

# Codegree problems for projective geometries

Peter Keevash \*

Yi Zhao †

## Abstract

The codegree density  $\gamma(F)$  of an  $r$ -graph  $F$  is the largest number  $\gamma$  such that there are  $F$ -free  $r$ -graphs  $G$  on  $n$  vertices such that every set of  $r - 1$  vertices is contained in at least  $(\gamma - o(1))n$  edges. When  $F = PG_2(2)$  is the Fano plane Mubayi showed that  $\gamma(F) = 1/2$ . This paper studies  $\gamma(PG_m(q))$  for further values of  $m$  and  $q$ . In particular we have an upper bound  $\gamma(PG_m(q)) \leq 1 - 1/m$  for any projective geometry. We show that equality holds whenever  $m = 2$  and  $q$  is odd, and whenever  $m = 3$  and  $q$  is 2 or 3. We also give examples of 3-graphs with codegree densities equal to  $1 - 1/k$  for all  $k \geq 1$ .

## 1 Introduction

A fundamental problem in extremal combinatorics is to determine the Turán number  $\text{ex}(n, F)$ , which is the maximum number of edges in an  $r$ -graph (i.e.  $r$ -uniform hypergraph) on  $n$  vertices that does not contain a copy of the fixed  $r$ -graph  $F$ . It is not hard to show that the limit  $\pi(F) = \lim_{n \rightarrow \infty} \text{ex}(n, F) / \binom{n}{r}$  exists. It is usually called the Turán density of  $F$ . For ordinary graphs, when  $r = 2$ , Turán determined the exact value  $\text{ex}(n, K_t)$  for the complete graphs  $K_t$ . For a general graph  $F$  Erdős, Stone and Simonovits showed that the Turán density  $\pi(F)$  is  $1 - 1/(\chi(F) - 1)$ , where  $\chi(F)$  denotes the chromatic number of  $F$ . By contrast, when  $r > 2$  there are very few hypergraphs for which the Turán density is known. A deceptively simple conjecture of Turán is that the Turán density of  $K_4^3$ , the complete 3-graph on 4 vertices, should be  $5/9$ .

A natural variation on the Turán problem is to ask how large the minimum degree can be in an  $F$ -free  $r$ -graph. For any  $r$ -graph  $G$  let  $C(G)$  be the minimum degree  $d(S)$  of a set  $S$  of  $r - 1$  vertices, where  $d(S)$  is the number of edges of  $G$  containing  $S$ . The codegree extremal number  $\text{co-ex}(n, F)$  is the maximum possible value of  $C(G)$  where  $G$  is an  $F$ -free  $r$ -graph on  $n$  vertices. This parameter was introduced by Mubayi and Zhao [16]. They showed that the limit  $\gamma(F) = \lim_{n \rightarrow \infty} \text{co-ex}(n, F) / n$  exists and called it the codegree density of  $F$ . It is not hard to see that  $\gamma(F) \leq \pi(F)$  and that equality holds when  $r = 2$ , i.e.  $F$  is an ordinary graph.

The situation for hypergraphs is quite different: the only non-trivial values previously known were  $\gamma(PG_2(2)) = 1/2$  (Mubayi [14]) and  $\gamma(C_3^{2k}) = 1/2$  (attributed to Sudakov in [16]). Here  $PG_2(2)$  denotes

---

\*Department of Mathematics, Caltech, Pasadena, CA 91125, USA. Email: keevash@caltech.edu. Research supported in part by NSF grant DMS-0555755.

†Department of Mathematics and Statistics, Georgia State University, Atlanta, GA 30303, USA. Email: yizhao@mathstat.gsu.edu. Research supported in part by NSA grant H98230-06-1-0140.

the Fano plane, and  $C_3^{2k}$  is the  $2k$ -uniform hypergraph obtained by letting  $P_1, P_2, P_r$  be pairwise disjoint sets of size  $k$  and taking as edges all sets  $P_i \cup P_j$  with  $i \neq j$ . (The Turán problem for this latter hypergraph was solved by Keevash and Sudakov in [12].) Again for the seemingly simple example  $K_4^3$  there is a conjecture of Czygrinow and Nagle [3] that the codegree density should be  $1/2$ . One can also define the codegree density of a family  $\mathcal{F}$  of  $r$ -graphs as  $\gamma(\mathcal{F}) = \lim_{n \rightarrow \infty} \text{co-ex}(n, \mathcal{F})/n$ , where  $\text{co-ex}(n, \mathcal{F})$  is the maximum possible value of  $C(G)$ , where  $G$  is an  $r$ -graph on  $n$  vertices that is  $F$ -free for all  $F \in \mathcal{F}$ . The main result of [16] is that for any  $r \geq 3$ , the values of  $\gamma(\mathcal{F})$ , where  $\mathcal{F}$  ranges over finite families of  $r$ -graphs, are dense in  $[0, 1]$ . This contrasts sharply with the situation for graphs ( $r = 2$ ), where  $\gamma(\mathcal{F}) = \pi(\mathcal{F})$  can only take a value  $1 - 1/k$  for some  $k \geq 1$ .

In this paper we will consider the codegree problem when the forbidden hypergraph is  $PG_m(q)$ , i.e. the projective geometry of dimension  $m$  over the field with  $q$  elements. For the Fano plane (the case  $m = q = 2$ ) we have already mentioned the result of Mubayi [14] that  $\gamma(PG_2(2)) = 1/2$ . The Turán density has a different value: de Caen and Füredi [4] showed that  $\pi(PG_2(2)) = 3/4$ . (Later the exact Turán number for the Fano plane was determined independently and simultaneously by Keevash and Sudakov [11] and Füredi and Simonovits [7]. They showed that  $\text{ex}(n, PG_2(2)) = \binom{n}{3} - \binom{\lfloor n/2 \rfloor}{3} - \binom{\lceil n/2 \rceil}{3}$  for  $n$  sufficiently large. The exact codegree extremal number  $\text{co-ex}(n, PG_2(2))$  is unknown: Mubayi [14] conjectures that it should be  $\lceil n/2 \rceil$  for  $n$  sufficiently large.) For general  $PG_m(q)$  some bounds on the Turán densities were given by Keevash [10]. Our main result is an upper bound on the codegree density for general projective geometries, which is tight in many cases.

