

Minimal Paths and Cycles in Set Systems

Dhruv Mubayi *

Jacques Verstraëte †

July 9, 2006

Abstract

A minimal k -cycle is a family of sets A_0, \dots, A_{k-1} for which $A_i \cap A_j \neq \emptyset$ if and only if $i = j$ or i and j are consecutive modulo k . Let $f_r(n, k)$ be the maximum size of a family of r -sets of an n element set containing no minimal k -cycle. Our results imply that for fixed $r, k \geq 3$,

$$\ell \binom{n-1}{r-1} + O(n^{r-2}) \leq f_r(n, k) \leq 3\ell \binom{n-1}{r-1} + O(n^{r-2}),$$

where $\ell = \lfloor (k-1)/2 \rfloor$. We also prove that $f_r(n, 4) = (1 + o(1)) \binom{n-1}{r-1}$ as $n \rightarrow \infty$. This supports a conjecture of Füredi [9] on families in which no two pairs of disjoint sets have the same union.

1 Introduction.

In this paper, we are interested in a generalization to hypergraphs of the extremal theory of paths and cycles in graphs. Following Berge [1], a k -cycle is a set system $\{A_i : i \in \mathbf{Z}_k\}$ such that the family $\{A_i \cap A_{i+1} : i \in \mathbf{Z}_k\}$ has a system of distinct representatives. It is convenient to represent a k -cycle $\{A_0, A_1, \dots, A_{k-1}\}$ as an ordered list $A_0 A_1 A_2 \dots A_{k-1} A_0$. A *minimal* k -cycle is a k -cycle $A_0 A_1 \dots A_{k-1} A_0$ such that $A_i \cap A_j \neq \emptyset$ if and only if $i = j$ or i and j are consecutive modulo k . A minimal k -path is a family of sets $\{A_0, A_1, \dots, A_{k-1}\}$ such that $A_i \cap A_j \neq \emptyset$ if and only if $|i - j| \leq 1$. In other words, no vertex in a minimal k -path or k -cycle belongs to two non-consecutive sets. We write \mathcal{C}_k and \mathcal{P}_k for the family of minimal k -cycles and k -paths, respectively. For example any path, in the traditional sense of graphs, is a minimal path.

Given vertices x, y in a set system, consider the minimum integer k such that there is a k -path $\{A_0, A_1, \dots, A_{k-1}\}$ with $x \in A_0$ and $y \in A_{k-1}$ (we call this a shortest xy -path). Then a shortest xy -path in a set system is a minimal path, and the shortest paths describe a natural notion of distance between vertices of a set system. Our interest lies in extremal problems for these structures, in particular, the determination of $\text{ex}_r(n, \mathcal{C}_k)$ ($\text{ex}_r(n, \mathcal{P}_k)$), defined as the maximum size of a family of r -element sets of an n element set containing no minimal k -cycle (k -path). One of the first noteworthy results in this direction is due to Erdős and Gallai [4], who proved the following theorem:

Theorem 1.1 (Erdős-Gallai [4]) *Let n and k be positive integers and suppose $n = qk + r$, where $0 \leq r < k$. Then $\text{ex}_2(n, \mathcal{P}_k) = q \binom{k}{2} + \binom{r}{2}$.*

*Department of Mathematical Statistics and Computer Science, University of Illinois, Chicago, Illinois 60607, USA; research partially supported by NSF grants DMS 9970325, 0400812, and an Alfred P. Sloan Research Fellowship. E-mail: mubayi@math.uic.edu

†Department of Combinatorics and Optimization, University of Waterloo, Waterloo, ON N2L 3G1, Canada. E-mail: jverstra@math.uwaterloo.ca

The extremal problem for k -cycles in graphs (which is the same as minimal k -cycles) is a notorious open problem in combinatorics when k is even (see [2] for history and results), and completely solved when k is odd [3, 17]. The situation for hypergraphs is somewhat different. For example, the following theorem about triangles (or minimal 3-cycles), solves an old problem of Erdős, and was proved by the current authors only recently.

Theorem 1.2 ([15]) *Let $r \geq 3$ and $n \geq 3r/2$. Then $\text{ex}_r(n, \mathcal{C}_3) = \binom{n-1}{r-1}$.*

Recall the Erdős-Ko-Rado theorem [6] which states that for $n \geq 2r$, the maximum size of an intersecting family of r -element sets on an n element set is $\binom{n-1}{r-1}$. We will deduce from the Erdős-Ko-Rado theorem that the extremal set system for a minimal path of length three is the same as that in Theorem 1.2. More generally, we will prove the following:

Theorem 1.3 *Let $r, k \geq 3, \ell = \lfloor \frac{k-1}{2} \rfloor$ and $n \geq (k+1)r/2$. Then*

$$\text{ex}_r(n, \mathcal{P}_k) \leq \begin{cases} \binom{n-1}{r-1} & \text{for } k = 3 \\ \frac{5k-1}{6} \binom{n}{2} & \text{for } r = 3 \\ 2\ell \binom{n-1}{r-1} + O(n^{r-2}) & \text{for } k, r > 3. \end{cases}$$

The result is sharp for $k = 3$.

Although the same problem for minimal k -cycles seems more difficult, our results are similar. Note that the case $k = 3$ is already completely solved by Theorem 1.2.

