Uniform Set Systems and their Diameters

David White

442615

Submitted to the School of Mathematics, University of the Witwatersrand

In fulfilment of the requirements for the degree of

Master of Science

Supervised by Professors Sheng Bau and Anne Love

31 May 2016

Declaration

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work. It is being submitted in fulfillment of the degree of Master of Science at the University of the Witwatersrand, Johannesburg. It has not previously been submitted to another University for any degree or examination, in its entirety or in part.

....

David Mark White

Signed on this the 31st day of May, 2016.

Abstract

This dissertation examines the existing literature on set systems (or hypergraphs) and conducts an investigation of their $(k-1)$-overlapping diameters. In the general case of set systems of given diameter, some bounds on the possible sizes are given. We then restrict our focus to acyclic set systems and provide a full classification for simple, connected, acyclic, uniform set systems of each positive diameter, extending some results on trees.

Contents

Part I. Preliminaries

1 Definitions and Concepts

We begin by defining set systems and a few basic properties. Let V be a set and E a collection of subsets of V. Then, if $\bigcup_{e \in E} = V$, the pair $G = (V, E)$ is said to be a *set system* (or hypergraph) on V. The elements of V are called *vertices* of G while the elements of E are called *edges*. A set system $H = (V, E')$ is said to be a set subsystem of $G = (V, E)$ if $E' \subset E$. As with graphs, the *order* of G is the cardinality of V, while the *size* of G is the cardinality of E.

A set system $G = (V, E)$ is said to be *simple* (or Spernerian), if for each $e, f \in E$, $e \not\subseteq f$. That is, a set system is simple if no edge is contained in another edge. If for each $e \in E$, $|e| = k$, then G is said to be k-uniform. Note that graphs are just 2-uniform set systems. Throughout this dissertation all set systems will be assumed to be simple and uniform.

For the purpose of this dissertation, we will use the definition from [12] of a vertex $v \in V$ being *isolated* if it contained in exactly one edge $e \in E$. Similar definitions and some further discussions around properties of set systems can be found in [3, 4, 12].

We now define the geodesics we will be using and some properties based on them, including our primary focus: diameter.

Let $G = (V, E)$ be a set system. In [11], Tomasta defines a *path* of *length* q connecting vertices u and v to be a sequence of distinct edges $e_1, e_2, ..., e_q$, such that $u \in e_1, v \in e_q$, and for each $i = 1, 2, ..., q-1$, $e_i \cap e_{i+1} \neq \emptyset$. A path in which the first and last edges also have non-empty intersection is called a *cycle*.

Research has recently been done (especially into Hamiltonian cycles in set systems, see [7, 8]) where paths (and cycles) were specified to be *ℓ-overlapping paths* (respectively *ℓ-overlapping cycles*) if each pair of consecutive edges in the path intersects in precisely ℓ vertices, that is if $|e_i \cap e_{i+1}| = \ell$ for all $i = 1, .., q-1$ (and for ℓ-overlapping cycles, the additional constraint $|e_1 \cap e_q| = \ell$). Note that in k-uniform set systems, successive vertices in the tracing of $(k-1)$-overlapping paths are uniquely defined by the list of edges. As with graphs, we define the following concepts in terms of paths:

- G is said to be *ℓ-overlap connected* if for each $u, v \in V$ there exists an ℓ-overlapping path connecting u and v.
- A shortest ℓ-overlapping path connecting u and v is said to be an *ℓ-overlapping geodesic*.

- The length of such a path is defined to be the (geodesic) *ℓ-overlapping distance* between u and v, denoted $d_\ell(u, v)$.
- The *ℓ-overlapping eccentricity* of a vertex v, $\epsilon_\ell(v)$, is the greatest ℓ-overlapping distance between v and any other vertex in the set system.
- The *ℓ-overlapping radius* of G, $r_\ell(G)$, is the minimum ℓ-overlapping eccentricity of any vertex.
- The *ℓ-overlapping diameter* of G, $d_\ell(G)$, is the maximum ℓ-overlapping eccentricity of any vertex, or equivalently the maximum ℓ-overlapping distance between any two vertices in G.

Note that traditional paths and cycles in graphs are simply 1-overlapping paths in 2-uniform set systems and these concepts all correspond with their graph theoretic definitions.

Now, we define the concept of acyclicity of set systems, as thoroughly investigated by Wang in [12]. To do this we first define the two types of *Graham reductions* that can be applied to a set system $G = (V, E)$:

(i) If there exists $e_1, e_2 \in E$, with $e_1 \subseteq e_2$, remove e_1. That is, $E \leftarrow E \backslash \{e_1\}$, $V \leftarrow V$.

(ii) If $v \in V$ is an isolated vertex of G, contained in $e \in E$, delete it. That is, $E \leftarrow E \backslash \{e\} \cup \{e \backslash \{v\}\}$, $V \leftarrow V \backslash \{v\}$.

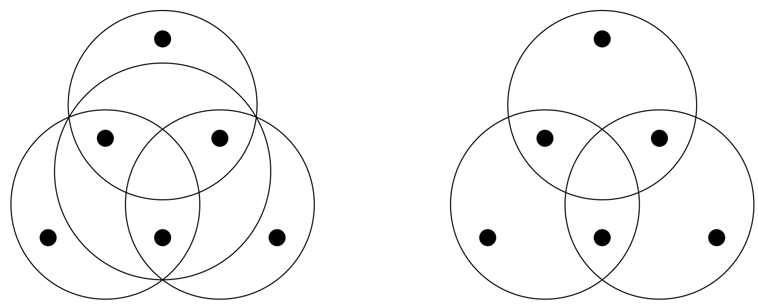


Fig. 1: $G = (V, E)$ and $G' = (V, E')$

A set system G is said to be *acyclic* if there exists a sequence of Graham reductions on G resulting in the empty set. A set system which is not acyclic is said to be *cyclic*.

