Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Combinatorial Theory, Series A 111 (2005) 289–309
Journal of Combinatorial Theory Series A
www.elsevier.com/locate/jcta

The Turán problem for projective geometries

Peter Keevash

Department of Mathematics, Princeton University, Princeton, NJ 08544, USA

Received 23 February 2004
Available online 12 February 2005

Abstract

We consider the following Turán problem. How many edges can there be in a $(q+1)$-uniform hypergraph on n vertices that does not contain a copy of the projective geometry $PG_m(q)$? The case $q=m=2$ (the Fano plane) was recently solved independently and simultaneously by Keevash and Sudakov (The Turán number of the Fano plane, Combinatorica, to appear) and Füredi and Simonovits (Triple systems not containing a Fano configuration, Combin. Probab. Comput., to appear). Here we obtain estimates for general q and m via the de Caen–Füredi method of links combined with the orbit-stabiliser theorem from elementary group theory. In particular, we improve the known upper and lower bounds in the case $q=2$, $m=3$.

MSC: 05C35; 05C65

Keywords: Extremal problems; Hypergraphs; Projective geometries

1. Introduction

For an r-uniform hypergraph $\mathcal{F}$, the Turán number $\text{ex}(n,\mathcal{F})$ is the maximum number of edges in an r-uniform hypergraph on n vertices that does not contain a copy of $\mathcal{F}$. Determining these numbers is one of the central problems in extremal combinatorics. For ordinary graphs (the case $r=2$) this is completely solved for many instances, including all complete graphs. Turán proved that the unique largest graph on n vertices not containing a copy of K_t (the complete graph on t vertices) is the complete $(t-1)$-partite graph with

E-mail address: keevash@math.princeton.edu.

0097-3165/$ - see front matter
doi:10.1016/j.jcta.2004.12.007

part sizes as equal as possible. Moreover, asymptotic results are known for all non-bipartite graphs.

In contrast, for nearly any r-uniform hypergraph $\mathcal{F}$ with $r > 2$, the problem of finding the numbers $ex(n, \mathcal{F})$ is notoriously difficult. Even the asymptotics of hypergraph Turán numbers are poorly understood. It is not hard to show that the limit $\pi(\mathcal{F}) = \lim_{n\to\infty} ex(n, \mathcal{F})/\binom{n}{r}$ exists. It is usually called the *Turán density*. It is a famous open problem of Turán to determine the numbers $ex(n, K_s^{(r)})$, where $K_s^{(r)}$ denotes the complete r-uniform hypergraph on s vertices. In particular, he conjectured that $\pi(K_4^{(3)})$ is equal to 5/9, and Erdős offered a \$1000 prize for the solution of even this case. There are very few exact results on hypergraph Turán numbers. Most of these are described in the excellent survey of Füredi [4]. More recently, there have been three new exact results (see [6–8,11,12] for details). Most of the progress has been for triple systems (the case $r = 3$), where there have also been some new results on Turán densities (see [14]).

In this paper we will consider the Turán 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 $m = q = 2$) the exact Turán number was determined independently and simultaneously by Keevash and Sudakov [11] and Füredi and Simonovits [8]. They showed that $ex(n, PG_2(2)) = \binom{n}{3} - \binom{\lfloor n/2\rfloor}{3} - \binom{\lceil n/2\rceil}{3}$ for n sufficiently large. In particular $\pi(PG_2(2)) = 3/4$, which was proved earlier by de Caen and Füredi [2]. The case $q = 2$, $m = 3$ was considered by Cioabă [1], who obtained the bounds $27/32 \leqslant \pi(PG_3(2)) \leqslant 27/28$.

Our first result gives general bounds for $\pi(PG_m(q))$.

Theorem 1.1. *The Turán density of $PG_m(q)$ satisfies*

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

For the case $q = 2$, our next theorem improves the general upper bound of Theorem 1.1.