Theorem 1.4 *Let $r \geq 3, k \geq 4, \ell = \lfloor \frac{k-1}{2} \rfloor$ and $n \geq kr/2$. Then*

$$\text{ex}_r(n, \mathcal{C}_k) \leq \begin{cases} \frac{5k-1}{6} \binom{n}{2} & \text{for } r = 3 \\ \frac{5k}{4} \binom{n}{3} & \text{for } r = 4 \\ 3\ell \binom{n-1}{r-1} + O(n^{r-2}) & \text{for } r > 4. \end{cases}$$

The following construction shows that the theorems above are not far from optimal. Let $\mathcal{A}[k, r]$ be a set system on an n -element set X , constructed as follows: put $\ell = \lfloor \frac{k-1}{2} \rfloor$, let L be an ℓ -element subset of X , and take all r -sets in X which contain at least one vertex of L . It is straightforward to verify that $\mathcal{A}[k, r]$ contains no minimal k -cycle, since for any set of ℓ vertices in a minimal k -cycle, there is a set in the cycle disjoint from these vertices, whereas no set in $\mathcal{A}[k, r]$ is disjoint from L . Furthermore

$$|\mathcal{A}[k, r]| = \binom{n}{r} - \binom{n-\ell}{r} \geq \ell \binom{n-1}{r-1} + O(n^{r-2}).$$

The foregoing discussion suggests the following conjecture:

Conjecture 1.5 *Let $n, k, r \geq 3$ be integers and $\ell = \lfloor \frac{k-1}{2} \rfloor$. Then, as $n \rightarrow \infty$,*

$$\text{ex}_r(n, \mathcal{C}_k) = \ell \binom{n-1}{r-1} + O(n^{r-2}).$$

In fact, it would be very interesting to determine the exact value of $\text{ex}(n, \mathcal{C}_k)$. Conjecture 1.5 holds for $k = 3$ by Theorem 1.2. We also settle the next case, namely $k = 4$. This problem was studied in [13] in the context of simple hypergraphs, where the asymptotics for the extremal function were determined. Another related extremal problem on minimal 4-cycles was posed in ([7], page 191). It appears there in the context of Natarajan dimension, itself a generalization of the well studied notion of VC dimension.

The extremal problem for minimal 4-cycles is also a relaxation of the question of estimating the maximum size $f_r(n)$ of a family of r -element sets on an n element set containing no two pairs of disjoint r -element sets with the same union. Since all the forbidden configurations in this question are minimal 4-cycles, $\text{ex}_r(n, \mathcal{C}_4) \leq f_r(n)$. Answering a question of Erdős, Füredi [9] proved that $f_r(n) \leq \frac{7}{2} \binom{n}{r-1}$. The authors [14] slightly improved Füredi's result by showing that $f_r(n) < 3 \binom{n}{r-1}$. Füredi further conjectured that $f_3(n) = \binom{n}{2}$ for infinitely many n and $f_r(n) = \binom{n-1}{r-1} + \lfloor \frac{n-1}{r} \rfloor$ for all sufficiently large n . We prove the following result, which supports Füredi's conjecture:

Theorem 1.6 *Let $r \geq 3$ be fixed. Then*

$$\binom{n-1}{r-1} + \left\lfloor \frac{n-1}{r} \right\rfloor \leq \text{ex}_r(n, \mathcal{C}_4) \leq \binom{n-1}{r-1} + O(n^{r-2}).$$

Notation and Terminology

We follow Bollobás [2] for set system notations, such as $X^{(r)}$ for the set system of all r -sets in X , \mathcal{A}_x for the family of sets in a set system \mathcal{A} containing x , and $d(x) = |\mathcal{A}_x|$. The residue of a set $S \subset X$ in a set system \mathcal{A} on X is defined by $\text{res}S = \{A \subset X \setminus S : A \cup S \in \mathcal{A}\}$. We write $d(S) = |\text{res}S|$.

An r -partite set system \mathcal{A} on X is a set system for which X has a partition (X_1, X_2, \dots, X_r) into parts X_i such that $|A \cap X_i| = 1$ for all $A \in \mathcal{A}$ and $1 \leq i \leq r$. In this case, we consider \mathcal{A} as a subset of the cartesian product $\prod_{i=1}^r X_i$, which denotes the complete r -partite system with parts X_1, X_2, \dots, X_r . For convenience, this set system is denoted $\prod^{(r)}$, when the parts X_1, X_2, \dots, X_r are specified.

If \mathcal{F} is a family of set systems and \mathcal{G} is a fixed set system, then $\text{ex}(\mathcal{G}, \mathcal{F})$ denotes the maximum size of a subsystem of \mathcal{G} that does not contain a copy of any member of \mathcal{F} . When $\mathcal{G} = [n]^{(r)}$, we write $\text{ex}_r(n, \mathcal{F})$, or $\text{ex}(n, \mathcal{F})$ when clear from context. Any set system achieving this maximum is called an extremal set system for \mathcal{F} , and any set system not containing any member \mathcal{F} is called \mathcal{F} -free. If some set system in \mathcal{F} is r -partite, then we write $z(\prod^{(r)}, \mathcal{F})$ for the maximum size of an \mathcal{F} -free set system with parts X_1, X_2, \dots, X_r ; when $|X_i| = n$ for all i , we again abbreviate as $z_r(n, \mathcal{F})$ or $z(n, \mathcal{F})$.

2 Proof Techniques

The aim of this section is to give an idea of how the theorems on minimal paths and cycles stated in the introduction will be proved. The first two inequalities of Theorems 1.3 and 1.4 will be proved separately. The last inequalities in Theorems 1.3 and 1.4 are implied by the following:

Theorem 2.1 *Let $n, k \geq 3$. Then*

$$\begin{aligned} z_r(n, \mathcal{P}_k) &\leq 2ln^{r-1} && \text{for } r \geq 2 \\ z_r(n, \mathcal{C}_k) &\leq 3ln^{r-1} && \text{for } r \geq 3. \end{aligned}$$

To establish this implication, we first consider the more general extremal problem for families which are said to be closed under extension (Section 2.1), and then we show how to reduce our extremal problems to the setting of r -partite set systems (Section 2.2). In Sections 3–5, we will use these ideas to prove our results.