An edge $e \in E$ is said to be an *ear* of G if there exists another edge $f \in E$ such that all the vertices of $e \backslash f$ are isolated, and a set system can equivalently be defined to be acyclic if there exists a sequence of ear removals ($E \leftarrow E \backslash \{e\}$, $V \leftarrow V \backslash (e \backslash f)$) resulting in the empty set. (Observe that for ear reductions, the edge f may be the empty set, so that the last edge is removable).

Note that unlike traditional graphs, acyclic set systems may contain cyclic set subsystems. Fig. 1 shows an example of this, provided by Wang in [12].

In this example,

$$V = \{1, 2, 3, 4, 5, 6\},$$

$$E = \{\{1, 2, 3\}, \{3, 4, 5\}, \{5, 6, 1\}, \{1, 3, 5\}\}, \text{ and}$$

$$E' = \{\{1, 2, 3\}, \{3, 4, 5\}, \{5, 6, 1\}\}.$$

Clearly, $G' = (V, E')$ is a set subsystem of $G = (V, E)$. Note that G is acyclic as its edges can be removed as ears in the order they are written. G', however, contains no ears and is therefore cyclic.

We now define some other objects that can be defined from set systems and some properties that they can give us, which will be consolidated in Theorem 2.1. First, for a general set system $G = (V, E)$, we may define a k-uniform set system $G_{(k)} = (V, E_{(k)})$ by setting $E_{(k)} = \{S \subset V : |S| = k, S \subset e \text{ for some } e \in E\}$. We will consider the case of $k = 2$, where $G_{(2)}$ is a graph such that two vertices are adjacent if they appear together in an edge of G.

A set system $G = (V, E)$ is said to be *chordal* if $G_{(2)}$ is chordal in the usual graph theoretic sense (namely that any cycle of length at least 4 must have a chord. A chord of a cycle is an edge connecting two non-successive vertices in the cycle).

We may also use $G_{(2)}$ to define the notion of conformality: G is said to be conformal if for every clique (complete subgraph) U of $G_{(2)}$ there exists an edge $e \in E$ such that $U \subset e$.

Now, we define another object that can be generated from G, namely the *intersection closure semilattice*, G^*, to be the smallest set containing $E(G)$ and closed under set intersection ([12]).

We then define the *Hassen digraph* of G^* to be the digraph with G^* as its vertex set and $x, y \in G^*$ adjacent if and only if $x \subsetneq y$ and there is no $z \in G^*$ such that $x \subsetneq z \subsetneq y$. The *Hassen graph* of G^*, denoted $H(G^*)$, is the underlying undirected graph of the Hassen digraph.

Now, for each $x \in G^*$, we denote by $c^+(x)$ the number of components of $H(G^*)$ when restricted to the vertex set $\{y \in G^* : x \subsetneq y\}$. In [9], Lee defines a generalization of the graph theoretic concept of cyclomatic number to set systems by

$$l(G) = 1 + \sum_{x \in G^*} (c^+(x) - 1).$$

For any vertex $x \in G^*$, denote by $G[x]$ the set of edges of G that contain x.

Finally, the *Line graph* of a set system G, denoted $L(G)$, is defined in [12] to be graph with

$$V(L(G)) = E(G) \text{ and } E(L(G)) = \{(e,f) : e,f \in E(G) \text{ and } e \cap f \neq \emptyset\},$$

where (e,f) has weight $|e \cap f|$. In [1], another generalization of the cyclomatic number is given:

$$\mu(G) = \sum_{e \in E(G)} |e| - \left| \bigcup_{e \in E(G)} e \right| - w(G),$$

where $w(G)$ denotes the maximum weight of a forest of $L(G)$.

In [13] and [6], Wang collaborated on papers that proved that both the cyclomatic numbers l and μ take value zero if and only if the set system is acyclic as defined by Graham reductions, and hence they certainly agree for zero values. This forms part of Theorem 2.1.

However, we now provide an example to that show that these two cyclomatic numbers are not necessarily equal for cyclic set systems:

Consider any set V of order n, and set $E = \binom{V}{k}$, the set of all subsets of V of order k, for some positive integer k. We call $G = (V, E)$ the *complete* k-uniform set system of order n.

Let $k \geq 3, n \geq k+1$ and let G be the complete k-uniform set system of order n.

Then $G^* = \{x \subset V : |x| \leq k\}$. Clearly, there are $\binom{n}{k}$ edges of G and $c^+(e) = 0$ for each edge $e \in G$. There are $\binom{n}{k-1}$ nodes x with $|x| = k-1$, and for each of these we have $c^+(x) = k-1$. For all other nodes x, $c^+(x) = 1$. Hence

$$\begin{aligned} l(G) &= 1 - \binom{n}{k} + (k-2)\binom{n}{k-1} \\ &= 1 + \binom{n}{k}\frac{k^2 - k - n - 1}{n - k + 1}. \end{aligned}$$

On the other hand, the maximum forest in the Line Graph of G will be a tree containing exactly $\binom{n}{k} - 1$ edges, each of weight $k-1$ so $w(G) = (k-1)(\binom{n}{k} - 1)$, giving

$$\begin{aligned} \mu(G) &= k\binom{n}{k} - n - (k-1)\left(\binom{n}{k} - 1\right) \\ &= \binom{n}{k} - n + k - 1. \end{aligned}$$

Clearly, these do not coincide in general. As a particular example, for $n = 4$ and $k = 3$ we get $l(G) = 3$ and $\mu(G) = 2$. These can be calculated from the Hassen graph and the Line graph of G depicted in Figures 2 and 3 respectively. Observe, though, that when restricted to graphs, these two cyclomatic numbers do agree and are equal to the traditional cyclomatic number of graphs.