Theorem 1.2. *The Turán density of $PG_m(2)$ satisfies*

$$\frac{(2^{m+1}-3)(2^{m+1}-4)}{(2^{m+1}-2)^2} \leqslant \pi(PG_m(2)) \leqslant \begin{cases} 1 - \dfrac{3}{2^{2m}-1}, & m \text{ odd},\\ 1 - \dfrac{6}{(2^m-1)(2^{m+1}+1)}, & m \text{ even}.\end{cases}$$

Note that in the case $m = 3$ this improves the known upper bound to 20/21. Next we concentrate further on this case, where we are able to improve both the upper and lower bounds.

Theorem 1.3. *The Turán density of $PG_3(2)$ satisfies*

$$3\sqrt{3} + 2\sqrt{2(9-5\sqrt{3})} - 6 \leqslant \pi(PG_3(2)) \leqslant 13/14.$$

For comparison purposes, note that $3\sqrt{3}+2\sqrt{2(9-5\sqrt{3})}-6 \sim 0.844778$ and $27/32 = 0.84375$.

Notation. Our graphs and multigraphs are denoted by the letter G or J, possibly with subscripts. If G is a graph or multigraph, then $e(G)$ denotes the number of edges it contains, counted with multiplicity. If X is a subset of the vertex set then G_X denotes the restriction of G to X (i.e. the induced subgraph) and $e(X) = e(G_X)$ is the number of edges there. In particular, if u, v are vertices we write $e(uv)$ for the multiplicity of the pair u, v. We write $d(u) = \sum_v e(uv)$ for the degree of u and $d_X(u) = \sum_{v \in X} e(uv)$ for the degree of u in X. Our hypergraphs and multihypergraphs are denoted by calligraphic letters such as $\mathcal{H}$ and $\mathcal{G}$. (In an r-uniform multihypergraph each r-subset A of the vertex set has some non-negative integer multiplicity $e(A)$.) Suppose $\mathcal{H}$ is an r-uniform hypergraph and x is a vertex. The *link* of x is an $(r-1)$-uniform hypergraph $L(x)$ on the same vertex set as $\mathcal{H}$, where A is an edge of $L(x)$ exactly when $A \cup \{x\}$ is an edge of $\mathcal{H}$. If X is a subset of the vertex set then the *link multihypergraph* of X is the $(r-1)$-uniform multihypergraph $L(X) = \sum_{x \in X} L(x)$. In other words, each $(r-1)$-tuple A has multiplicity in $L(X)$ equal to the number of vertices $x \in X$ such that $A \cup \{x\}$ is an edge of $\mathcal{H}$.

The rest of this paper is organised as follows. In the next section we will define the projective geometries $PG_m(q)$ and prove Theorems 1.1 and 1.2. Section 3 contains the proof of Theorem 1.3. The upper bound requires a technical lemma, the proof of which we postpone to Section 4. The final section contains some concluding remarks and open problems.

2. General bounds

In this section we prove Theorems 1.1 and 1.2. We start by recalling some elementary algebra. Let $\mathbb{F}_q$ denote the field with q elements. The *projective geometry* of dimension m over $\mathbb{F}_q$, denoted $PG_m(q)$, is the following hypergraph. Its vertex set is the set of all one-dimensional subspaces of $\mathbb{F}_q^{m+1}$. Its edges correspond to 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 $PG_m(q)$.

We can identify a one-dimensional subspace by picking one of its non-zero vectors. There are $q-1$ choices of this representative, which are equivalent in the sense that they are scalar multiples of one another. A two-dimensional subspace contains $q+1$ one-dimensional subspaces, for which we can choose representatives of the form $\{x, y, x+y, 2x+y, \dots, (q-1)x+y\}$. To count the two-dimensional subspaces consider picking a non-zero vector and then another which is not equivalent. There are $(q^{m+1}-1)(q^{m+1}-q)$ such choices, and each two-dimensional subspace is generated by $(q^2-1)(q^2-q)$ choices, giving $\frac{(q^{m+1}-1)(q^{m+1}-q)}{(q^2-1)(q^2-q)}$ subspaces. Therefore $PG_m(q)$ is a $(q+1)$-uniform hypergraph with $\frac{q^{m+1}-1}{q-1} = \sum_{j=0}^{m} q^j$ vertices and $\frac{(q^{m+1}-1)(q^m-1)}{(q^2-1)(q-1)}$ edges.