2.1 Families Closed Under Extension

Let W be a part of an r -partite set system \mathcal{A} on X , let Y be a set disjoint from X with $|Y| = |W|$, and let $\phi : W \rightarrow Y$ be a bijection. The extension of \mathcal{A} from W to Y under ϕ is the $(r+1)$ -partite system

$$\mathcal{A}^\phi = \bigcup_{w \in W} \{A \cup \phi(w) : A \in \mathcal{A}_w\}.$$

In other words, we extend each set $A \in \mathcal{A}$ to a new part Y by deciding, using ϕ , which point from Y to add to A . We say that a family \mathcal{F} of set systems is closed under extension if the following holds for every r : for every r -partite set system $\mathcal{A} \in \mathcal{F}$ with ground set X , and every bijection ϕ from a part W of \mathcal{A} to a set Y disjoint from X , $\mathcal{A}^\phi \in \mathcal{F}$. The notion of extension is useful in conjunction with the following inductive lemma:

Lemma 2.2 *Let X_1, X_2, \dots, X_{r+1} be disjoint sets each of size n , and let \mathcal{F} be a family of set systems which is closed under extension. Then, for any integer $r \geq 2$,*

$$z(\Pi^{(r+1)}, \mathcal{F}) \leq n \cdot z(\Pi^{(r)}, \mathcal{F}).$$

Proof. Let $\mathcal{B} \subset \Pi^{(r+1)}$ be an \mathcal{F} -free set system with parts X_1, X_2, \dots, X_{r+1} and $|\mathcal{B}| = z(\Pi^{(r+1)}, \mathcal{F})$. Let $\phi_1, \phi_2, \dots, \phi_n$ denote bijections $X_r \rightarrow X_{r+1}$ such that $\phi_i(x) \neq \phi_j(x)$ for all $x \in X_r$ and for $1 \leq i < j \leq n$. Then

$$\sum_{i \leq n} \sum_{x \in X_r} |\text{res}\{x, \phi_i(x)\}| = |\mathcal{B}|.$$

By the Pigeonhole Principle, there exists a bijection $\phi \in \{\phi_1, \phi_2, \dots, \phi_n\}$ such that

$$\sum_{x \in X_r} |\text{res}\{x, \phi(x)\}| \geq \frac{|\mathcal{B}|}{n}.$$

Now let \mathcal{B}' denote the family of sets $B \in \mathcal{B}$ which contain some pair $\{x, \phi(x)\}$, and let

$$\mathcal{A} = \bigcup_{x \in X_r} \text{res}\{\phi(x)\},$$

where the residue is taken in \mathcal{B}' . By the above inequality, $|\mathcal{B}'| \geq |\mathcal{B}|/n$. Also, $\mathcal{B}' = \mathcal{A}^\phi$, and \mathcal{A} is r -partite and of size $|\mathcal{B}'| \geq |\mathcal{B}|/n$. If \mathcal{A} contains an r -partite $\mathcal{G} \in \mathcal{F}$, then $\mathcal{G}^\phi \in \mathcal{F}$, since \mathcal{F} is closed under extension. But then

$$\mathcal{G}^\phi \subset \mathcal{A}^\phi = \mathcal{B}' \subset \mathcal{B},$$

which is a contradiction. This shows that \mathcal{A} contains no r -partite member of \mathcal{F} , and

$$\frac{1}{n}z(\Pi^{(r+1)}, \mathcal{F}) \leq |\mathcal{A}| \leq z(\Pi^{(r)}, \mathcal{F}).$$

■

2.2 Partitioning Set Systems

The following lemma of Erdős and Kleitman's [5] establishes an explicit relationship between $\text{ex}_r(n, \mathcal{F})$ and $z_r(n, \mathcal{F})$, when the family \mathcal{F} contains an r -partite set system. We omit the proof, which is now a standard counting argument.

Lemma 2.3 *Let \mathcal{A} be a system of r -sets in X , where $|X| = rn$. Then \mathcal{A} contains an r -partite set system \mathcal{B} in which each part has size n and such that $|\mathcal{B}| \geq (r!/r^r) \cdot |\mathcal{A}|$.*

Corollary 2.4 *Let \mathcal{F} be a family of set systems containing an r -partite set system, let X_1, \dots, X_r be disjoint sets of the same size, and $X = \bigcup X_i$. Then*

$$\text{ex}(X^{(r)}, \mathcal{F}) \leq \frac{r^r}{r!} \cdot z(\Pi^{(r)}, \mathcal{F}).$$

In particular, $\text{ex}_r(tr, \mathcal{F}) \leq (r^r/r!)z_r(t, \mathcal{F})$.

Proof. Let \mathcal{A} be an \mathcal{F} -free family of r -sets in X , such that $|\mathcal{A}| = \text{ex}(X^{(r)}, \mathcal{F})$. By Lemma 2.3, \mathcal{A} contains an r -partite set system, \mathcal{B} , of size at least $(r!/r^r) \cdot \text{ex}(X^{(r)}, \mathcal{F})$. Clearly $|\mathcal{B}| \leq z(\Pi^{(r)}, \mathcal{F})$, from which Corollary 2.4 follows. ■

Proposition 2.5 *The last bounds of Theorems 1.3 and 1.4 are implied by Theorem 2.1 for $r = 2$ and $r = 3$, respectively.*

Proof. It is straightforward to verify that the families \mathcal{P}_k and \mathcal{C}_k of minimal k -paths and minimal k -cycles are both closed under extension. The proof of the proposition is the same for paths and cycles, so we just consider the case of cycles. Suppose Theorem 2.1 has been proved for $r = 3$. Let $n' = tr$ be the smallest integer greater than n that is divisible by r . By Lemma 2.2, applied with $\mathcal{F} = \mathcal{C}_k$,

$$z_r(t, \mathcal{C}_k) \leq t^{r-3}z_3(t, \mathcal{C}_k) \leq t^{r-3}3lt^2 = 3lt^{r-1}.$$

Now Corollary 2.4 implies that

$$\text{ex}_r(n, \mathcal{C}_k) \leq \text{ex}_r(n', \mathcal{C}_k) \leq \frac{r^r}{r!} z_r(t, \mathcal{C}_k) \leq \frac{r^r}{r!} 3\ell t^{r-1} = 3\ell \binom{n-1}{r-1} + O(n^{r-2}).$$

This completes the proof. ▀

3 Minimal Paths

In this section we prove the first and last inequalities in Theorem 1.3. We will verify Theorem 1.3 for $r = 3$ and Theorem 1.4 simultaneously at the end of Section 4.

Proof of Theorem 1.3 for $k = 3$ and for $k, r > 3$.

Case 1 : $k = 3$. Let $|X| = n$ and $\mathcal{A} \subset X^{(r)}$ be a \mathcal{P}_3 -free family of maximum size. Let K_1, \dots, K_t be the components of \mathcal{A} , where K_i has n_i vertices, and suppose $n_1, n_2, \dots, n_s \leq 2r - 1$ and $n_{s+1}, n_{s+2}, \dots, n_t \geq 2r$.