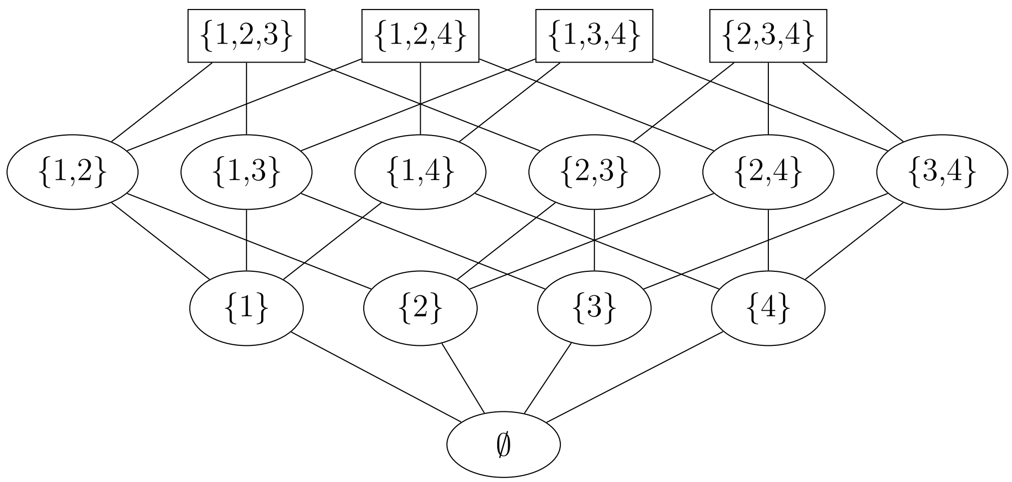


Fig. 2: Hasse graph of the complete 3-uniform set system of order 4

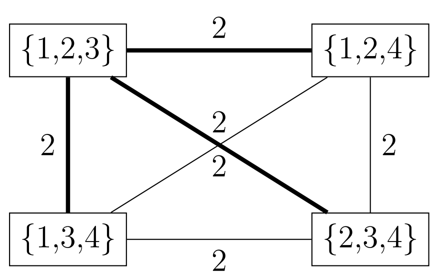


Fig. 3: Line graph of the complete 3-uniform set system of order 4, with a maximum spanning forest in bold

2 Literature Review

Work in [3, 4, 12] provides a thorough introduction to the concepts of set systems and their properties. In addition to wide discussions of the theory of set systems, Berge and Duchet ([3, 4]) conduct investigations around colouring set systems and prove results on chromatic numbers. In section 4.1 of [12], Wang provides a discussion of the acyclicity of set systems and proves four characterizations of acyclic set systems, summarized in the following theorem.

Theorem 2.1 ([12], p57)**.** Let G be a set system. Then the following are equivalent:

1. G is acyclic as defined by Graham reductions,
2. G is chordal and conformal
3. $l(G) = 0$,
4. $\mu(G) = 0$, and
5. For any $x \in G^*$, $G(x)$ is acyclic.

Summary of Proof. In [12], Wang provides proofs for each equivalence. We shall provide a summary of these results.

1. $\Leftrightarrow$ 2.) (proved first in [2]. We cite it from [12], p50) Assume G is acyclic, i.e. there exists a sequence of Graham reductions on G leading to the empty set.

   Let C be a cycle in $G_{(2)}$ of length at least 4, and let v be the first vertex of C to be removed during this sequence. Since v is isolated before removal, and adjacent to both the preceding and succeeding vertices in C, these two vertices must be contained in the single edge containing v and so they must be adjacent in $G_{(2)}$, giving that C has a chord.

   Now, let U be a clique of $G_{(2)}$ and let v be the first vertex of U to be removed during the sequence of Graham reductions. As before, immediately before removal, v is isolated and adjacent to all of U hence U must be wholly contained in an edge of the set system in its present state. Clearly, this this edge containing U is contained in some edge of the original set system.

   Now, assume G is chordal and conformal. A classical result by Fulkerson and Gross in [5] states that since $G_{(2)}$ is a chordal graph, it contains a simplicial vertex u, i.e. a vertex u such that $G_{(2)}|_{N(u)}$, the subgraph obtained by restricting $G_{(2)}$ to the neighbourhood of u, is a clique.

Since G is conformal, we have that there exists an edge $e \in E(G)$ such that $N(u) \cup \{u\} \subset e$. Suppose u is contained in another edge $f \in E(G)$. Then $f \subset N(u) \subset e$ so f can be removed by a Graham reduction (i). Hence we can assume e is the only edge containing u, and so u can be removed by a Graham reduction (ii).

Note that the removal of a simplicial vertex will preserve chordality and conformality and hence this can be repeated until G is empty.

1. $\Leftrightarrow$ 3.) In [13], Wang and Lee proved that the removal of an ear from a set system preserves the parameter l, by use of an algorithm. In [12], Wang then proved that a set system, G, with the property $l(G) = 0$ will contain an ear. Together these two results prove the equivalence.

1. $\Leftrightarrow$ 4.) (proved first in [6]. We cite it from [12], p55) As above, Haizhu and Wang proved that the removal of an ear from a set system also preserves the parameter μ, this time by considering a few cases. They then proved that a set system, G, with the property $\mu(G) = 0$ will contain an ear.

3. $\Rightarrow$ 5.) (proved first in [13]. We cite it from [12], p54) Assume 3. and suppose 5. does not hold, i.e. there exists $x \in G^*$ such that $G(x)$ is cyclic. By 1. $\Leftrightarrow$ 3. we know that $l(G(x)) \geq 1$.

We may choose x to be minimal, i.e. a node such that $G(x)$ is cyclic but $G(y)$ is acyclic for any $y < x$. But then $G(x)$ would be obtainable from G by the removal of ears, and hence $l(G(x)) = l(G) = 0$, a contradiction.

5. $\Rightarrow$ 1. is clear by letting $x = \emptyset$.