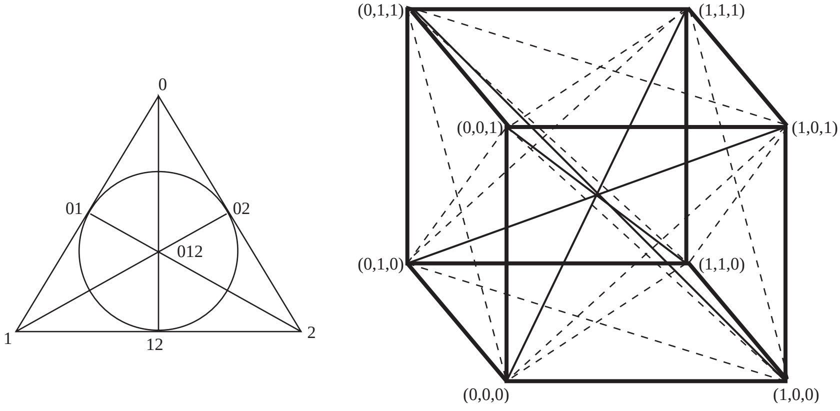

Fig. 1. The Fano plane in $PG_3(2)$.

Let $GL_m(q)$ denote the *general linear* group of invertible linear maps from $\mathbb{F}_q^m$ to itself. The set of non-zero multiples of the identity matrix forms a normal subgroup D in $GL_m(q)$; the quotient $GL_m(q)/D$ is the *projective linear group* $PGL_m(q)$. This is the automorphism group of $PG_{m-1}(q)$. Note that in the case $q = 2$ there is only one non-zero field element. Here we can identify $PG_{m-1}(2)$ with the non-zero elements of $\mathbb{F}_2^m$, and D consists only of the identity element. Then $PGL_m(2) = GL_m(2)$, and it is customary to denote both by $L_m(2)$. (It is also equal to the special linear group and projective special linear group in this case, but these will not be relevant to our discussion.)

From our discussion of the case of general q we know that $PG_m(2)$ is a 3-uniform hypergraph, with $2^{m+1} - 1$ vertices and $\frac{1}{3}(2^m - 1)(2^{m+1} - 1)$ edges. The automorphism group of $PG_m(2)$ is $L_{m+1}(2)$, the group of invertible linear maps from $\mathbb{F}_2^{m+1}$ to itself. For illustrative purposes, and because it will be important later, we will describe the specific example $PG_2(2)$. We will use the following concise notation for its elements. A non-zero element $x = (x_1, x_2, x_3)$ of $\mathbb{F}_2^3$ is described by a string containing some combination of the symbols 0, 1, 2, where symbol i appears in the string for x exactly when $x_{i+1} = 1$. With this notation, the vertices of $PG_2(2)$ are 0, 1, 2, 01, 02, 12, 012 and the lines are as shown on the left of Fig. 1 (the circle is also a line). This configuration is commonly known as the *Fano plane*. Its automorphism group $L_3(2)$ has 168 elements. It will be helpful later to note that $L_3(2)$ is doubly transitive, transitive on lines and transitive on non-collinear triples.

An obvious property of $PG_m(q)$ is that every pair of points x, y belongs to exactly one edge, for which we can choose representatives $\{x, y, x+y, 2x+y, \dots, (q-1)x+y\}$. Note that $PG_m(q)$ contains a copy of $PG_{m-1}(q)$, for example that consisting of the equivalence classes of all non-zero vectors $x = (x_1, \dots, x_{m+1})$ in $\mathbb{F}_q^{m+1}$ which have $x_{m+1} = 0$. Let X be the set of vertices of this copy of $PG_{m-1}(q)$ and let Y be the remaining vertices. Then Y consists of the equivalence classes of all vectors with $x_{m+1} \neq 0$; we can pick representatives so that $x_{m+1} = 1$, and then all possible q^m vectors appear in the first m co-ordinates. Any edge of $PG_m(q)$ that intersects Y has the form $\{x, y, x + y, 2x + y, \dots, (q - 1)x + y\}$, where $x \in X$ and $y \in Y$, so it intersects Y in exactly $q - 1$ points.