We first claim that $|K_i| \leq \binom{n_i-1}{r-1}$ for $s+1 \leq i \leq t$. To see this we observe that if $|K_i| > \binom{n_i-1}{r-1}$, then by the Erdős-Ko-Rado theorem, there are disjoint sets A_1, A_2 in K_i . Since K_i is a component, we may select two points in A_1 and A_2 which are the endpoints of a minimal path \mathcal{P} . If the path has length at least two, then $\mathcal{P} \cup \{A_i\}$ contains a minimal path of length three for some $i \in \{1, 2\}$. If \mathcal{P} has length one, then $\mathcal{P} \cup \{A_1, A_2\}$ is a minimal path of length three in \mathcal{A} . Therefore $|K_i| \leq \binom{n_i-1}{r-1}$ for $s+1 \leq i \leq t$. It follows that

$$|\mathcal{A}| \leq \sum_{i=1}^s \binom{n_i}{r} + \sum_{i=s+1}^t \binom{n_i-1}{r-1}.$$

The next claim is that this is less than $\binom{n-1}{r-1}$ unless \mathcal{A} has only one component. To prove this claim, let A_1, A_2, \dots, A_s be sets of size n_1, n_2, \dots, n_s and let $A_{s+1}, A_{s+2}, \dots, A_t$ be sets of size $n_{s+1} - 1, n_{s+2} - 1, \dots, n_{t-1}$, respectively, chosen so that $A_p \cap A_q = \{x\}$ for all $p \neq q$. Then the sum above is precisely sum of the number of $(r-1)$ -element subsets of each A_i . Furthermore, since $r \geq 3$, there is an $(r-1)$ -set in $A_1 \cup A_2 \cup \dots \cup A_t$ which is not contained in any A_i , and this union has size at most $n - 1$ if \mathcal{A} has more than one component. Finally, if \mathcal{A} has one component, then $|\mathcal{A}| \leq \binom{n-1}{r-1}$ if $n \geq (k+1)r/2 = 2r$. This completes the proof for $k = 3$.

Case 2 : $k, r > 3$. By Proposition 2.5, it suffices to prove that $z_2(n, \mathcal{P}_k) \leq 2ln$. In fact, we will show more generally that a bipartite graph \mathcal{H} with parts X and Y of sizes m and n , respectively, and with $|\mathcal{H}| > (m+n)\ell$, contains a k -path. We proceed by induction on $m+n+k$. If $m \leq \ell$ or $n \leq \ell$, the claim is vacuously true, since $|\mathcal{H}| > mn$ in this case. Suppose $m > \ell$ and $n > \ell$. If $d(x) \leq \ell$ in \mathcal{H} , for some $x \in X \cup Y$, then

$$\mathcal{H}' = \{E \in \mathcal{H} : x \notin E\}$$

has size more than $(m+n-1)\ell$ and therefore contains a k -path, by induction. So we may assume $d(x) > \ell$ for all $x \in X \cup Y$.

By induction on $m+n+k$, \mathcal{H} contains a $(k-1)$ -path $\mathcal{P} = \{A_1, A_2, \dots, A_{k-1}\}$. The endvertices of a path are those vertices x such that $d(x) = 1$. If there exists an edge $E = \{x, y\}$ where x is an endvertex of \mathcal{P} and $y \notin \bigcup A_i$, then $\mathcal{P} \cup \{E\}$ is a k -path in \mathcal{H} , as required. Suppose this is not the case. Then, as $d(x) > \ell$ for each endvertex x of \mathcal{P} , $|\mathcal{P}| \geq 2\ell + 1$ if k is even and $|\mathcal{P}| \geq 2\ell + 2 \geq k$ if k is odd. In the latter case, we have the required minimal k -path in \mathcal{H} , namely \mathcal{P} . In the case k is even, let x and y be the endvertices of \mathcal{P} , and suppose $|\mathcal{P}| = 2\ell + 1$. Then $d(x) > \ell$ and $d(y) > \ell$ imply $\{x, y\} \in \mathcal{H}$. Therefore $\mathcal{P} \cup \{\{x, y\}\}$ is a minimal k -cycle $\mathcal{C} \subset \mathcal{H}$.

Let $V = \bigcup \mathcal{C}$. If there exists $\{u, v\} \in \mathcal{H}$ with $v \in V$ and $u \notin V$, then let $\mathcal{Q} \subset \mathcal{C}$ be a minimal $(k-1)$ -path such that v is an endvertex of \mathcal{Q} . It follows that $\mathcal{Q} \cup \{\{u, v\}\}$ is a minimal k -path in \mathcal{H} , as required. So we assume that for all $E \in \mathcal{H}$, $E \subset V$ or $E \cap V = \emptyset$. In particular,

$$\mathcal{G} = \{E \in \mathcal{H} : E \cap V = \emptyset\}$$

has size at least

$$(m+n)\ell - (\ell+1)^2 > (m+n-2(\ell+1))\ell.$$

By induction, \mathcal{G} contains a k -path. This completes the proof of Theorem 1.3 for $k > 3$. ▀

Remark 3.1 *The proof of $z_2(X \times Y, \mathcal{P}_k) \leq (m+n)\ell$ for X, Y of size m, n respectively was actually only needed for the case when k is even. When k is odd, Theorem 1.1 implies that $\text{ex}_2(n, \mathcal{P}_k) \leq \ell n$, and so one immediately gets $z_2(n, \mathcal{P}_k) \leq \text{ex}_2(2n, \mathcal{P}_k) \leq 2\ell n$ which is the bound we sought.*

4 Minimal Cycles

The aim of this section is to prove Theorem 1.4. We already saw that it is sufficient to prove Theorem 2.1 in the case $r = 3$ for cycles. Before we prove Theorem 2.1, we require a lemma. The following definition is needed for this lemma: given a graph G and a collection of sets $\mathcal{S} = \{S_v : v \in V(G)\}$ indexed by the vertex set of G , a G -system of representatives for \mathcal{S} is a multiset $X = \{x_v : v \in V(G)\}$ indexed by the vertex set of G such that $x_v \in S_v$ for all $v \in V(G)$ and $x_u \neq x_v$ whenever $\{u, v\}$ is not an edge of G . For example, if G is the empty graph then X is a system of distinct representatives. In fact, as pointed out by a referee, the notion of G -system of representatives corresponds to list coloring of the complement of G .