□

In the remainder of [12], Wang also introduces the idea of Hamiltonian cycles and proves some results on the decomposition of complete set systems into Hamiltonian cycles.

In [11], Tomasta extends to the case of complete regular set systems some results by Palumbíny ([10]) on the existence and size of decompositions of complete graphs into isomorphic factors with equal diameter.

As stated in Section 1, research done on ℓ-geodesics in set systems has focused on Hamiltonian cycles. In [7], Hàn and Schacht define the minimum $(k-1)$-degree of a set system to be the minimum number of edges containing a fixed set of $k-1$ vertices. They then prove a Dirac-type condition on this for the existence of Hamiltonian ℓ-cycles. In particular, they show that for each $\ell < k/2$, every k-uniform set system of order n and minimum $(k-1)$-degree $n/(2(k-\ell)) + o(n)$ contains an Hamiltonian ℓ-cycle.

In [8], Kühn, Mycroft and Osthus prove a conjecture by Hán and Schacht by showing that for each $\ell < k$ with $k-\ell$ not a factor of k, any k-uniform set system of order n and minimum degree at least $(n/\lceil k/(k-\ell)\rceil(k-\ell))+o(n)$ contains an Hamiltonian ℓ-cycle.

Part II. The Size of Uniform Set Systems of Given Diameter

In this part, we prove a few simple results on the possible sizes that simple, ℓ-overlap connected uniform set systems can take for each possible diameter.

3 Lower Bounds

Let G be an ℓ-overlap connected, k-uniform set system of order $n > k$, where $1 \leq \ell \leq k-1$. Consider any edge and label its vertices $e = \{1, ..., k\}$. There must exist some other edge, f, which intersects with e in exactly ℓ vertices. In turn, there must be an edge which intersects with either e or f in exactly ℓ vertices.

We may continue to trace the edges in this way, connecting at most $k - \ell$ new vertices to the system with each edge. Hence there must be at least $\lceil 1 + \frac{n-k}{k-\ell} \rceil = \lceil \frac{n-\ell}{k-\ell} \rceil$ edges.

We now show that for k-uniform set systems with diameter greater than 1, this bound is attained.

Lemma 3.1. Let n,k,ℓ and d be positive integers satisfying $k \geq 2$, $n \geq k+1$, $1 \leq \ell \leq k-1$ and $2 \leq d \leq \lceil \frac{n-\ell}{k-\ell} \rceil$. Then there exists a k-uniform set system of order n and ℓ-overlapping diameter d with size exactly $\lceil \frac{n-\ell}{k-\ell} \rceil$.

Proof. Consider the set system, G, with vertex set $V(G) = \{1, ..., n\}$ and edge set, $E(G)$ consisting of a single ℓ-overlapping path of length $d-1$ and all other vertices connected to the last edge in the path with ℓ-overlapping edges, as follows:

$$\begin{aligned}
E(G) = &\Big\{ \{a(k-\ell)+1, a(k-\ell)+2, ..., a(k-\ell)+k\} : 0 \leq a \leq d-2 \Big\} \\
&\cup \Big\{ \{(d-1)(k-\ell)+1, (d-1)(k-\ell)+2, ..., (d-1)(k-\ell)+\ell\} \\
&\qquad \cup \{bk-(b-1)\ell+1, bk-(b-1)\ell+2, ..., bk-(b-1)\ell+k-\ell\} \\
&\qquad : d-1 \leq b \leq c \Big\} \\
&\cup \Big\{ \{(d-1)(k-\ell)+1, (d-1)(k-\ell)+2, ..., (d-1)(k-\ell)+\ell\} \\
&\qquad \cup \{(c+1)k-c\ell+1, (c+1)k-c\ell+2, ..., n\} \\
&\qquad \cup \{(d-1)k-(d-2)\ell+1, ..., (d+c+1)(k-\ell)-n\} \Big\}
\end{aligned}$$

where $c = \lfloor \frac{n-k}{k-\ell} \rfloor$, and the last edge is omitted if (and only if) $(c+1)k-c\ell = n$ i.e. if $k-\ell$ divides $n-k$ exactly. Clearly, then, we have exactly $1 + \lceil \frac{n-k}{k-\ell} \rceil = \lceil \frac{n-\ell}{k-\ell} \rceil$ edges, and an ℓ-overlapping diameter d. $\square$

Example 3.2. Figure 4 illustrates the above construction for the case of $n = 13, k = 4, d = 4$ and $\ell = 1, 2, 3$.

It now remains to consider k-uniform set systems of ℓ-diameter 1. Note that in this case ℓ does not matter. To have ℓ-diameter 1 is equivalent to saying $G_{(2)}$ is complete, for any ℓ. We will hence just refer to these set systems as having diameter 1. In fact, in this case we can find a much better lower bound on the size of the set system.

Lemma 3.3. Every k-uniform set system of order n and diameter 1 must have size at least

$$\left\lceil \frac{n \left\lceil \frac{n-1}{k-1} \right\rceil}{k} \right\rceil.$$

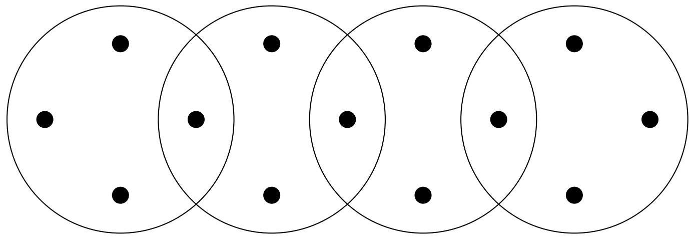
Proof. In order for a k-uniform set system of order n to have diameter 1, every vertex must appear in an edge with every other vertex. But each vertex can only be paired with $k - 1$ other vertices per edge, and so each vertex must appear in at least $\lceil (n-1)/(k-1) \rceil$ edges. Therefore, the edges must contain at least $n\lceil (n-1)/(k-1) \rceil$ vertices in total, proving the result. □