**Theorem 1.1** *The codegree density of projective geometries satisfies  $\gamma(PG_m(q)) \leq 1 - 1/m$ . Equality holds whenever  $m = 2$  and  $q$  is 2 or odd, and whenever  $m = 3$  and  $q$  is 2 or 3.*

Our proof uses some ideas from [14] but is independent of that paper, so we have a new (and slightly simpler) proof that  $\gamma(PG_2(2)) = 1/2$ .

For other values of  $m$  and  $q$  we cannot determine  $\gamma(PG_m(q))$  exactly, but we have some non-trivial bounds, which we will summarise with a table in the final section of the paper. Most of the bounds come either from Theorem 1.1 or previously known results on projective geometries;  $PG_2(4)$  is an exception, and it will become clearer later why the bound expressed in the following theorem is interesting.

**Theorem 1.2**  $\gamma(PG_2(4)) \geq 1/3$ .

With our next theorem, which arises as a byproduct of our method of proof, we can exhibit an infinite family of hypergraphs for which we can determine the codegree densities.

**Theorem 1.3** *For every  $r$  and  $i$  there is an  $r$ -graph  $F_i^r$  with  $\gamma(F_i^r) = 1 - 1/i$ .*

The rest of this paper is organised as follows. In the next section we prove the upper bound in Theorem 1.1, deriving it from a bound for a general hypergraph construction. Section 3 contains some constructions for lower bounds, including those asserted by Theorems 1.1 and Theorem 1.2. We prove Theorem 1.3 in section 4 and the final section contains some concluding remarks and an open problem.

**Definitions and notation.** An  $r$ -graph  $G$  has a vertex set  $V(G)$  and an edge set  $E(G)$ , where each element of  $E(G)$  is a subset of  $V(G)$  of size  $r$ . If  $G$  is an  $r$ -graph and  $S \subseteq V(G)$  then  $G[S]$  denotes the

restriction of  $G$  to  $S$ , which has vertex set  $S$  and edge set those edges of  $G$  that are contained in  $S$ . Let  $\mathbb{F}_q$  be the field with  $q$  elements, for any prime power  $q$ . The projective geometry  $PG_m(q)$  of dimension  $m$  over  $\mathbb{F}_q$  is a  $(q+1)$ -graph with vertex set equal to the one-dimensional subspaces of  $\mathbb{F}_q^{m+1}$  and edges corresponding to the two-dimensional subspaces of  $\mathbb{F}_q^{m+1}$ , in that for each two-dimensional subspace, the set of one-dimensional subspaces that it contains is an edge of the hypergraph  $PG_m(q)$ . A vertex of  $PG_m(q)$  can be described by projective co-ordinates as  $(x_1 : \cdots : x_{m+1})$ , where  $(x_1, \cdots, x_{m+1})$  is any non-zero vector of  $\mathbb{F}_q^{m+1}$  and  $(x_1 : \cdots : x_{m+1})$  denotes the one-dimensional subspace that it generates. This description is not unique, except in characteristic 2, and then we prefer to use  $(x_1, \cdots, x_{m+1})$  to denote both the vector and the space it generates. Given an  $r$ -graph  $G$  and  $S \subseteq V(G)$  the degree  $d_G(S)$ , or simply  $d(S)$ , of  $S$  is the number of edges containing  $S$ . When  $|S| = r - 1 \geq 2$  this is also called the codegree of  $S$ . We also define the neighbourhood  $N(S) = \{x \in V(G) : S \cup \{x\} \in E(G)\}$ , so that  $d(S) = |N(S)|$ .

## 2 The upper bound

In this section we prove the upper bound in Theorem 1.1. We will deduce it from a general bound on the codegree density of a hypergraph generated by a certain construction that we will now describe.

Given any  $r$ -graph  $F$  and a positive integer  $t$  the  $t$ -blowup  $F(t)$  is defined as follows. For each vertex  $x$  of  $F$  there are vertices  $x^1, \cdots, x^t$  of  $F(t)$ . For each edge  $x_1 \cdots x_r$  of  $F$  we have all  $t^r$  edges  $x_1^{i_1} \cdots x_r^{i_r}$  with  $1 \leq i_1, \cdots, i_r \leq t$  in  $F(t)$ . Also, for a family  $\mathcal{F}$  of  $r$ -graphs we define  $\mathcal{F}(t) = \{F(t) : F \in \mathcal{F}\}$ .

Let  $F^+$  be the  $r$ -graph obtained from  $F(r-1)$  by adding a new vertex  $y$  and all edges  $yx^1 \cdots x^{r-1}$  where  $x$  is a vertex of  $F$ .

**Theorem 2.1** *Suppose  $F$  is an  $r$ -graph. Then  $\gamma(F^+) \leq \frac{1}{2-\gamma(F)}$ .*

To prove this theorem we need two lemmas.

**Lemma 2.2** *(see Lemma 2.1 in [16]) For any  $\epsilon > 0$  there is an integer  $K_1$ , so that if  $n \geq k \geq K_1$  and  $G$  is an  $r$ -graph on a set of  $n$  vertices with  $C(G) > (\delta + 2\epsilon)n$ , then at least  $\frac{1}{2} \binom{n}{k}$  induced subhypergraphs  $H$  of  $G$  on  $k$  vertices have  $C(H) > (\delta + \epsilon)k$ .*

For the convenience of the reader we include a brief summary of the proof given in [16].