To continue our discussion we introduce a definition. Consider any r-uniform hypergraph $\mathcal{H}$ and a vertex x. The *link* of x is an $(r-1)$-uniform hypergraph $L(x)$ on the same vertex set as $\mathcal{H}$, where A is an edge of $L(x)$ exactly when $A \cup \{x\}$ is an edge of $\mathcal{H}$. Suppose that $\mathcal{H} = PG_m(q)$ and x is a vertex of X (with the notation of the previous paragraph). Then the link $L(x)$ restricted to Y is a perfect matching M_x of Y, i.e. a set of q^{m-1} mutually disjoint q-tuples. Indeed, for each $y \in Y$ the unique edge of $L(x)$ containing y is $\{y, x+y, 2x+y, \ldots, (q-1)x+y\}$. As x ranges over $PG_{m-1}(q)$ we obtain a set of $\frac{q^m-1}{q-1}$ perfect matchings of Y such that each pair of vertices of Y appear in exactly one edge. Fig. 1 illustrates this in the case $q = 2, m = 3$. Here X is the Fano plane and we can think of $Y \cong \mathbb{F}_2^3$ as the vertices of a cube. The edges of the cube (in bold) comprise the three matchings M_0, M_1, M_2, the long diagonals form the matching M_{012}, and the face diagonals (dotted lines) form the remaining three matchings M_{01}, M_{02}, M_{12}.

Our strategy for finding a copy of $PG_m(q)$ in a sufficiently dense $(q+1)$-uniform hypergraph is as follows. By induction we will be able to assume that there is a copy of $PG_{m-1}(q)$, which we denote X and label as before with the equivalence classes of all non-zero vectors $x \in \mathbb{F}_q^{m+1}$ with $x_{m+1} = 0$. We will find a particular set Y of q^m vertices, which can be labelled with the vectors $x \in \mathbb{F}_q^{m+1}$ with $x_{m+1} = 1$ and define the matchings $\{M_x : x \in X\}$ as above, i.e. M_x contains all q-tuples of the form $\{y, x+y, 2x+y, \ldots, (q-1)x+y\}$ for $y \in Y$. This labelling will have the property that there is an automorphism $\pi \in PGL_m(q)$ so that $M_x \subset L(\pi(x))$ for every $x \in PG_{m-1}(q)$. Clearly, this gives a copy of $PG_m(q)$, as we can relabel $\pi(x)$ as x without altering the hypergraph. In order to make this argument we need the links of the vertices in $PG_{m-1}(q)$ to be large (i.e. we need large degrees). This is achieved by the following fact.

Fact 2.1. *Let $\delta, \varepsilon > 0$ and let $k \geqslant 1$, $r \geqslant 2$ and n_0 be positive integers. Then there is n_1 so that for all $n \geqslant n_1$ any r-uniform multihypergraph $\mathcal{H}$ on n vertices with at least $(\delta + 2\varepsilon)\binom{n}{r}$ edges and maximum multiplicity at most k contains an r-uniform multihypergraph $\mathcal{H}'$ with m vertices and minimum degree at least $(\delta+\varepsilon)\binom{m-1}{r-1}$, for some $m \geqslant n_0$.*

This follows from a standard argument involving deleting vertices of small degree, but for the convenience of the reader we will give the following brief proof.