Remark. Lemma 3.3 can equivalently be stated as: We require at least

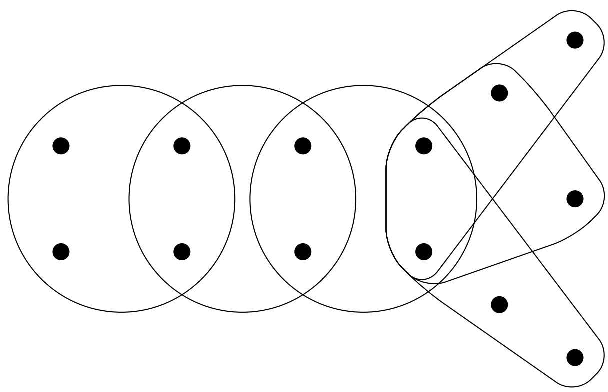
$$\left\lceil \frac{n \left\lceil \frac{n-1}{k-1} \right\rceil}{k} \right\rceil$$

k-cliques to cover the complete graph K_n.

$\ell = 1$:



$\ell = 2$:



$\ell = 3$:

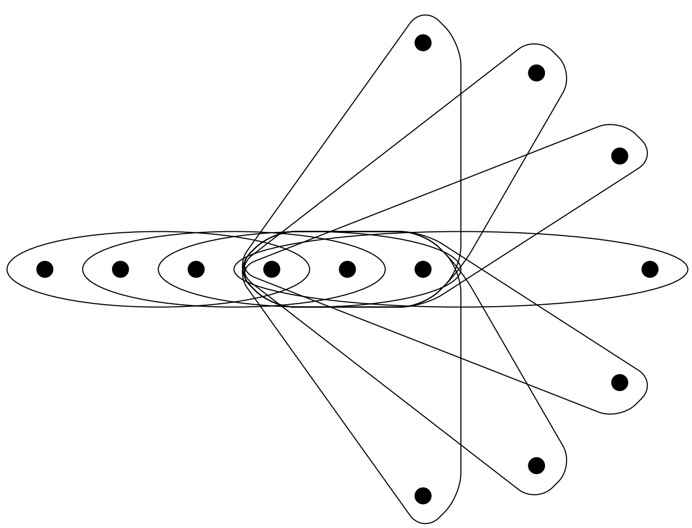


Fig. 4: Minimal 4-uniform set systems of order 13 and ℓ-overlapping diameter 4 for $\ell = 1, 2, 3$

The tightness of Lemma 3.3 remains an open problem. It appears to be tight for the case of $k = 2, 3$:

Example 3.4. In the case of $k = 2$, this bound is simply $\binom{n}{2}$. This indicates that for a graph to have diameter 1, it must be complete - a well-known fact. Clearly, this graph exists and is unique.

In the case of $k = 3$ this bound can be expressed as

$$\left\lceil \frac{n\left\lceil \frac{n-1}{2} \right\rceil}{3} \right\rceil = \begin{cases} \frac{n^2}{6}, & n \equiv 0 \pmod 6 \\ \frac{n(n-1)}{6}, & n \equiv 1 \pmod 6 \text{ or } n \equiv 3 \pmod 6 \\ \frac{n^2+2}{6}, & n \equiv 2 \pmod 6 \text{ or } n \equiv 4 \pmod 6 \\ \frac{n(n-1)+4}{6}, & n \equiv 5 \pmod 6 \end{cases}$$

Table 1 motivates the conjecture that for each $n \geq 3$, there exists a 3-uniform set system of order n, diameter 1, and size exactly

$$\left\lceil \frac{n\left\lceil \frac{n-1}{2} \right\rceil}{3} \right\rceil.$$

n	$\left\lceil \frac{\mathbf{n}\left\lceil\frac{\mathbf{n-1}}{\mathbf{2}}\right\rceil}{\mathbf{3}} \right\rceil$	**Example of an edge set for a set system with vertex set** $\{1, ..., n\}$**, size** $\left\lceil \frac{\mathbf{n}\left\lceil\frac{\mathbf{n-1}}{\mathbf{2}}\right\rceil}{\mathbf{3}} \right\rceil$ **and diameter 1**
3	1	$\Big\{\{1,2,3\}\Big\}$
4	3	$\Big\{\{1,2,3\},\{1,2,4\},\{2,3,4\}\Big\}$
5	4	$\Big\{\{1,2,3\},\{1,4,5\},\{2,3,4\},\{2,4,5\}\Big\}$
6	6	$\Big\{\{1,2,3\},\{1,4,5\},\{1,2,6\},\{2,4,5\},\{3,4,5\},\{3,4,6\}\Big\}$
7	7	$\Big\{\{1,2,3\},\{1,4,5\},\{1,6,7\},\{2,4,6\},\{2,5,7\},\{3,4,7\},\{3,5,6\}\Big\}$
8	11	$\Big\{\{1,2,3\},\{1,4,5\},\{1,6,7\},\{1,2,8\},\{2,4,6\},\{2,5,7\},\{3,4,7\},\{3,5,6\},\{3,5,8\},\{4,5,8\},\{6,7,8\}\Big\}$
9	12	$\Big\{\{1,2,3\},\{1,4,5\},\{1,6,7\},\{1,8,9\},\{2,4,7\},\{2,5,9\},\{2,6,8\},\{3,4,8\},\{3,5,6\},\{3,7,9\},\{4,6,9\},\{5,7,8\}\Big\}$

Tab. 1: Examples of minimal 3-uniform set systems of diameter 1

In general, we know that the bound in Lemma 3.3 is not tight for $k \geq 4$. Consider, for example, the case of $n = 7, k = 4$. A 4-uniform set system of order n and size $\lceil 7 \lceil \frac{6}{3} \rceil / 4 \rceil = 4$ cannot have diameter 1.