**Proof.** Given sets of vertices  $S, T$  with  $|S| = k$ ,  $|T| = r - 1$ ,  $T \subseteq S$  we say that  $T$  is bad for  $S$  if  $|N(T) \cap S| \leq (\delta + \epsilon)k$ . Suppose  $S$  is a random set of size  $k$ . Consider any  $T \subseteq S$  of size  $r - 1$ . Using the terminology of [9] (page 29)  $|N(T) \cap S|$  has the hypergeometric distribution with parameters  $n$ ,  $k$ ,  $|N(T)|$ , which has mean  $\lambda = k|N(T)|/n$ . Since  $|N(T)| = d(T) \geq C(G) > (\delta + 2\epsilon)n$  we have  $\lambda > (\delta + 2\epsilon)k$  and we can apply Theorem 2.10 of [9] to get  $\mathbb{P}(|N(T) \cap S| \leq (\delta + \epsilon)k) \leq 2e^{-(\epsilon/\delta)^2 \lambda/3} = 2^{-\Theta(k)}$ . By the union bound, the probability that there is any  $T$  bad for  $S$  is at most  $\binom{k}{r-1} 2^{-\Theta(k)} < 1/2$  for large  $k$ . If no  $T$  is bad for  $S$  then  $H = G[S]$  satisfies  $C(H) > (\delta + \epsilon)k$ .  $\square$

**Lemma 2.3** *Suppose  $\mathcal{F}$  is a finite family of  $r$ -graphs and  $t$  is a positive integer. Then  $\gamma(\mathcal{F}(t)) = \gamma(\mathcal{F})$ .*

**Proof.** Our argument combines the ideas of Proposition 1.4 of [16] and Lemma 4 of [15]. Write  $\delta = \gamma(\mathcal{F})$  and  $\ell = |\mathcal{F}|$ . Suppose  $\epsilon > 0$ . Choose  $K_1$  so that Lemma 2.2 applies, and then  $k \geq K_1$  large enough that any  $r$ -graph  $H$  on  $k$  vertices with  $C(H) > (\delta + \epsilon)k$  contains some  $F \in \mathcal{F}$ . Let  $G$  be an  $r$ -graph on  $n$  vertices with  $C(G) > (\delta + 2\epsilon)n$ , where  $n$  is sufficiently large. By Lemma 2.2 at least  $\frac{1}{2} \binom{n}{k}$  induced subhypergraphs  $H$  of  $G$  on  $k$  vertices have  $C(H) > (\delta + \epsilon)k$ . Each such subhypergraph contains a hypergraph from  $\mathcal{F}$ , by choice of  $k$ , and so some  $F \in \mathcal{F}$  is contained in at least  $\frac{1}{2\ell} \binom{n}{k}$  induced subhypergraphs  $H$  of  $G$  on  $k$  vertices. Write  $f = |V(F)|$  and let  $J$  be the  $f$ -graph consisting of all  $f$ -subsets of  $V(G)$  that span a copy of  $F$ . Then  $|E(J)| \geq \frac{1}{2\ell} \binom{n}{k} \binom{n-f}{k-f}^{-1} = \left(2\ell \binom{k}{f}\right)^{-1} \binom{n}{f}$ . Now by a result of Erdős [5], for any fixed  $T$  and  $n$  large  $J$  contains the  $T$ -blowup of a single  $f$ -set. A standard result of Ramsey theory implies that for large  $T$  this will contain the  $t$ -blowup of a single  $f$ -set for which all the copies of  $F$  align to form  $F(t)$ .  $\square$

**Proof of Theorem 2.1.** Write  $\gamma = \gamma(F)$  and  $\gamma^+ = \frac{1}{2-\gamma}$ . Suppose  $0 < \epsilon < 2 - \gamma - \gamma^+$  is given. Let  $K_1$  be the integer given by Lemma 2.2, and choose  $K_2$  so that if  $k \geq K_2$  and  $G$  is an  $r$ -graph on a set of  $k$  vertices with  $C(G) > (\gamma + \epsilon)k$  then  $F \subseteq G$ . Choose  $k > \max\{K_1, 2K_2, (2 - \gamma - \gamma^+ - \epsilon)^{-1}\}$ . Let  $\mathcal{H}$  be the set of all  $r$ -graphs  $H$  on  $k$  vertices with  $C(H) \geq (\gamma^+ + \epsilon)k$ . Now  $\gamma(\mathcal{H}) \leq \gamma^+ + 2\epsilon$ , since if  $G$  is an  $r$ -graph on  $n$  vertices with  $n$  large and  $C(G) \geq (\gamma^+ + 2\epsilon)n$  then by Lemma 2.2 there is an induced subgraph  $H$  of  $G$  on  $k$  vertices with  $C(H) \geq (\gamma^+ + \epsilon)k$ , i.e.  $G$  contains a member of  $\mathcal{H}$ . Now by Lemma 2.3 we have  $\gamma(\mathcal{H}(r-1)) \leq \gamma^+ + 2\epsilon$ .

Let  $G$  be an  $r$ -graph on  $n$  vertices with  $C(G) \geq (\gamma^+ + 3\epsilon)n$ , and suppose  $n$  is large. By the previous paragraph we can find  $H(r-1) \subseteq G$ , for some  $H \in \mathcal{H}$ . Recall that for a vertex  $x$  of  $H$  we denote its copies in  $H(r-1)$  by  $x^1, \dots, x^{r-1}$ . For each  $x \in V(H)$  let  $C_x = \{y \in V(G) \setminus V(H(r-1)) : x^1 \dots x^{r-1}y \in E(G)\}$ , and for each  $y \in V(G) \setminus V(H(r-1))$  let  $D_y = \{x \in V(H) : x^1 \dots x^{r-1}y \in E(G)\}$ . By the codegree assumption  $|C_x| \geq (\gamma^+ + 3\epsilon)n - (r-1)k > (\gamma^+ + \epsilon)n$  for all  $x$ . Now

$$k \cdot (\gamma^+ + \epsilon)n < \sum_{x \in V(H)} |C_x| = \sum_{y \in V(G) \setminus V(H(r-1))} |D_y|,$$

so we can find some  $y$  with  $|D_y| > (\gamma^+ + \epsilon)k$ . Choose  $X \subseteq D_y$  with  $|X| = \lceil (\gamma^+ + \epsilon)k \rceil$ .

