r(r-1)(r-2) \geq 6 \end{aligned}$$

for $r \geq 3$. Hence, we indeed find $\text{int}_t(\mathcal{L}) > \text{int}_t(\mathcal{F})$, as required. ■

5. Concluding remarks and open problems

In this paper, we have provided a partial solution to a problem of Ahlswede on the minimum number of disjoint pairs in set systems. For small systems, we verified Bollobás and Leader’s conjecture by showing that the initial segment of the lexicographical ordering is optimal. By considering the complementary set systems, this also resolves the problem for very large set systems. However, it remains to determine which systems are optimal in between.

When $k=2$, Ahlswede and Katona showed that the optimal system was always either a union of stars or its complement. For $k \geq 3$, Bollobás and Leader suggest a larger family of possible extremal systems. We note that for systems of size $s = \frac{1}{2} \binom{n}{k}$, the lexicographical system is at least near-optimal. A straightforward calculation shows $dp(n, k, s) \leq dp(\mathcal{L}_{n,k}(s)) \leq \frac{1}{2} \left(1 - \frac{2^{1/k} k^2}{n} + O(n^{-2}) \right) s^2$. On the other hand, exploiting the connection to the Kneser graph, we can use spectral techniques to obtain the bound $dp(n, k, s) \geq \frac{1}{2} \left(1 - \frac{k(k+2)}{n} + O(n^{-2}) \right) s^2$.

While our focus has been showing that a system with more than $\binom{n-1}{k-1}$ sets must contain many disjoint pairs, a closely related problem is to determine whether such a system must have any sets disjoint from many other sets. This type of question has been studied before in other settings. For example, when one is considering the number of triangles in a graph, Erdős showed in [11] that any graph with $\lfloor \frac{n^2}{4} \rfloor + 1$ edges must contain an edge in at least $\frac{n}{6} + o(n)$ triangles. It is well-known and easy to see that the hypercube, a graph whose vertices are subsets of $[n]$, with two vertices adjacent if they are comparable and differ in exactly one element, has independence number 2^{n-1} . In [8], it is proved that any induced subgraph on $2^{n-1} + 1$ vertices contains a vertex of degree at least $(\frac{1}{2} + o(1)) \log_2 n$. It is an open problem to determine whether or not this bound is tight (the corresponding upper bound is $O(\sqrt{n})$), and the answer to this question has ramifications in theoretical computer science.

In the context of the Erdős–Ko–Rado Theorem, it is trivial to show that in a system of $\binom{n-1}{k-1} + 1$ sets, there must be a set disjoint from at least $\frac{1}{2} \left(1 - \frac{k^3}{n} \right) \binom{n-1}{k-1}$ other sets. Indeed, by the Erdős–Ko–Rado Theorem, there exists a pair F_1, F_2 of disjoint sets. At most $k^2 \binom{n-2}{k-2} < \frac{k^3}{n} \binom{n-1}{k-1}$ sets can intersect both F_1 and F_2 , and so either F_1 or F_2 must be disjoint from at least half of the remaining sets, resulting in the above bound. Furthermore, this is easily seen to be asymptotically tight, as one may take all sets contain-

ing $\{1, 2\}$, and then take half the remaining sets to contain 1, and half to contain 2. It may be of interest to obtain sharper estimates for this problem, especially as the aforementioned construction shows that this is closely related to the original problem when $s \approx \frac{1}{2} \binom{n}{k}$, since one should choose the sets containing 1 or 2 optimally. A related problem, that of determining the largest possible system where no set is disjoint from more than ℓ other sets, was studied in [18].

We find most exciting the prospect of studying Erdős–Rademacher-type problems in other settings. In an earlier paper [9], we presented an Erdős–Rademacher-type strengthening of Sperner’s Theorem, a problem that was also studied in [10]. However, as one can investigate similar extensions for any extremal result, there is truly no end to the number of directions in which this project can be continued. We hope that further work of this nature will lead to many interesting results and a greater understanding of classical theorems in extremal combinatorics.

Acknowledgment. We would like to thank the anonymous referees for their valuable comments and suggestions to improve the presentation of this paper.