Proof. Let $V = \{1, 2, 3, 4, 5, 6, 7\}$. Suppose e_1, e_2, e_3, e_4 are all 4-subsets of V such that $G = (V, \{e_1, e_2, e_3, e_4\})$ is a set system of diameter 1. Without loss of generality, assume $e_1 = \{1, 2, 3, 4\}$. Since each vertex must be adjacent to five other vertices, it must be contained in at least two edges. Since in total the edges contain 16 vertices, there must be at least five vertices that appear in exactly two edges. Without loss of generality assume vertex 1 is such a vertex, with $1 \in e_1 \cap e_2$. Then since vertex 1 must be adjacent to all other vertices, we must have $e_2 = \{1, 5, 6, 7\}$. Assume without loss of generality that vertex 2 is also only contained in two edges, with $2 \in e_1 \cap e_3$. Then similarly we must have $e_3 = \{2, 5, 6, 7\}$. Now, we know vertices 3 and 4 must be contained in at least two edges, so we must have $3, 4 \in e_1 \cap e_4$. But we also know that vertex 3 must be adjacent to all other vertices, so we would need $5, 6, 7 \in e_4$, a contradiction. □

4 Upper Bounds

Clearly, the complete k-uniform set system of order n has diameter 1 and so we cannot find a better upper bound for the size of k-uniform set systems of diameter 1 than the trivial one $\binom{n}{k}$.

It remains to determine tight upper bounds on the sizes of k-uniform set systems of ℓ-overlapping diameter greater than 1. Finding a general tight upper remains an open problem but we provide here a few special cases.

For the case of $k = 2$, it is clear that a graph of diameter d can have at most $\binom{n}{2} - \binom{d}{2}$ edges since it must contain a path of length d containing $d+1$ vertices that are otherwise non-adjacent, i.e. there are $\binom{d+1}{2} - d = \binom{d}{2}$ edges that cannot be in the graph.

We now consider the case of $k = 3$. For a 3-uniform set system of order $n \geq 4$ to have ℓ-overlapping diameter 2, it must contain two vertices x and y which are an ℓ-overlapping distance of 2 apart. This means that the system cannot contain any edge that includes both vertex x and vertex y, since otherwise x and y would only be an ℓ-distance of 1 apart. That is, the set system cannot contain any of the $n-2$ edges of the form $\{x, y, z\}$, i.e. it can have size at most $\binom{n}{3} - (n-2) = \frac{1}{6}(n-2)(n+2)(n-3)$.

We now show that this bound is tight by providing a construction of a 3-uniform set system of order n, size exactly $\binom{n}{3} - (n-2)$ and 1-overlapping and 2-overlapping diameter both equal to 2.

For each $n \geq 5$, consider the 3-uniform set system, G, with vertex set $V(G) = \{1, ..., n\}$ and edge set

$$E(G) = \Big\{\{x, y, z\} : x, y, z \in \mathbb{Z}, 1 \leq x < y < z \leq n\Big\} \setminus \Big\{\{1, 2, a\} : a = 3, ..., n\Big\}.$$

Clearly this system has order n, 2-overlapping diameter 2 (vertices 1 and 2 are a 1-overlapping distance and a 2-overlapping distance apart) and size exactly $\binom{n}{3} - (n-2)$.

Similarly, for a 3-uniform set system of order $n \geq 5$ to have ℓ-diameter 3, it must contain two vertices x and y which are an ℓ-distance of 3 apart. For this example, we will just consider the case of $\ell = 2$.

Certainly, we also cannot have any of the $n-2$ edges containing x and y since then they would be a distance of 1 apart. But we also cannot have x and y being 2 apart, so for each pair of distinct vertices $z_1, z_2 \in V$ we cannot have both $\{x, z_1, z_2\} \in E$ and $\{y, z_1, z_2\} \in E$, so for each of these $\binom{n-2}{2}$ pairs of vertices there is at least one edge that cannot be in the system.

Hence the set system can have size at most

$$\binom{n}{3} - (n-2) - \binom{n-2}{2} = \frac{1}{6}(n-1)(n-2)(n-3).$$

We provide an example to show that this bound is tight. For each $n \geq 5$, consider the 3-uniform set system, $G = (V, E)$, with vertex set $V = \{1, ..., n\}$ and edge set

$$\begin{aligned} E = & \big\{\{x, y, z\} : x, y, z \in \mathbb{Z}, 1 \leq x < y < z \leq n\big\} \\ & \cup \big\{\{n-3, n-2, n\}\big\} \\ & \setminus \Big(\big\{\{a, n-1, n\} : a \in \mathbb{Z}, 1 \leq a \leq n-2\big\} \\ & \cup \quad \big\{\{a, b, n\} : a, b \in \mathbb{Z}, 1 \leq a < b \leq n-2\big\} \\ & \cup \quad \big\{\{n-3, n-2, n-1\}\big\}\Big). \end{aligned}$$

Clearly this system has order n, 2-overlapping diameter 3 (vertices $n-1$ and n are a 2-overlapping distance of 3 apart) and size exactly $\binom{n}{3} - (n-2) - \binom{n-2}{2}$.

Part III. Diameters of Acyclic, Uniform Set Systems

5 Cycles in Acyclic Set Systems

We now discuss the existence of different types of cycles in acyclic set systems. First, we define a *union covered* cycle as one in which every edge is contained in the union of the preceding edge and the succeeding edge in the cycle. That is, a cycle $e_1, e_2, ..., e_q$ is union covered if for each $i = 1..., q$, $e_i \subset e_{i-1} \cup e_{i+1}$ (where indices are taken modulo q). Clearly, in k-uniform set systems, ℓ-overlapping cycles are union covered for all integers $\ell \geq k/2$. Our first result shows that simple acyclic set systems cannot contain union covered cycles.