Proof. Suppose that $\mathcal{H}$ is a counterexample to the statement. Then we can construct a sequence $\mathcal{H}_n, \mathcal{H}_{n-1}, \ldots, \mathcal{H}_{n_0}$ where $\mathcal{H}_{m-1}$ is obtained from $\mathcal{H}_m$ by deleting a vertex of degree at most $(\delta+\varepsilon)\binom{m-1}{r-1}$. Then

$$k\binom{n_0}{r} \geqslant e(\mathcal{H}_{n_0}) \geqslant (\delta+2\varepsilon)\binom{n}{r} - \sum_{m=n_0+1}^{n} (\delta+\varepsilon)\binom{m-1}{r-1}$$
$$= (\delta+2\varepsilon)\binom{n_0}{r} + \varepsilon \sum_{m=n_0+1}^{n} \binom{m-1}{r-1}.$$

From the crude estimate $kn_0^r > \varepsilon n$ we obtain a contradiction with $n_1 = \varepsilon^{-1} k n_0^r$, so the Fact is true. □

We will need the following piece of elementary group theory. (See [10] for an introduction to this subject.) Suppose a group G acts on a set X, with the action written on the left. For an element x in X, the stabiliser of x is the subgroup G_x of G that fixes x. The orbit of x is $Gx = \{gx : g \in G\}$, i.e. the set of all images of x under the action of G. The orbit-stabiliser theorem states that $|G| = |G_x||Gx|$. In particular, if G acts transitively on X (i.e. the whole of X is a single orbit) then $|G_x| = |G|/|X|$. Let $G(x, y)$ denote the set of elements of G that take element x to element y. If g is any element taking x to y then $G(x, y)$ is the coset gG_x, so it also has size $|G|/|X|$. We will apply this to the action of $PGL_m(q)$ on $PG_{m-1}(q)$, which is obviously transitive, as one can map any line to any other line via an invertible linear map. In this case there are $|PGL_m(q)|/|PG_{m-1}(q)|$ elements mapping x to y, for any $x, y \in PG_{m-1}(q)$.

Now we give the proof of Theorem 1.1, which gives bounds on the Turán density of $PG_m(q)$ for general m and q.

Proof of Theorem 1.1. We start with the lower bound. Let $t = |PG_m(q)| - 1 = \sum_{j=1}^m q^j$ and let $\mathcal{H}(n)$ be the 'blow-up' of $K_t^{(q+1)}$, i.e. we divide a set of n vertices into parts $X_1, \dots, X_t$ with $||X_i| - |X_j|| \leqslant 1$ for all i, j and take as edges all $(q+1)$-tuples of the form $x_1 \cdots x_{q+1}$ with $x_i \in X_{a_i}$ for some pairwise distinct $a_1, \dots, a_{q+1}$. Recall that $PG_m(q)$ is a $(q+1)$-uniform hypergraph on $t+1$ points in which every pair of points belong to some edge. It is clear that $\mathcal{H}(n)$ does not contain a copy of $PG_m(q)$, as for any set of $t+1$ points some two will fall into the same X_i, so there will not be an edge through them. Since $K_t^{(q+1)}$ has $\binom{t}{q+1}$ edges, we get a lower bound on the density $\pi(PG_m(q))$ of $\lim_{n\to\infty} \binom{n}{q+1}^{-1} e(\mathcal{H}(n)) = (q+1)! \binom{t}{q+1} / t^{q+1} = \prod_{i=1}^q (1 - i/t)$.

Now we prove the upper bound by induction, the case $m = 1$ being trivial. Suppose $m \geqslant 2$ and define $\delta = 1 - 1/\binom{q^m}{q}$. Suppose $\varepsilon > 0$ and let $\mathcal{H}$ be a $(q+1)$-uniform hypergraph with minimum degree at least $(\delta + \varepsilon)\binom{n-1}{q}$. By Fact 2.1 it suffices to prove that $\mathcal{H}$ contains a copy of $PG_m(q)$ when n is sufficiently large.

Note that $\mathcal{H}$ contains a copy of $PG_{m-1}(q)$ by induction hypothesis. Let X be its set of vertices and let $Z = V(\mathcal{H}) \setminus X$. Let $\mathcal{G}$ be the *link multihypergraph* of X on Z, i.e. each q-tuple $A \subset Z$ appears in $\mathcal{G}$ with multiplicity equal to the number of vertices $x \in X$ such that $A \cup \{x\}$ is an edge of $\mathcal{H}$. Note that there are less than $|X|n^{q-1}$ q-tuples that intersect X, each of which can have multiplicity at most $|X|$ in $\mathcal{G}$. By the minimum degree assumption we have

$$e(\mathcal{G}) > \sum_{x\in X} |L(x)| - |X|^2 n^{q-1} > |PG_{m-1}(q)|(\delta + \varepsilon/2)\binom{n-1}{q},$$

for large n. By averaging there must be a subset $Y \subset Z$ with $|Y| = q^m$ so that

$$e(\mathcal{G}_Y) > |PG_{m-1}(q)|\delta\binom{q^m}{q} = |PG_{m-1}(q)|\left(\binom{q^m}{q} - 1\right).$$