References

- [1] R. AHLWEDE: Simple hypergraphs with maximal number of adjacent pairs of edges, *J. Combinatorial Theory (B)* **28** (1980), 164–167.
- [2] R. AHLWEDE and N. CAI: A counterexample to Kleitman’s conjecture concerning an edge-isoperimetric problem, *Combinatorics, Probability and Computing* **8** (1999), 301–305.
- [3] R. AHLWEDE and N. CAI: Appendix: on edge-isoperimetric theorems for uniform hypergraphs, *General Theory of Information Transfer and Combinatorics*, Springer Berlin Heidelberg (2006), 979–1005.
- [4] R. AHLWEDE and G. O. H. KATONA: Graphs with maximal number of adjacent pairs of edges, *Acta Mathematica Hungarica* **32** (1978), 97–120.
- [5] R. AHLWEDE and L. H. KHACHATRIAN: The complete intersection theorem for systems of finite sets, *European Journal of Combinatorics* **18** (1997), 125–136.
- [6] I. ANDERSON: *Combinatorics of Finite Sets*, Courier Dover Publications, 1987.
- [7] B. BOLLOBÁS and I. LEADER: Set systems with few disjoint pairs, *Combinatorica* **23** (2003), 559–570.
- [8] F. R. K. CHUNG, Z. FÜREDI, R. L. GRAHAM and P. SEYMOUR: On induced subgraphs of the cube, *J. Combinatorial Theory (A)* **49** (1988), 180–187.
- [9] S. DAS, W. GAN and B. SUDAKOV: Sperner’s theorem and a problem of Erdős, Katona and Kleitman, *Combinatorics, Probability and Computing* **24** (2015), 585–608.
- [10] A. P. DOVE, J. R. GRIGGS, R. J. KANG and J. S. SERENI: Supersaturation in the Boolean lattice, *Integers* **14A** (2014), #A4.

- [11] P. ERDŐS: On a theorem of Rademacher–Turán, *Illinois Journal of Math* **6** (1962), 122–127.
- [12] P. ERDŐS: On the number of complete subgraphs contained in certain graphs, *Magy. Tud. Akad. Mat. Kut. Int. Közl.* **7** (1962), 459–474.
- [13] P. ERDŐS: A problem on independent r -tuples, *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.* **8** (1965), 93–95.
- [14] P. ERDŐS, C. KO and R. RADO: Intersection theorems for systems of finite sets, *The Quarterly Journal of Mathematics* **12** (1961), 313–320.
- [15] P. FRANKL: On the minimum number of disjoint pairs in a family of finite sets, *J. Combinatorial Theory (A)* **22** (1977), 249–251.
- [16] P. FRANKL: Improved bounds for Erdős’ Matching Conjecture, *J. Combinatorial Theory (A)* **120** (2013), 1068–1072.
- [17] P. FRANKL, V. RÖDL and A. RUCIŃSKI: On the maximum number of edges in a triple system not containing a disjoint family of a given size, *Combinatorics, Probability and Computing* **21** (2012), 141–148.
- [18] D. GERBNER, N. LEMONS, C. PALMER, B. PATKÓS and V. SZÉCSI: Almost intersecting families of sets, *SIAM Journal of Discrete Math.* **26** (2012), 1657–1669.
- [19] H. HUANG, P. LOH and B. SUDAKOV: The size of a hypergraph and its matching number, *Combinatorics, Probability and Computing* **21** (2012), 442–450.
- [20] G. O. H. KATONA, G. Y. KATONA and ZS. KATONA: Most probably intersecting families of subsets, *Combinatorics, Probability and Computing* **21** (2012), 219–227.
- [21] T. ŁUCZAK and K. MIECZKOWSKA: On Erdős’ extremal problem on matchings in hypergraphs, *J. Combinatorial Theory (A)* **124** (2014), 178–194
- [22] W. MANTEL: Problem 28, *Wiskundige Opgaven* **10** (1907), 60–61.
- [23] S. POLJAK and ZS. TUZA: Maximum bipartite subgraphs of Kneser graphs, *Graphs and Combinatorics* **3** (1987), 191–199.
- [24] R. M. WILSON: The exact bound on the Erdős–Ko–Rado Theorem, *Combinatorica* **4** (1984), 247–257.

Shagnik Das

*Department of Mathematics and
Computer Science
Freie Universität Berlin
14195 Berlin, Germany
shagnik@mi.fu-berlin.de*

Wenyang Gan

*Department of Mathematics
UCLA
Los Angeles, CA, 90095
wgan@ucla.edu*

Benny Sudakov

*Department of Mathematics
ETH Zürich
8092 Zürich, Switzerland
benjamin.sudakov@math.ethz.ch*