Lemma 5.1. Let $G = (V, E)$ be a simple set system containing a union covered cycle $C = e_1, e_2, ..., e_q$ for some $q \geq 3$. Then G is cyclic.

Proof. Observe that none of the vertices contained in the edges of C are isolated since they are all contained in at least two edges of C. Hence no Graham reduction (i) can take place on any of these vertices unless one of the edges of C is removed first by a Graham reduction (ii).

Moreover, since G is simple, none of the edges of C will be contained in another edge of G. Hence no Graham reduction (ii) can take place on any of the edges of C unless one of their vertices is removed first by a Graham reduction (i). Hence any sequence of Graham reductions will preserve all the vertices and edges of C. □

Corollary 5.2. Let $G = (V, E)$ be a simple, k-uniform set system. For $\ell \geq k/2$, if G contains an ℓ-overlapping cycle, then G is cyclic.

Example 5.3. We can, however, have k-uniform, simple, acyclic set systems containing ℓ-overlapping cycles for $\ell < k/2$. As a straightforward example, consider $\ell = 1$. Let $k, c \geq 3$ and consider the set system $G = (V, E)$ with $V = \{1, ..., c(k-1)\}$ and

$$
\begin{aligned}
E = &\Big\{\{1, ..., k\}, \{k, ..., 2k-1\}, ..., \{(c-1)(k-1)+1, ..., c(k-1), 1\}\Big\} \\
&\cup \Big\{\{1, k, 2k-1, ..., k(k-1)-(k-2)\}\Big\} \\
&\cup \Big\{\{k^2-(k-1), k(k+1)-k, ..., k(2k-1)-(2k-2)\}\Big\} \\
&\qquad\vdots \\
&\cup \Big\{\{\left\lceil \frac{c}{k} \right\rceil k^2 - (\left\lceil \frac{c}{k} \right\rceil k - 1), ..., k\left(\left(\left\lceil \frac{c}{k} \right\rceil + 1\right)k - 1\right) - \left(\left(\left\lceil \frac{c}{k} \right\rceil + 1\right)k - 2\right)\Big\}
\end{aligned}
$$

The first line of this edge set describes a 1-overlapping cycle in the system. The edges can be removed as leaves sequentially as written above.

We now restrict our focus to $(k-1)$-overlapping geodesics. As noted in Section 1, successive vertices in the tracing of $(k-1)$-overlapping paths are uniquely defined by the list of edges, making these paths an intuitive field of study. Moreover, for $k = 2, 3$, we have shown that $(k-1)$-overlapping cycles are the only types of cycles that correspond to our notion of acyclicity. We will now generalize the relationship between the order and the size of trees to the case of $(k-1)$-overlap connected, acyclic, k-uniform set systems.

Theorem 5.4. Let $G = (V, E)$ be a simple, $(k-1)$-overlap connected, k-uniform set system of order n. Then G is acyclic if and only if it has size $n - k + 1$.

Proof. Suppose G is acyclic. Then G contains an ear, $e \in E$ and another edge $f \in E$ such that all the vertices of $e \backslash f$ are isolated. Since G is simple, $(k-1)$-overlap connected and k-uniform, we must have that $|e \cap f| = k - 1$ and $|e \backslash f| = 1$. Hence removing this ear will remove exactly one vertex and one edge. Clearly, removal of this ear will result in another simple, $(k-1)$-overlap connected, k-uniform set system, this time of order $n-1$. Inductively, we may remove ears in this way until we obtain a single edge containing k vertices. At this point we will have removed $n - k$ vertices and $n - k$ edges. Hence G must have originally contained n vertices and $n - k + 1$ edges.

Now, suppose G has size $n-k+1$. Since G is $(k-1)$-overlap connected and k-uniform, by counting we get that G must contain at least one isolated vertex v, contained in exactly one edge, say e. Now, since G is $(k-1)$-overlap connected and k-uniform, we also get that there must exist an edge f so that $|e \cap f| = k-1$ and hence $e \backslash f = \{v\}$. Thus e is an ear so we can remove it, and v, to obtain a simple, $(k-1)$-overlap connected, k-uniform set system of order $n-1$ and size $n-k$. Inductively, we may remove ears in this way until we obtain a single edge containing k vertices. All the vertices of this edge will then be isolated so the edge can be removed. □

Example 5.5. Observe that if $\ell < k-1$, we may have acyclic, ℓ-overlap connected, k-uniform set systems of various sizes. Indeed, consider the k-uniform ℓ-overlapping path with edge set:

$$\{1,2,...,k\}, \{k-\ell+1, k-\ell+2, ..., 2k-\ell\}, \{2k-2\ell+1, 2k-2\ell+2, ..., 3k-2\ell\}, ...$$

Clearly this set system is acyclic. But we may also successively add edges of the form $\{a, a+1, ..., a+k-1\}$ and preserve acyclicity since removing edges in lexicographic order from such a set system will provide a sequence of ear reductions leading to the empty set. Observe that for $\ell = k-1$, all such edges are already in the set system so this cannot be done.

6 Classifications of Acyclic, Uniform Set Systems of fixed Diameter

In this section, we classify those k-uniform sets systems of each positive $(k-1)$-overlap diameter that are acyclic. Our first result shows that (for sufficiently large n) no such set systems exist for $d = 1$.

Proposition 6.1. Let $G = (V, E)$ be a simple, k-uniform set system of order $n \geq k+1$ and ℓ-overlap diameter $d = 1$ for any $\ell = 1, ..., k-1$. Then G is cyclic.

Proof. Let $v \in V$. For each of the $n-1$ vertices $u \in V \backslash \{v\}$, we have $d_\ell(v, u) = 1$, so there must exist some $e_u \in E$ such that $u, v \in e_u$. But each e_u contains only k vertices so there must be at least two such edges containing v. Hence G has no isolated vertices (and is simple) and hence no ears. □

Note that for $n = k$, we have the trivial k-uniform set system of size 1 which is acyclic and has diameter 1.