Gravier and Maffray [11] conjectured that every claw-free graph has chromatic number equal to its list chromatic number. Ohba [12] conjectured that a graph with chromatic number χ and at most $2\chi + 1$ vertices has list chromatic number χ as well. Since the complement of a k -vertex path is both claw-free, and has at most $2\chi + 1$ vertices, the following result would follow from either of these two conjectures. Since both these conjectures are open (although the second has been proved asymptotically by Reed and Sudakov), we cannot apply earlier results to the lemma below.

Lemma 4.1 *Let P be a path on k vertices, and let $\mathcal{S} = \{S_v : v \in V(P)\}$ be a set system of sets of size at least $k/2$. Then there is a P -system of representatives for \mathcal{S} .*

Proof. If k is odd, then the sets S_i each have size at least $(k+1)/2$. Therefore if we can prove the statement of the lemma for all even values of k , we also obtain all the odd values of k . Throughout the rest of the proof, k is even. In fact, we will prove something slightly more general: we will prove that for any graph F that is the union of any number of vertex-disjoint paths of odd length, there is an F -system of representatives for any family $\mathcal{S} = \{S_v : v \in V(F)\}$, where each S_v has size at least $k/2$. The case $k = 2$ is trivial, so we suppose $k \geq 4$ and proceed by induction on k . Now by Hall's Theorem, we can find a system of distinct representatives for \mathcal{S} if and only if Hall's condition holds: for all $X \subset V(F)$

$$\left| \bigcup_{v \in X} S_v \right| \geq |X|.$$

A system of distinct representatives for F is also an F -system of representatives, so we are done if Hall's condition holds. Therefore we assume that there is an $X \subset V(F)$ which violates Hall's condition. This set must have size greater than $k/2$, since each set in \mathcal{S} has size at least $k/2$. Furthermore, the sets in $\{S_x : x \in X\}$ are pairwise intersecting, otherwise $\bigcup_{x \in X} S_x$ has size at least k , and Hall's condition would hold for X . Since $|X| > k/2$, we can choose $\{u, v\} \subset X$ such that $\{u, v\}$ is an edge of F and the graph F' obtained by removing the vertices u and v from F consists of a union of vertex-disjoint paths of odd length. Now choose any $x \in S_u \cap S_v$, and replace \mathcal{S} by the family

$$\mathcal{S}' = \{S_v - \{x\} : v \in V(F')\}.$$

Since $|V(F')| = k - 2$, \mathcal{S}' has an F' -system of representatives, by induction. Adding x to this set gives an F -system of representatives for \mathcal{S} , in which S_u and S_v are represented by x . ▀

Proof of Theorem 1.3 for $r = 3$ and Theorem 1.4. Let X, Y and Z be disjoint sets of size n . We start with the case $r > 4$:

Case 1 : $r > 4$. By Proposition 2.5, to prove Theorem 1.4 for $r > 4$, it suffices to prove

$$z(X \times Y \times Z, \mathcal{C}_k) \leq 3\ell n^2.$$

We will show that if $\mathcal{A} \subset X \times Y \times Z$ and $|\mathcal{A}| > 3\ell n^2$, then \mathcal{A} contains a minimal k -cycle. Let \mathcal{A}' denote the family of all sets $\{x, y, z\} \in \mathcal{A}$ with $y \in Y$ and $z \in Z$, such that $\{y, z\}$ is a subset of at least $\lceil \frac{1}{2}k \rceil$ sets in \mathcal{A} . Then

$$|\mathcal{A}| \leq (\lceil \frac{1}{2}k \rceil - 1)n^2 + |\mathcal{A}'| \leq \ell n^2 + |\mathcal{A}'|.$$

Since $|\mathcal{A}| > 3\ell n^2$, we have $|\mathcal{A}'| > 2\ell n^2$. So, for some $x_1 \in X$, we must have

$$|\mathcal{A}'_{x_1}| > 2\ell n.$$

By Theorem 2.1, with $r = 2$, \mathcal{A}'_{x_1} contains a k -path $\mathcal{P} = \{A_1, A_2, \dots, A_k\}$. The next step is to apply Lemma 4.1. Take P to be the path with consecutive vertices $2, 3, 4, \dots, k-1$, and let \mathcal{S} be the family of sets defined by

$$S_i = \{y \in X \setminus \{x_1\} : A_i \in \mathcal{A}'_y\}$$

for $2 \leq i \leq k-1$. By definition of \mathcal{A}' , we have $|S_i| \geq (k-2)/2$ for all i . By Lemma 4.1, there exists a P -system of representatives for \mathcal{S} . This means we can find vertices $x_2, \dots, x_{k-1} \in X$ representing S_2, S_3, \dots, S_{k-1} such that x_i and x_j are distinct whenever $|i-j| > 1$. Now $A_i \in \mathcal{A}_{x_i}$ and $A_1, A_k \in \mathcal{A}_{x_1}$, so the sets $A_1 \cup \{x_1\}, A_2 \cup \{x_2\}, \dots, A_k \cup \{x_1\}$ form a minimal k -cycle in \mathcal{A} . This completes the proof of Theorem 1.4 for $r > 4$.