By definition of  $\mathcal{H}$ , we have  $C(H) \geq (\gamma^+ + \epsilon)k$ . So for any vertices  $x_1, \dots, x_{r-1}$  in  $X$  the number of vertices  $w$  in  $X$  such that  $x_1 \dots x_{r-1}w$  is an edge is at least

$$(\gamma^+ + \epsilon)k - |V(H) \setminus X| \geq (2\gamma^+ - 1 + 2\epsilon)k = \left(\frac{\gamma}{2-\gamma} + 2\epsilon\right)k.$$

It follows that  $C(H[X]) \geq \left(\frac{\gamma}{2-\gamma} + 2\epsilon\right)k > (\gamma + \epsilon)|X|$ , as  $|X| = \lceil (\gamma^+ + \epsilon)k \rceil$  and  $1/k < 2 - \gamma - \gamma^+ - \epsilon$ . Since  $|X| > \gamma^+k > K_2$  we can find a copy of  $F$  in  $H[X]$ . Then  $F(r-1) \cup \{y\}$  spans a copy of  $F^+$  in  $G$ . Since  $\epsilon$  was arbitrary, we have  $\gamma(F^+) \leq \gamma^+$ , as required.  $\square$

**Proof that  $\gamma(PG_m(q)) \leq 1 - 1/m$ .** We argue by induction on  $m$ . Note that  $PG_1(q)$  consists of a single edge, and so trivially  $\gamma(PG_1(q)) = 0$ . For  $m > 1$  we will show that  $PG_m(q) \subseteq PG_{m-1}(q)^+$ . Then by the bound  $\gamma(F^+) \leq \frac{1}{2-\gamma(F)}$  just proved we will have  $\gamma(PG_m(q)) \leq \frac{1}{2-(1-1/(m-1))} = 1 - 1/m$ , as required. Construct  $G = PG_{m-1}(q)^+$  from  $F = PG_{m-1}(q)$  with vertex set  $X$  and an extra vertex  $y$  by adding all edges  $yx^1 \dots x^q$  with  $x \in X$  to  $F(q)$ . Choose projective co-ordinates for  $X$  and for each  $(a_1 : \dots : a_m)$  in  $X$  (not all  $a_i$  equal to zero) identify its  $q$  copies in  $F(q)$  with the points  $(a_1 : \dots : a_m : b)$  in  $PG_m(q)$ ,

as  $b$  ranges over  $\mathbb{F}_q$ . Identify the point  $y$  with  $(0 : \cdots : 0 : 1)$  in  $PG_m(q)$ . Thus we have identified all points of  $PG_m(q)$ . Furthermore, every projective line is present as an edge of  $G$ . For in any line containing  $y = (0 : \cdots : 0 : 1)$  the other  $q$  points are of the form  $(a_1 : \cdots : a_m : b)$  as  $b$  ranges over  $\mathbb{F}_q$ , and these are all edges by definition of  $G$ . Also, for any line  $\{a^1, \dots, a^{q+1}\}$  of  $PG_m(q)$  not containing  $y$ , where  $a^i = (a_1^i : \cdots : a_{m+1}^i)$ , the line  $\{(a_1^i : \cdots : a_m^i)\}_{i=1}^{q+1}$  of  $PG_{m-1}(q)$  is present as an edge of  $F$ , and so  $\{(a_1^i : \cdots : a_{m+1}^i)\}_{i=1}^{q+1}$  is present as an edge of  $G$  by definition of the blowup  $F(q)$ . Therefore we have a copy of  $PG_m(q)$ .  $\square$

### 3 Constructions

Now we will describe the constructions giving what lower bounds we know on  $\gamma(PG_m(q))$ . Many of these are derived from previously known results, but the bounds expressed in Theorems 1.1 and 1.2 require new constructions and analysis. First we give some definitions.

Given a set of vertices  $V$  and an integer  $t \geq 1$  we define an  $r$ -graph  $G_t(V)$  on  $V$  as follows. Partition  $V$  as  $V_1 \cup \cdots \cup V_t$  so that  $||V_i| - |V|/t| < 1$  for  $1 \leq i \leq t$ . The edges are all  $r$ -tuples that are not contained in any  $V_i$ . The chromatic number  $\chi(F)$  of an  $r$ -graph  $F$  is the smallest  $t$  such that  $F \subseteq G_t(V)$  for some  $V$ .

We can bound the codegree density of  $F$  in terms of its chromatic number by  $\gamma(F) \geq 1 - 1/(\chi(F) - 1)$ . This follows as  $F$  is not contained in  $G_t(V)$  with  $t = \chi(F) - 1$  and  $C(G_t(V)) > (1 - 1/t)|V| - 1$ . Known values of  $\chi(PG_m(q))$  are:  $\chi(PG_2(q)) = 2$  for all  $q > 2$ ,  $\chi(PG_3(2)) = \chi(PG_2(2)) = 3$ ,  $\chi(PG_4(2)) = 4$  (Pelikán [17]), and  $\chi(PG_5(2)) = 5$  (Fugère, Haddad and Wehlau [6]). This supplies the additional bounds  $\gamma(PG_4(2)) \geq \frac{2}{3}$  and  $\gamma(PG_5(2)) \geq \frac{3}{4}$ . Since  $PG_a(q) \subseteq PG_b(q)$  for  $a \leq b$  we always have  $\gamma(PG_a(q)) \leq \gamma(PG_b(q))$  for  $a \leq b$ . It follows from work of Haddad [8] that  $\lim_{m \rightarrow \infty} \chi(PG_m(q)) = \infty$ . Therefore, for any  $k, q$  there is  $m_{k,q}$  such that for  $m > m_{k,q}$  we have  $\chi(PG_m(q)) > k$ , and so  $\gamma(PG_m(q)) \geq 1 - 1/k$ . In particular  $\lim_{m \rightarrow \infty} \gamma(PG_m(q)) = 1$ .