Choose an arbitrary labelling of the vertices of Y with the vectors y in $\mathbb{F}_q^{m+1}$ with $y_{m+1} = 1$. Define the perfect matchings $\{M_x : x \in X\}$ as in the discussion before the theorem. As we remarked earlier, it suffices to show that there is an automorphism $\pi \in PGL_m(q)$ so that $M_x \subset L(\pi(x))$ for every $x \in PG_{m-1}(q)$. Consider any pair (A, x) such that $x \in PG_{m-1}(q)$, $A \in M_y$ for some $y \in PG_{m-1}(q)$ and $A \notin L(x)$. We cannot use any π that maps y to x; there are exactly $|PGL_m(q)|/|PG_{m-1}(q)|$ such π by the piece of group theory discussed earlier. Since $e(\mathcal{G}_Y) > |PG_{m-1}(q)| \left(\binom{q^m}{q} - 1\right)$ there are at most $|PG_{m-1}(q)| - 1$ such pairs (A, x) and so at most $\frac{|PG_{m-1}(q)|-1}{|PG_{m-1}(q)|}|PGL_m(q)|$ automorphisms that violate any condition $M_a \subset L(\pi(a))$. This leaves at least one automorphism that satisfies the required conditions, so we are done. □

Next we prove Theorem 1.2, which gives an improvement to the upper bound when $q = 2$. The proof method is the same as for Theorem 1.1, except that we also use the result of Füredi and Kündgen on the Turán problem for integer-weighted graphs, which was an important ingredient in [1,8,11,14]. An integer-weighted graph is a graph G together with an assignment of integral weights to its edges. The weight of the graph is the sum of the weights of its edges. Define $f_{\mathbb{Z}}(n, k, r)$ to be the maximum weight of an integer weighted graph on n vertices in which every subset of k vertices induces a subgraph of weight at most r. Let $f(k, r)$ denote the smallest number t so that $\sum_{i=1}^{k-1} \lfloor 1 + it \rfloor > r$. Füredi and Kündgen [5] showed that $f_{\mathbb{Z}}(n, k, r) = f(k, r)\binom{n}{2} + O(n)$. We will only use the upper bound, applied to multigraphs (which in particular are integer-weighted graphs). In the following lemma we calculate $f(k, r)$ in the case that we will use.

Lemma 2.2.

$$f\left(2^m, (2^m - 1)\left(\binom{2^m}{2} - 1\right)\right) = \begin{cases} 2^m - 1 - \frac{3}{2^m+1}, & m \geqslant 3 \text{ odd}, \\ 2^m - 1 - \frac{6}{2^{m+1}+1}, & m \geqslant 4 \text{ even}. \end{cases}$$

The proof of this lemma is an easy but slightly tedious calculation, which we give in Appendix A.

Proof of Theorem 1.2. The lower bound is given by Theorem 1.1. We prove the upper bound by induction for $m \geqslant 2$. For $m = 2$, i.e. the Fano plane, we have $\pi(PG_2(2)) = 3/4$ (by de Caen and Füredi [2]) which is less than the bound of 7/9 claimed by our theorem. Now suppose $m \geqslant 3$, and define δ_m to be equal to $1 - \frac{3}{2^{2m}-1}$ if m is odd and $1 - \frac{6}{(2^m-1)(2^{m+1}+1)}$ if m is even. Note that from Lemma 2.2 we have

$$(2^m - 1)\delta_m = f\left(2^m, (2^m - 1)\left(\binom{2^m}{2} - 1\right)\right). \tag{1}$$