Case 2 : $r = 3$ and $r = 4$. We now derive the upper bounds in Theorem 1.4 for $r = 3$ and $r = 4$ as well as the second bound in Theorem 1.3. The main claim is that if Y is a set of size $|X| - 1$, then

$$\text{ex}(X^{(r)}, \mathcal{C}_k) \leq \frac{1}{r} \left(\binom{n}{r-1} \left(rk - \left\lfloor \frac{3k}{2} \right\rfloor \right) + n \cdot \text{ex}(Y^{(r-1)}, \mathcal{P}_k) \right).$$

To prove this claim, let $\mathcal{A} \subset X^{(r)}$ be a set system containing no minimal k -cycle, and let

$$\mathcal{L} = \{E \in X^{(r-1)} : d(E) \geq (r-2)k + \lceil \frac{k}{2} \rceil + 1\}.$$

Set $\mathcal{S} = X^{(r-1)} \setminus \mathcal{L}$. Then

$$r|\mathcal{A}| = \binom{r}{r-1} |\mathcal{A}| = \sum_{E \in X^{(r-1)}} d(E) = \sum_{E \in \mathcal{S}} d(E) + \sum_{E \in \mathcal{L}} d(E).$$

We bound each term on the right separately. By definition of \mathcal{S} ,

$$\sum_{E \in \mathcal{S}} d(E) \leq \binom{n}{r-1} \left((r-2)k + \left\lfloor \frac{k}{2} \right\rfloor \right) = \binom{n}{r-1} \left(rk - \left\lfloor \frac{3k}{2} \right\rfloor \right).$$

We next prove that $\sum_{E \in \mathcal{L}} d(E) \leq n \cdot \text{ex}(Y^{(r-1)}, \mathcal{P}_k)$. Suppose, for a contradiction, that this is false. Then there is a vertex $x_1 \in X$ for which $|\text{res}(x_1) \cap \mathcal{L}| > \text{ex}(Y^{(r-1)}, \mathcal{P}_k)$ where $Y = X \setminus \{x_1\}$. Consequently, there exist sets $A_1, \dots, A_k \in \text{res}(x_1) \cap \mathcal{L}$ forming a k -path and satisfying $A_i \cup \{x_1\} \in \mathcal{A}$. Let

$$X_i = \{y \in X - \{x_1\} - \bigcup_{j \neq i} A_j : A_i \in \text{res}(y)\}.$$

Now $|\bigcup_{j \neq i} A_j| \leq (r-2)(k-1) + 1$. So, by definition of \mathcal{L} ,

$$|X_i| \geq d(A_i) - \left| \bigcup_{j \neq i} A_j \right| \geq \lceil \frac{k}{2} \rceil + r - 3 \geq \lceil \frac{k}{2} \rceil$$

for each $i \in \{1, 2, \dots, k\}$. We now apply Lemma 4.1 and proceed as in the proof of Case 1 to find a minimal k -cycle of the form $\{A_i \cup \{y_i\} : 1 \leq i \leq k\}$ where $y_1 = y_k = x_1$ and $y_i \in X_i$ for

$1 < i < k$. This proves the claim. We remark that a minimal k -path is obtained too, of the form $\{A_i \cup y_i : 1 \leq i \leq k\}$ where $y_i \in X_i$.

Now the bounds for $r = 3$ and $r = 4$ follow from the claim. To see this, we first let $r = 3$ in the claim. Then

$$\text{ex}(X^{(3)}, \mathcal{C}_k) \leq \frac{1}{3} \left[\binom{n}{2} \left\lceil \frac{3k}{2} \right\rceil + n \cdot \text{ex}(Y^{(2)}, \mathcal{P}_k) \right].$$

By Theorem 1.1, $\text{ex}(Y^{(2)}, \mathcal{P}_k) \leq \frac{k-1}{2}(n-1)$, which gives

$$\text{ex}(X^{(3)}, \mathcal{C}_k) \leq \frac{1}{3} \left(\left\lceil \frac{3k}{2} \right\rceil + k - 1 \right) \binom{n}{2} \leq \frac{5k-1}{6} \binom{n}{2}.$$

It may be checked, as we remarked above, that the same bound holds for paths:

$$\text{ex}(X^{(3)}, \mathcal{P}_k) \leq \frac{5k-1}{6} \binom{n}{2},$$

which was the second upper bound in Theorem 1.3. Using this fact and the claim for $r = 4$, we obtain

$$\text{ex}(X^{(4)}, \mathcal{C}_k) \leq \frac{1}{4} \left\lceil \frac{5k}{2} \right\rceil \binom{n}{3} + \frac{3}{4} \left(\frac{5k-1}{6} \right) \binom{n}{3} \leq \frac{5k}{4} \binom{n}{3}$$

This completes the proof of Theorem 1.4. ■

5 Minimal 4-Cycles

In this section we prove Theorem 1.6. This is achieved via the following result in the case $r = 3$, and a suitable application of Lemma 2.2 and Corollary 2.4 (see Proposition 2.5 for the details). The proof involves a substantially more delicate analysis than the case of k -cycles.

Theorem 5.1 *Let $r \geq 3$, let $n \geq 4$, and let X_1, X_2, \dots, X_r be disjoint sets of size n . Then $z(\prod^{(r)}, \mathcal{C}_4) \leq n^{r-1} + 11n^{r-2}$.*

To prove Theorem 5.1 in the case $r = 3$, let $|X| = |Y| = |Z| = n$ for some disjoint sets X, Y and Z , and let $\mathcal{A} \subset X \times Y \times Z$ contain no minimal 4-cycle; we must show that $|\mathcal{A}| \leq n^2 + 11n$.

A *star* is a graph consisting of a positive number of edges incident with a fixed vertex. We define a *double star* to be a graph comprising a pair of disjoint stars whose centers are joined by a new edge. Note that a double star contains a path of length three, and therefore at least one leaf in X and in Y . Any vertex of degree one in a star or double star is called a *leaf*. Finally, the edge of a double star joining the centers of two stars is called the *central edge*.

For $z \in Z$, let \mathcal{S}_z denote the set of edges in $\text{res}\{z\}$ which are contained in precisely one set in \mathcal{A} . Let \mathcal{L}_z denote the set of edges in $\text{res}\{z\}$ not in \mathcal{S}_z . Finally, let $\mathcal{S} = \bigcup_{z \in Z} \mathcal{S}_z$ – this is the set of edges with one end in X and the other in Y , which appear in exactly one triple in \mathcal{A} .

Claim 1. For $z \in Z$, each component of \mathcal{L}_z is contained in a double star.