The theory of blocking sets in projective geometries can be used to give a non-trivial lower bound for some other cases. Say that  $S \subseteq PG_m(q)$  is a blocking set if  $0 < |S \cap L| < |L|$  for every line  $L$  of  $PG_m(q)$ . Clearly  $PG_m(q)$  has a blocking set if and only if it has chromatic number 2. Cassetta [2] verified by computer search that  $PG_3(4)$  does not have a blocking set, and so we have  $\gamma(PG_m(4)) \geq 1/2$  for  $m \geq 3$ . On the other hand, Tallini [18] showed that  $PG_3(q)$  has a blocking set for any  $q \geq 5$ . We are not aware of any other values of  $m$  and  $q$  for which it is known that  $PG_m(q)$  does not have a blocking set.

This completes the description of bounds following from previously known results, and now we will give some new constructions to prove Theorems 1.1 and 1.2.

#### 3.1 $PG_2(q)$ , $q$ odd

Let  $V$  be a set of  $n$  vertices. Partition it as  $V = V_1 \cup V_2$  so that  $||V_i| - n/2| < 1$  for  $i = 1, 2$ . Let  $G$  be the  $(q + 1)$ -graph in which the edges are all  $(q + 1)$ -tuples  $e$  with both  $|e \cap V_1|$  and  $|e \cap V_2|$  odd. Then  $C(G) > (1/2 - o(1))n$ . To see that  $G$  does not contain  $PG_2(q)$ , suppose to the contrary that it is possible to label the vertices of  $PG_2(q)$  from  $\{1, 2\}$  so that on each edge the number of 1's and number

of 2's are both odd. Label the lines as  $L_1, \dots, L_r$ , where  $r = q(q+1) + 1$  is odd. Let  $s_i$  be the number of 1's on  $L_i$  and  $s$  be the total number of 1's. Each point belongs to  $q+1$  lines, so  $(q+1)s = \sum_{i=1}^r s_i$ . However the left side is even (since  $q$  is odd) and the right side is odd (a sum of an odd number of odd numbers), contradiction.

### 3.2 $PG_3(2)$

Let  $V$  be a set of  $n$  vertices. Partition it as  $V = V_1 \cup V_2 \cup V_3$  so that  $||V_i| - n/3| < 1$  for  $i = 1, 2, 3$ . Let  $G$  be the 3-graph consisting of all triples with 2 points in  $V_i$  and 1 point in  $V_j$ , for any  $i \neq j$ . Then  $C(G) > (2/3 - o(1))n$ . To see that  $G$  does not contain  $PG_3(2)$  it suffices to show that in any 3-colouring of the vertices of  $PG_3(2)$  there will either be a monochromatic edge (having 3 vertices of the same colour) or a rainbow edge (having vertices of all different colours). Consider a 3-colouring of the vertices of  $PG_3(2)$  with no monochromatic edge. Pelikán [17] (Theorem 3.1) showed that in such a colouring there are 5 points of each colour, and any Fano subhypergraph contains at most 3 points in each colour. Consider any Fano subhypergraph, and identify it with  $\{a \in \mathbb{F}_2^3 : a \neq 0\}$  in such a way that its edges are all triples  $abc$  with  $a + b = c$ . Some colour, say red, occurs on at least 3 points, so on exactly 3 points of the Fano plane. These 3 points do not form an edge (it would be monochromatic), so without loss of generality they are  $(1, 0, 0)$ ,  $(0, 1, 0)$  and  $(0, 0, 1)$ . Without loss of generality, the edge  $(0, 1, 1)$ ,  $(1, 0, 1)$ ,  $(1, 1, 0)$  is coloured blue, blue, green. Now if  $(1, 1, 1)$  is blue then  $(0, 0, 1)$ ,  $(1, 1, 1)$ ,  $(1, 1, 0)$  is a rainbow edge, and if  $(1, 1, 1)$  is green then  $(1, 0, 0)$ ,  $(0, 1, 1)$ ,  $(1, 1, 1)$  is a rainbow edge.

### 3.3 $PG_3(3)$

Let  $V$  be a set of  $n$  vertices. Partition it as  $V = V_1 \cup V_2 \cup V_3$  so that  $||V_i| - n/3| < 1$  for  $i = 1, 2, 3$ . Let  $G$  be the 4-graph consisting of all quadruples  $e$  that are *odd*, meaning that at least one of  $|e \cap V_i|$  is odd. In fact, if  $e$  is odd then  $(|e \cap V_1|, |e \cap V_2|, |e \cap V_3|)$  is a rearrangement of  $(3, 1, 0)$  or  $(2, 1, 1)$ . Then  $C(G) > (2/3 - o(1))n$ . Say that a 3-colouring of  $PG_m(3)$  is odd if each line is odd, in that some colour appears on an odd number of points of that line. We need to show that there is no odd colouring of  $PG_3(3)$ .