We now consider the case of $d = 2$. In this case we restrict our focus to $\ell = k-1$ and show that under this assumption there is a unique acyclic set system for each n and k.

Proposition 6.2. Let $G = (V, E)$ be a simple, k-uniform set system of order $n \geq k+2$ and $(k-1)$-overlapping diameter $d = 2$. Then G is either cyclic or isomorphic to $G_{n,k}$ where $V(G_{n,k}) = \{1, 2, ...n\}$ and $E(G_{n,k}) = \{\{1, 2, ..., k-1, j\} : j = k, ..., n\}$.

Proof. Suppose G is acyclic. Then there exists a pair $e, f \in E$ such that all the vertices of $e \backslash f$ are isolated. Since G is $(k-1)$-overlap connected, we must have $|e \backslash f| = 1$ and $|e \cap f| = k-1$. Label the vertices of $e \cap f$ by $v_1, ..., v_{k-1}$ and the vertex in $e \backslash f$ by v_k. Label the remaining vertices $v_{k+1}, ..., v_n$. Now, since v_k is isolated in e and G has $(k-1)$-diameter 2, we must have $d_{k-1}(v_k, v_j) = 2$ for each $j = k+1, ..., n$, so there must exist $e_j = \{v_1, v_2, ..., v_{k-1}, v_j\} \in E$ for each $j = k+1, ..., n$, i.e.

$$\{e_j = \{v_1, v_2, ..., v_{k-1}, v_j\} : j = k, ..., n\} \subset E(G).$$

By Theorem 5.4, we get that $E(G)$ cannot contain any other edges. □

Corollary 6.3. By setting $k = 2$, we confirm that trees of diameter 2 have exactly one internal node that appears in every edge, with the remaining vertices all leaves.

We now turn our attention to the case of $d \geq 3$.

Theorem 6.4. Let $G = (V, E)$ be a simple, k-uniform set system of order $n > k + 2$ and $(k - 1)$-overlap diameter $d \geq 3$. Then G is acyclic if and only if there exists a $(k - 1)$-overlapping path $e_1, ..., e_d \in E$ such that:

(i) For every $f \in E \backslash \{e_1, ..., e_d\}$, there exists some $j \in \{2, ..., d - 1\}$ such that $|f \cap e_j| = k - 1$; and

(ii) All the vertices of $V \backslash \left(\bigcup_{j=2}^{d-1} e_j \right)$ are isolated.

Proof. Suppose G is acyclic. Since G has $(k - 1)$-overlap diameter d, it must contain an a $(k - 1)$-overlapping path $e_1, ..., e_d \in E$. Clearly this path contains $k + d - 1$ vertices. Thus

$$|V \backslash \left(\bigcup_{j=2}^{d-1} e_j \right) | = n - k - d + 1.$$

Moreover, by 5.4 we have that E contains $n - k + 1$ vertices. Hence

$$|E \backslash \{e_1, ..., e_d\}| = n - k - d + 1.$$

Since G is $(k - 1)$-overlap connected, (i) and (ii) follow. The converse is clearly true. $\square$

Corollary 6.5. By setting $k = 2$, we obtain that trees of diameter $d \geq 3$ contain a path connecting exactly $d + 1$ vertices such that the middle $d - 1$ vertices of this path are the only internal nodes of the tree (all others are leaves) and every edge contains at least one of these $d - 1$ vertices.

Remark. While Proposition 6.1 holds for all values of $\ell = 1, ..., k-1$, Proposition 6.2 and Theorem 6.4 cannot be generalized for lower values of ℓ-overlapping connectivity since the set systems may contain many other edges and fewer isolated vertices.

References

[1] B.D. Acharya and M.L. Vergnal. Hypergraphs with cyclomatic number zero, triangulated graphs, and an inequality. *J. Comb. Theory, Series B*, 33:52–56, 1992.

[2] C. Beeri, R. Fagin, D. Maier, and M. Yannakakis. On the desirability of acyclic database schemes. *JACM*, 30:479–513, 1983.

[3] C. Berge. *Hypergraphs.* North-Holland, Amsterdam, 1989.

[4] P. Duchet. Hypergraphs. In R. L. Graham, G. Grötschel, and L. Lovász, editors, *Handbook of combinatorics (vol. 1)*, pages 381–432. MIT Press, USA, 1995.

[5] D. R. Fulkerson and O. A. Gross. Incidence matrices and interval graphs. *Pacific J. Math.*, 15:838–858, 1965.

[6] L. Haizhu and J. Wang. On acyclic and cyclic hypergraphs. *J. Systems Science and Systems Engineering*, 15(4):253–262, 2002.

[7] H. Hàn and M. Schacht. Dirac-type results for loose hamilton cycles in uniform hypergraphs. *J. Comb. Theory, Series B*, 100(3):332–346, 2010.

[8] D. Kühn, R. Mycroft, and D. Osthus. Hamilton ℓ-cycles in uniform hypergraphs. *J. Comb. Theory, Series A*, 117(7):910–927, 2010.

[9] T.T. Lee. An information theoretic analysis of relational databases.Part II: information structures of database schemas. *IEEE Trans. Soft. Eng.*, SE-13(10):1062–1072, 1983.

[10] D. Palumbíny. On decompositions of complete graphs into factors with equal diameters. *Boll. Unione Mat. Ital.*, 7:420–428, 1973.

[11] P. Tomasta. Decomposition of graphs and hypergraphs into isomorphic factors with a given diameter. *Czech Math. J.*, 27(4):598–608, 1977.

[12] J. Wang. *Information Hypergraph Theory.* Science Press, Beijing, 2008.

[13] J. Wang and T.T. Lee. An invariant for hypergraphs. *Acta Mathematicae Apllicatae Sinica*, 12(2):113–120, 1996.