Proof. It suffices to show that \mathcal{L}_z contains no 4-path and no 4-cycle. Suppose, for a contradiction, that \mathcal{L}_z contains edges A_1, A_2, A_3, A_4 forming a 4-path or 4-cycle. Then, by definition of \mathcal{L}_z , there exist (not necessarily distinct) vertices $v, w \in Z$, distinct from z , such that $A_2 \in \mathcal{L}_v$ and $A_3 \in \mathcal{L}_w$. Consequently $\{A_1 \cup \{z\}, A_2 \cup \{v\}, A_3 \cup \{w\}, A_4 \cup \{z\}\}$ is a minimal 4-cycle in \mathcal{A} , a contradiction. Therefore \mathcal{L}_z contains no 4-path or 4-cycle. The proof of Claim 1 is complete.

Claim 2. For $w \in Z$, let $\mathcal{S}_{\neg w}$ denote the graph of all edges of $\mathcal{S} \setminus \mathcal{S}_w$ which intersect at least one edge of \mathcal{L}_w . Then $|\mathcal{S}_{\neg w}| \leq 2n$.

Proof. A vertex of \mathcal{L}_w is an element in some edge of \mathcal{L}_w . We assert that for each $x \in X$, there is at most one edge of $\mathcal{S}_{\neg w}$ joining x to a vertex of \mathcal{L}_w . The same assertion is made for each $y \in Y$; the proof of the latter statement will be the same as that for each $x \in X$. Suppose, for a contradiction, $\{w, a, x\}, \{w, b, x\} \in \mathcal{A}$, with $\{a, v\}, \{b, x'\} \in \mathcal{L}_w$, for some $v, x, x' \in X$. To prove the assertion, we must show that $a = b$. First note that $\mathcal{S}_{\neg w} \cap \mathcal{L}_w = \emptyset$, and therefore $\{a, x\}, \{b, x\} \notin \mathcal{L}_w$. Let us suppose that $\{a, x\} \in \mathcal{S}_t$ and $\{b, x\} \in \mathcal{S}_u$, where t, u are distinct from w . Then $\{(x, a, t), (x, b, u), (v, a, w), (x', b, w)\}$ is a minimal 4-cycle in \mathcal{A} . This contradiction shows that $a = b$, as required.

By definition, each edge of $\mathcal{S}_{\neg w}$ is incident with an element of \mathcal{L}_w . Combining the above assertions, we find that:

$$|\mathcal{S}_{\neg w}| \leq \sum_{x \in X} 1 + \sum_{y \in Y} 1 = |X| + |Y| = 2n.$$

This completes the proof of Claim 2.

Claim 3. For $z \in Z$, let X_z, Y_z denote the sets of leaves of \mathcal{L}_z in X, Y respectively. Then

$$n|\mathcal{A}| \leq n^3 + 3n^2 + 2 \sum_{w \in Z} |X_w||Y_w|.$$

Proof. Let X_z^* and Y_z^* denote the sets of vertices of X and Y incident with central edges of \mathcal{L}_z , respectively. By Claim 1, the central edges form a matching, so

$$|\mathcal{L}_z| \leq |X_z| + |Y_z| + |X_z^*|.$$

So, for each $w \in Z$, we have

$$\begin{aligned} |\mathcal{A}| &= \sum_{z \in Z} |\mathcal{L}_z| + \sum_{z \in Z} |\mathcal{S}_z| \\ &\leq \sum_{z \in Z} (|X_z| + |Y_z| + |X_z^*|) + |\mathcal{S}_w| + \sum_{\substack{z \in Z \\ z \neq w}} |\mathcal{S}_z| \end{aligned}$$

$$\leq \sum_{z \in Z} (|X_z| + |Y_z| + |X_z^*|) + |\mathcal{S}_w| + |\mathcal{S}_{\neg w}| + \sum_{\substack{z \in Z \\ z \neq w}} |\mathcal{S}_z \setminus \mathcal{S}_{\neg w}|.$$

Now the last sum is at most $(n - |X_w| - |X_w^*|)(n - |Y_w| - |Y_w^*|)$, as every edge of $\mathcal{S}_z \setminus \mathcal{S}_{\neg w}$ is disjoint from all edges in \mathcal{L}_w , by definition of $\mathcal{S}_{\neg w}$. By Claim 2, $|\mathcal{S}_{\neg w}| \leq 2n$. Therefore

$$\begin{aligned} |\mathcal{A}| &\leq \sum_{z \in Z} (|X_z| + |Y_z| + |X_z^*|) + |\mathcal{S}_w| + 2n + (n - |X_w| - |X_w^*|)(n - |Y_w| - |Y_w^*|) \\ &\leq \sum_{z \in Z} (|X_z| + |Y_z| + |X_z^*|) + |\mathcal{S}_w| + 2n + (n - |X_w| - |X_w^*|)(n - |Y_w|) \\ &\leq \sum_{z \in Z} (|X_z| + |Y_z| + |X_z^*|) + |\mathcal{S}_w| + 2n + n^2 - n(|X_w| + |Y_w| + |X_w^*|) + 2|X_w||Y_w|. \end{aligned}$$

In the last line, we used the fact $|X_w^*| \leq |X_w|$. This follows from the fact that each double star has at least one leaf in X . Summing over $w \in Z$ gives

$$n|\mathcal{A}| \leq n^3 + 3n^2 + 2 \sum_{w \in Z} |X_w||Y_w|.$$

Here we used the fact $\sum_{w \in Z} |\mathcal{S}_w| \leq n^2$. This completes Claim 3.

The proof of Theorem 1.6 is complete, once we have verified the following claim:

Claim 4. $\sum_{w \in Z} |X_w||Y_w| \leq 4n^2$.

Proof. The terms in the sum above can be interpreted as the number of pairs $\{u, v\}$ with $u \in X_w$ and $v \in Y_w$. In words, this means that both u and v are leaves of \mathcal{L}_w . Let \mathcal{M}_w denote the set of edges of \mathcal{L}_w which are not incident with any other edges of \mathcal{L}_w . Then \mathcal{M}_w is a matching. Let $X_w^{**} \subset X_w$ and $Y_w^{**} \subset Y_w$ be the sets of vertices of \mathcal{M}_w in X and Y . For each $w \in Z$, define a bipartite graph H_w with parts X_w and Y_w and in which $\{u, v\} \in H_w$ whenever $u \in X_w, v \in Y_w$ and $(u, v) \notin X_w^{**} \times Y_w^{**}$, and let H be the bipartite multigraph with parts X and Y consisting of the sum of all the graphs H_w . In other words, an edge $e \in H$ has multiplicity equal to the number of w for which $e \in H_w$. We claim that H has no multiple edges.