First we claim that in any odd 3-colouring of  $PG_2(3)$  as  $C_1 \cup C_2 \cup C_3$  each  $|C_i|$  is odd. To see this, note first that an odd colouring gives a partition of the 13 lines of  $PG_2(3)$  into 3 classes  $L_1, L_2, L_3$ , where  $L_i$  is the set of lines determined by  $C_i$ , i.e. the set of lines with at least 2 points of colour  $i$ . For each  $i$ , colour  $i$  does not contain any line, so  $|L_i| = \ell_2(C_i) + \ell_3(C_i)$ , where  $\ell_j(C_i)$  denotes the number of lines with exactly  $j$  points of colour  $i$ . Note that  $\ell_2(C_i) + \ell_3(C_i)$  has the same parity as  $\ell_2(C_i) + 3\ell_3(C_i) = \binom{|C_i|}{2}$ . This is even if  $|C_i|$  is 0 or 1 mod 4, odd if  $|C_i|$  is 2 or 3 mod 4. Now  $|C_1| + |C_2| + |C_3| = 13 = 1 \pmod{4}$ , so the possibilities for  $|C_i| \pmod{4}$  up to permutation are  $(0, 0, 1)$ ,  $(0, 2, 3)$ ,  $(1, 2, 2)$ ,  $(1, 1, 3)$ ,  $(3, 3, 3)$ . These respectively give the following possibilities for  $|L_i| \pmod{2}$ :  $(0, 0, 0)$ ,  $(0, 1, 1)$ ,  $(0, 1, 1)$ ,  $(0, 0, 1)$ ,  $(1, 1, 1)$ . Since  $|L_1| + |L_2| + |L_3| = 13$  is odd, only the last two possibilities may hold, so each  $|C_i|$  is odd.

Now suppose there is an odd 3-colouring of  $PG_3(3)$  as  $C_1 \cup C_2 \cup C_3$ . Fix a point  $x \in C_1$  and count pairs  $(y, P)$  where  $y \in C_2$  and  $P$  is a plane<sup>1</sup> containing  $x$  and  $y$ . For each  $y$  there are 4 planes containing  $x$  and  $y$  so the number of pairs  $(y, P)$  is  $4|C_2|$ , which is even. However, there are 13 planes containing  $x$ ,

<sup>1</sup>A plane is a set of projective points corresponding to some 3-dimensional subspace of  $\mathbb{F}_3^4$ .

and each contains an odd number of points from  $C_2$  by the above claim, which implies that the number of pairs  $(y, P)$  is odd, contradiction.

### 3.4 $PG_2(4)$

Consider a set of  $n$  vertices  $V$  and partition  $V$  into  $V_1, V_2, V_3$  with  $||V_i| - n/3| < 1$  for  $i \in \{1, 2, 3\}$ . We will define a 5-graph  $G$  on  $V$  that does not contain  $PG_2(4)$  with the property that for any set  $S$  of 4 vertices, there is some  $i \in \{1, 2, 3\}$  such that  $S \cup \{v\}$  is an edge for each  $v \in V_i$ . Say that a set of vertices has type  $abc$  if it has  $a$  points in  $V_1$ ,  $b$  points in  $V_2$  and  $c$  points in  $V_3$ . The construction of  $G$  is to take all 5-tuples with one of the following types: 410, 140, 104, 401, 131, 113, 122, 212, 221.

First we verify the codegree property, i.e. that all possible types  $abc$  for a set of 4 vertices are obtained from a type of  $G$  by decreasing one number by 1. If  $abc$  is a permutation of 400 it can be obtained from 410, 140, or 104. If  $abc$  is a permutation of 310, then if the 3 is in position 1 then it can be obtained from 410 or 401, and otherwise from 131 or 113. If  $abc$  is a permutation of 220 or 211 it can be obtained from a permutation of 221.

To show that  $G$  does not contain  $PG_2(4)$  we need to describe the classification of blocking sets in  $PG_2(4)$  given by Berardi and Eugeni [1]. A blocking set is a set of points that meets every line but does not contain a line. Note that the complement of a blocking set is a blocking set. The three main ingredients are the following constructions:

1. A Baer subplane  $F$ . One way to describe this is to represent  $PG_2(4)$  as  $\{A + x : x \in \mathbb{Z}_{21}\}$ , where  $A = \{3, 6, 7, 12, 14\}$ . An example of a Baer subplane is  $F = \{x : x \equiv 0 \pmod{3}\}$ . Note that every line intersects  $F$  in 1 or 3 points; in particular it is a blocking set. If one takes the lines that intersect  $F$  in 3 points and restrict them to  $F$  the resulting configuration can be regarded as  $\{B + x : x \in \mathbb{Z}_7\}$ , where  $B = \{1, 2, 4\}$ , i.e. the Fano plane.
2. Fix 3 lines  $L, M, N$  with no common point. Choose a point  $w \in N \setminus (L \cup M)$ . Then  $\Delta = ((L \cup M) \setminus N) \cup \{w\}$  is a blocking set.
3. Fix 3 lines  $L, M, N$  with no common point. Let  $L' = L \setminus (M \cup N)$ ,  $M' = M \setminus (N \cup L)$  and  $N' = N \setminus (L \cup M)$ . Then  $H = L' \cup M' \cup N'$  is a blocking set.

The classification of Berardi and Eugeni [1] is that if  $B$  is a blocking set, then  $B$  or its complement is one of the following:  $F, F \cup \{x\}, \Delta, \Delta \cup \{x\}, H, \Delta \cup \{x, y\}$ . (Here  $x$  and  $y$  denote arbitrary additional points, subject to the condition that the resulting set does not contain a line.)

Now we show that  $G$  does not contain  $PG_2(4)$ . Suppose to the contrary that we have a copy  $P$  of  $PG_2(4)$  in  $G$  and write  $P_i = V(P) \cap V_i$ . Note that  $P_1$  is a blocking set of  $P$ , so we can divide into cases using the classification. We will use the notation  $\ell(u, v)$  for the line that contains  $u$  and  $v$ .

1.  $F \subseteq P_1 \subseteq F \cup \{x\}$ . Then any line that has 3 points of  $F$  and does not contain  $x$  has exactly 3 points in  $P_1$ , so is not in  $G$ .
2.  $P \setminus P_1 = F$ . Since  $F$  is a Fano plane it is not 2-colourable, so without loss of generality one of its lines (of size 3) is contained in  $P_2$ , since  $V_2$  and  $V_3$  are interchangeable. In  $PG_2(4)$  this extends to an edge of type 230, which is not in  $G$ .