Suppose $\varepsilon > 0$ and let $\mathcal{H}$ be a 3-uniform hypergraph with minimum degree at least $(\delta_m + \varepsilon)\binom{n-1}{2}$. By Fact 2.1 it suffices to prove that $\mathcal{H}$ contains a copy of $PG_m(2)$ when n is sufficiently large. It is straightforward to verify the inequality

$$1 - \frac{3}{2^{2(m-1)} - 1} \leqslant 1 - \frac{6}{(2^m - 1)(2^{m+1} + 1)} \leqslant 1 - \frac{3}{2^{2(m+1)} - 1},$$

i.e. that δ_m is an increasing sequence. Therefore $\mathcal{H}$ contains a copy of $PG_{m-1}(2)$ by the induction hypothesis.

Let X be the vertex set of this $PG_{m-1}(2)$ and let $Z = V(\mathcal{H})\backslash X$. Let G be the link multigraph of X on Z (defined as before). Note that there are less than $|X|n$ pairs of vertices that intersect X, each having multiplicity at most $|X|$ in G. By the minimum degree assumption we have

$$e(G) > \sum_{x\in X} |L(x)| - |X|^2 n > (2^m - 1)(\delta_m + \varepsilon/2)\binom{n-1}{2}.$$

From Eq. (1) and the forementioned result of Füredi and Kündgen it follows that there is some subset $Y \subset Z$ with $|Y| = 2^m$ so that $e(G_Y) > (2^m - 1)\left(\binom{2^m}{2} - 1\right)$. Choose an arbitrary labelling of the vertices of Y with the elements of $\mathbb{F}_2^m$. For $x \neq 0$ let M_x be the perfect matching of Y in which any element y is matched with $y + x$. As in the proof of Theorem 1.1 we can find an automorphism $\pi \in L_m(2)$ so that $M_x \subset L(\pi(x))$ for every $x \in X$. Thus $X \cup Y$ spans a copy of $PG_m(2)$, so we are done. □

3. Dimension three

In this section we give the proof of Theorem 1.3. First we give the lower bound.

Theorem 3.1. *The Turán density of $PG_3(2)$ is at least* $3\sqrt{3} + 2\sqrt{2(9-5\sqrt{3})} - 6 \sim 0.844778$.

Proof. Consider the following construction. Divide a set of n vertices into parts X_1, X_2, Y, Z so that $\big||X_1| - \alpha n\big| \leqslant 1$, $\big||X_2| - \alpha n\big| \leqslant 1$, $\big||Y| - \beta n\big| \leqslant 1$ and $\big||Z| - \gamma n\big| \leqslant 1$, where α, β, γ are constants that we will specify later with $2\alpha + \beta + \gamma = 1$. Let $X = X_1 \cup X_2$. Define a 3-uniform hypergraph $\mathcal{H}_n$, whose edges are all triples that do not lie entirely within one of the sets X, Y, Z and are not of the form abc with $a \in Y$ and $b, c \in X_i$ for some i.

To see that this does not contain a copy of $PG_3(2)$ we will use a result of Pelikán [15] that in any weak 3-colouring of $PG_3(2)$ there must be (exactly) 5 points of each colour. (A weak colouring of a hypergraph is a colouring of the vertices in which there is no edge where all vertices have the same colour.) Indeed, suppose that there is a copy of $PG_3(2)$ in $\mathcal{H}_n$. Then (X, Y, Z) defines a proper 3-colouring of $PG_3(2)$, which therefore has 5 points in each part. Without loss of generality it has at least 3 points in X_1, call them x, y, z. There are no edges of $\mathcal{H}_n$ entirely within X_1, so x, y, z cannot be collinear. Consider the lines $(x, y, x+y)$, $(x, z, x+z)$, $(y, z, y+z)$. These are edges of $PG_3(2)$, and there are no edges in $\mathcal{H}_n$ of the form abc with $a \in Y$ and $b, c \in X_1$, so we must have $x+y, x+z, y+z \in Z$. However $(x+y)+(x+z) = y+z$, so there are three collinear points in Z. There are no edges of $\mathcal{H}_n$ entirely within Z, so we have a contradiction.