Suppose, for a contradiction, that $\{u, v\} \in H$ has edge-multiplicity at least two. Then, for some $w_1, w_2 \in Z$, u and v are leaves of stars in \mathcal{L}_{w_i} for $i \in \{1, 2\}$. Since $(u, v) \notin X_{w_i}^{**} \times Y_{w_i}^{**}$, $\{u, v\} \notin \mathcal{L}_{w_i}$, for $i \in \{1, 2\}$. Therefore we find vertices $u_i \in X$ and $v_i \in Y$ distinct from u and v , and such that $\{u, u_i\}, \{v, v_i\} \in \mathcal{L}_{w_i}$. Note that we may have $u_1 = u_2$ or $v_1 = v_2$. In any case,

$$\{(u, u_1, w_1), (u, u_2, w_2), (v_1, v, w_1), (v_2, v, w_2)\}$$

is a minimal 4-cycle in \mathcal{A} , which is a contradiction. Therefore H has no multiple edges. This implies that

$$|H| = \sum_{w \in Z} (|X_w||Y_w| - |X_w^{**}||Y_w^{**}|) \leq n^2.$$

Finally, to complete the proof of Claim 4, we show that $\sum_w |X_w^{**}||Y_w^{**}| \leq 3n^2$. Now as $|Y_w^{**}| = |X_w^{**}|$, we have

$$\sum_{w \in Z} |X_w^{**}||Y_w^{**}| = \sum_{w \in Z} |X_w^{**}|^2 = 2 \sum_{w \in Z} \binom{|X_w^{**}|}{2} + \sum_{w \in Z} |X_w^{**}|.$$

The last term is at most n^2 , so it suffices to show that the first term is less than $2n^2$. To do this, we will show that $\{u, v\} \subset X_w^{**}$ for at most two $w \in Z$. Suppose, for a contradiction, that this is not the case. Then there are pairs of disjoint edges $\{u, u_i\}, \{v, v_i\} \in \mathcal{M}_{w_i}$ for $i \in \{1, 2, 3\}$, and for some distinct vertices $w_1, w_2, w_3 \in Z$. If $\{u_i, u_j\} \cap \{v_i, v_j\} = \emptyset$ for some distinct $i, j \in \{1, 2, 3\}$, then

$$\{(w_i, u, u_i), (w_j, u, u_j), (w_j, v, v_j), (w_i, v, v_i)\}$$

is a minimal 4-cycle in \mathcal{A} , which is a contradiction. So $\{u_i, u_j\} \cap \{v_i, v_j\} \neq \emptyset$ for all $i, j \in \{1, 2, 3\}$. It follows that the graph K spanned by all the edges $\{u, u_i\}$ and $\{v, v_i\}$ is a complete bipartite graph. We may assume $v_1 = u_2, v_2 = u_3, v_3 = u_1$ and $u_1 \neq u_3$. Then

$$\{(w_1, u, u_1), (w_3, u, u_3), (w_2, v, u_3), (w_1, v, u_1)\}$$

is a minimal 4-cycle in \mathcal{A} . This contradiction completes the proof. ■

6 Acknowledgments

The authors thank the referee for helpful comments which included shortening the proof of Theorem 1.3 Case 1.

References

- [1] Berge, C. Hypergraphs. Combinatorics of finite sets. Translated from the French. North-Holland Mathematical Library, 45. North-Holland Publishing Co., Amsterdam, 1989.
- [2] Bollobás, B. Combinatorics. Set systems, hypergraphs, families of vectors and combinatorial probability. Cambridge University
- [3] A. J. Bondy, Large cycles in graphs, *Discrete Math* **1** 1971/1972 no. 2, 121–132.
- [4] Erdős, P., Gallai, T. On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hungar* **10** (1959) 337–356.
- [5] Erdős, P., Kleitman, D. On coloring graphs to maximize the proportion of multicolored k -edges. *J. Combinatorial Theory* **5** (1968) 164–169.
- [6] Erdős, P., Ko, H., Rado R. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)* **12** (1961) 313–320.
- [7] Fischer, P., Matoušek, J. A lower bound for families of Natarajan dimension d . *J. Combin. Theory Ser. A* 95 (2001), no. 1, 189–195.
- [8] Frankl, P. Extremal set systems. Handbook of combinatorics, Vol. 1, 2, 1293–1329, Elsevier, Amsterdam, 1995

- [9] Füredi, Z. Hypergraphs in which all disjoint pairs have distinct unions. *Combinatorica* **4** (1984), no. 2-3, 161–168.
- [10] Füredi, Z. Turán type problems. Surveys in combinatorics, 1991 (Guildford, 1991), 253–300, *London Math. Soc. Lecture Note Ser.*, **166**, Cambridge Univ. Press, Cambridge, 1991.
- [11] Gravier, S., Maffray, F. On the choice number of claw-free perfect graphs, 6th International Conference on Graph Theory, *Discrete Math.* 276 (2004), no. 1-3, 211–218.
- [12] Ohba, K. On chromatic-choosable graphs, *J. Graph Theory* 40 (2002), no. 2, 130–135.
- [13] Lazebnik, F., Verstraëte, J. On Hypergraphs of Girth Five, *Electronic Journal of Combinatorics*, **10** (2003) #R25.
- [14] Mubayi, D., Verstraëte, J. A hypergraph extension of the bipartite Turán problem. *J. Combin. Theory Ser. A* 106 (2004), no. 2, 237–253.
- [15] Mubayi, D., Verstraëte, J. Proof of a conjecture of Erdős on triangles in set systems, *Combinatorica*, 25 (2005), no. 5, 599–614..
- [16] Simonovits, M. A method for solving extremal problems in graph theory, stability problems. 1968 Theory of Graphs (Proc. Colloq., Tihany, 1966) pp. 279–319 Academic Press, New York
- [17] D. R. Woodall, Sufficient conditions for circuits in graphs, *Proc. London Math. Soc (3)* **24** (1972), 739–755.