3.  $P \setminus P_1 = F \cup \{x\}$ . Note that  $x$  lies on at most one line that intersects  $F$  in more than one point, as any two such lines intersect in a point of  $F$ . Thus we can choose  $y \in F$  so that  $\ell(x, y)$  does not contain any other point of  $F$ . Then  $\ell(x, y)$  has exactly 3 points in  $P_1$ , so is not in  $G$ .
4.  $\Delta \subseteq P_1 \subseteq \Delta \cup \{x, y\}$ . Choose  $a \in L \setminus (M \cup N)$  that does not belong to either  $\ell(w, x)$  or  $\ell(w, y)$ . Then  $\ell(w, a)$  has exactly 3 points in  $P_1$ , so is not in  $G$ .
5.  $\Delta \subseteq P \setminus P_1 \subseteq \Delta \cup \{x, y\}$ . If  $\{x, y\} = N \setminus (L \cup M \cup \{w\})$  then  $\ell(L \cap M, w)$  has exactly 3 points in  $P_1$ . Otherwise we can choose  $a \in N \setminus (L \cup M \cup \{w, x, y\})$  and  $b \in L \setminus (M \cup N)$  that does not belong to either  $\ell(a, x)$  or  $\ell(a, y)$ , and then  $\ell(a, b)$  has exactly 3 points in  $P_1$ , so is not in  $G$ .
6.  $P_1 = H$ . Any line that does not contain  $L \cap M$ ,  $M \cap N$  or  $N \cap L$  has exactly 3 points in  $P_1$ , so is not in  $G$ .
7.  $P \setminus P_1 = H$ . Note that none of  $L', M', N'$  is contained in  $P_2$  or  $P_3$ , as this would give a line of type 230 or 203. So without loss of generality  $|L' \cap P_2| = |M' \cap P_2| = 2$ . Write  $L' \cap P_2 = \{x, y\}$  and choose  $z \in N' \cap P_2$ . Now one of  $\ell(x, z)$  and  $\ell(y, z)$  must contain a point of  $M' \cap P_2$ . But then it has type 230, so is not in  $G$ .

In all cases we have a contradiction, so  $G$  does not contain  $PG_2(4)$ .

## 4 Proof of Theorem 1.3

Suppose  $r \geq 2$  and let  $F_1$  be the  $r$ -graph that consists of a single edge. Define  $F_{i+1} = F_i^+$  for  $i \geq 1$ . We will show by induction on  $i$  that  $\chi(F_i) = i + 1$  and  $\gamma(F_i) = 1 - 1/i$  for all  $i$ .

The base case  $i = 1$  is obvious. Now  $F_i$  is obtained from a copy of  $F_{i-1}(r - 1)$  and an extra vertex  $y$  by adding all edges  $yx^1 \cdots x^{r-1}$  with  $x \in F_{i-1}$ . By induction hypothesis there is a proper colouring  $c_{i-1} : V(F_{i-1}) \rightarrow \{1, \dots, i\}$ . Then we can construct a proper colouring  $c_i : V(F_i) \rightarrow \{1, \dots, i + 1\}$  of  $F_i$  given by  $c_i(x^j) = c(x)$  for  $x \in F_{i-1}$  and  $c_i(y) = i + 1$ , so  $\chi(F_i) \leq i + 1$ . Now suppose for a contradiction that we have a proper colouring  $c : V(F_i) \rightarrow \{1, \dots, i\}$ . Without loss of generality  $c(y) = i$ . For each  $x \in F_{i-1}$  there is at least one copy  $x^j$  with  $c(x^j) \neq j$ , since  $yx^1 \cdots x^{r-1}$  is not monochromatic. Choose such a copy  $x'$ . Then  $c$  restricts to a  $(i - 1)$ -colouring of  $\{x' : x \in F_{i-1}\}$ , which is a copy of  $F_{i-1}$ . This contradiction shows that  $\chi(F_i) = i + 1$ .

As observed above, this bound on the chromatic number implies that  $\gamma(F_i) \geq 1 - 1/i$ . But by the induction hypothesis and Theorem 2.1 we have  $\gamma(F_i) \leq \frac{1}{2 - \gamma(F_{i-1})} = \frac{1}{2 - (1 - 1/(i-1))} = 1 - 1/i$ . This completes the proof.  $\square$

## 5 Concluding remarks

1. Our results can be summarised by the following table, in which the entry in the cell indexed by row  $m$  and column  $q$  is either a number indicating the exact value of  $\gamma(PG_m(q))$  or an interval in which  $\gamma(PG_m(q))$  lies.

| $m \setminus q$ | 2              | 3              | 4              | $2^t, t \geq 3$ | $p^t, p$ odd   |
|-----------------|----------------|----------------|----------------|-----------------|----------------|
| 2               | 1/2            | 1/2            | [1/3, 1/2]     | [0, 1/2]        | 1/2            |
| 3               | 2/3            | 2/3            | [1/2, 2/3]     | [0, 2/3]        | [1/2, 2/3]     |
| 4               | [2/3, 3/4]     | [2/3, 3/4]     | [1/2, 3/4]     | [0, 3/4]        | [1/2, 3/4]     |
| $m \geq 5$      | [3/4, 1 - 1/m] | [2/3, 1 - 1/m] | [1/2, 1 - 1/m] | [0, 1 - 1/m]    | [1/2, 1 - 1/m] |

It is interesting to compare our bounds on codegree densities of projective geometries with those obtained for the Turán densities  $\pi(PG_m(q))$  in [10]. There it is shown that

$$\prod_{i=1}^q \left( 1 - \frac{i}{\sum_{j=1}^m q^j} \right) \leq \pi(PG_m(q)) \leq 1 - \frac{1}{\binom{q^m}{q}}.$$

Slightly better bounds are given for  $q = 2$ , and a more detailed analysis is given of the case  $PG_3(2)$ , with a lower bound of approximately 0.844 and an upper bound of 13/14. We see that the Turán densities are always considerably larger than the codegree densities.

It seems harder to give good constructions for codegree densities than Turán densities, and the contrast in our knowledge is particularly striking when the field size  $q$  is a power of 2. Here the lower bounds on the Turán density given in [10] are quite respectable, but we cannot show in general that the codegree density is even non-zero! Even to obtain a non-trivial lower bound on  $PG_2(4)$  we needed an ad hoc construction and some complicated analysis. But we are left with the following natural question.

**Open problem.** Is  $\gamma(PG_m(q)) > 0$  for all  $m$  and  $q$ ?

**2.** The argument in the proof of Theorem 1.3 can be generalised to show that, for any  $r$ -graph  $F$  and integer  $t$ , if  $\chi(F) = t + 1$  and  $\gamma(F) = 1 - 1/t$  then  $\chi(F^+) = t + 2$  and  $\gamma(F^+) = 1 - 1/(t + 1)$ . This can be used to give some other (closely related) infinite sequences of 3-graphs with codegree densities equal to  $1 - 1/i$ ,  $i \geq 2$  by taking  $G_2$  to be any 3-graph with  $PG_2(2) \subset G_2 \subset PG_1(2)^+$  and  $G_{i+1} = G_i^+$  for  $i \geq 2$ .

**3.** For future investigations of codegree densities it may be helpful to note another general property for an  $r$ -graph  $F$  that implies  $\gamma(F^+)$  is equal to the upper bound  $\frac{1}{2-\gamma(F)}$  given by Theorem 1.1. Consider any  $r$ -graph  $F$  with the property that  $F \subseteq F[B]^+$  for any  $B$  that intersects all edges of  $F$ . Given a vertex set  $V$ , construct an  $r$ -graph  $G$  by choosing a partition  $V = V_1 \cup V_2$ , letting  $G[V_1]$  be an  $F$ -free  $r$ -graph with largest possible minimum codegree on  $V_1$ , and adding all  $r$ -tuples that intersect both  $V_1$  and  $V_2$ . It is not hard to show that  $G$  is  $F^+$ -free and the partition can be chosen appropriately to give the lower bound  $\gamma(F^+) \geq \frac{1}{2-\gamma(F)}$ : we omit the details. It is not clear whether this property holds for any projective geometry  $PG_m(q)$ . It is certainly not applicable if  $PG_m(q)$  contains a blocking set, but it might apply to other cases. Even if it does not, it may prove useful in determining codegree densities for other hypergraphs.

**Acknowledgement.** We thank the referees for some helpful comments on the presentation of this paper.

## References

- [1] L. Berardi and F. Eugeni, Blocking sets in the projective plane of order four, Combinatorics '86 (Trento, 1986), 43–50, *Ann. Discrete Math.* **37**, North-Holland, Amsterdam, 1988.
- [2] N. Cassetta, The non-existence of blocking sets in  $\text{PG}(3, 4)$ , *Ital. J. Pure Appl. Math.* **3** (1998), 79–84 (1999).
- [3] A. Czygrinow and B. Nagle, A note on codegree problems for hypergraphs, *Bull. Inst. Combin. Appl.* **32** (2001), 63–69.
- [4] D. de Caen and Z. Füredi, The maximum size of 3-uniform hypergraphs not containing a Fano plane, *J. Combinatorial Theory B* **78** (2000), 274–276.
- [5] P. Erdős, On extremal problems of graphs and generalized graphs. *Israel J. Math.* **2** (1964), 183–190.
- [6] J. Fugère, L. Haddad and D. Wehlau, 5-chromatic Steiner triple systems, *J. Combin. Des.* **2** (1994), 287–299.
- [7] Z. Füredi and M. Simonovits, Triple systems not containing a Fano Configuration, *Combin. Probab. Comput.*, **14** (2005), 467–484.
- [8] L. Haddad, On the chromatic numbers of Steiner triple systems, *J. Combin. Des.* **7** (1999), 1–10.
- [9] S. Janson, T. Łuczak, A. Ruciński, Random graphs. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience, New York, 2000.
- [10] P. Keevash, The Turán problem for projective geometries, *J. Combin. Theory Ser. A* **111** (2005), 289–309.
- [11] P. Keevash and B. Sudakov, The exact Turán number of the Fano plane, *Combinatorica* **25** (2005), 561–574.
- [12] P. Keevash and B. Sudakov, On a hypergraph Turan problem of Frankl, *Combinatorica* **25** (2005), 673–706.
- [13] F. Mazzocca and G. Tallini, On the nonexistence of blocking-sets in  $\text{PG}(n, q)$  and  $\text{AG}(n, q)$ , for all large enough  $n$ , *Simon Stevin* **59** (1985), 43–50.
- [14] D. Mubayi, The co-degree density of the Fano plane, *J. Combin. Theory Ser. B* **95** (2005), 333–337.
- [15] D. Mubayi, A hypergraph extension of Turán’s theorem, *J. Combin. Theory Ser. B* **96** (2006), 122–134.
- [16] D. Mubayi and Y. Zhao, Co-degree density of hypergraphs, submitted.
- [17] J. Pelikán, Properties of balanced incomplete block designs, *Combinatorial theory and its applications*, Balatonfüred, Colloq. Math. Soc. János Bolyai, **4** (1969), 869–889.
- [18] G. Tallini, On blocking sets in finite projective and affine spaces, Combinatorics '86 (Trento, 1986), 433–450, *Ann. Discrete Math.* **37**, North-Holland, Amsterdam, 1988.