The number of edges in $\mathcal{H}_n$ is

$$e(\mathcal{H}_n) = \binom{n}{3} - \binom{|X|}{3} - \binom{|Y|}{3} - \binom{|Z|}{3} - |Y|\binom{|X_1|}{2} - |Y|\binom{|X_2|}{2}.$$

Therefore we have a lower bound for the density $\pi(PG_3(2))$ of

$$\lim_{n\to\infty} \binom{n}{3}^{-1} e(\mathcal{H}_n) = 1 - 8\alpha^3 - \beta^3 - \gamma^3 - 6\beta\alpha^2.$$

This lower bound is optimised by the following choice of parameters:

$$\alpha = \tfrac{1}{2}(\sqrt{3} - 1 - \sqrt{6 - 10/\sqrt{3}}) \sim 0.128067,$$
$$\beta = 1 - \sqrt{1 - 1/\sqrt{3}} \sim 0.349885,$$
$$\gamma = 1 + \sqrt{3 - \sqrt{3}} - \sqrt{3} \sim 0.393982.$$

This gives the lower bound

$$\pi(PG_3(2)) \geqslant 3\sqrt{3} + 2\sqrt{2(9 - 5\sqrt{3})} - 6 \sim 0.844778,$$

as required. □

Next, we want to improve the upper bound on $\pi(PG_3(2))$ which comes from Theorem 1.2. Our broad strategy is the same. We find a copy X of the Fano plane $PG_2(2)$, and a set Y of 8 vertices labelled with the elements of $\mathbb{F}_2^3$, so that defining the matchings $\{M_x : x \in X\}$ as before there is an automorphism $\pi \in L_3(2)$ so that $M_x \subset L(\pi(x))$. The improvement comes from a closer analysis of the conditions under which we can find a set Y that has such a labelling. This is achieved by the following technical lemma, whose proof we postpone to the next section.

Lemma 3.2. *Suppose* $\varepsilon > 0$. *Let G be a multigraph on n vertices with maximum multiplicity* 7 *and with at least* $(6.5 + \varepsilon)\binom{n}{2}$ *edges. Suppose also that for each pair of vertices* x, y *we have a set* $T_{x,y} \subset PG_2(2)$ *and that* $|T_{x,y}|$ *is the multiplicity of the pair* x, y *in G. Then we can find* 8 *vertices of G and label them with the elements of* $\mathbb{F}_2^3$ *so for every* $a \in PG_2(2)$ *and* $x \in \mathbb{F}_2^3$ *we have* $a \in T_{x,x+a}$.

Now we can prove the upper bound in Theorem 1.3, which we restate as the following theorem.

Theorem 3.3. *The Turán density of* $PG_3(2)$ *is at most* $13/14$.

Proof. Suppose $\varepsilon > 0$ and let $\mathcal{H}$ be a 3-uniform hypergraph with minimum degree at least $(13/14 + \varepsilon)\binom{n}{2}$. By Fact 2.1 it suffices to prove that $\mathcal{H}$ contains a copy of $PG_3(2)$ when n is sufficiently large. Note that $\mathcal{H}$ contains a copy of the Fano plane $PG_2(2)$, since this has Turán density 3/4 (see [2]). Let X be its set of vertices and let $Z = V(\mathcal{H})\backslash X$. For $y, z \in Z$ let $T_{y,z}$ be the set of vertices x in X such that xyz is an edge of $\mathcal{H}$. Let $\mathcal{G}$ be the link multigraph of X on Z, i.e. the multiplicity of yz is $|T_{y,z}|$. By the minimum degree assumption we have $e(\mathcal{G}) > \sum_{x\in X} |L(x)| - |X|^2 n > (6.5 + \varepsilon)\binom{n}{2}$. By Lemma 3.2 we can find 8 vertices $Y \subset Z$ and label them with the elements of $\mathbb{F}_2^3$ so for every $x \in X$ and for every $y \in Y$ the pair $(y, y + x)$ belongs to $L(x)$. Then $X \cup Y$ spans a copy of $PG_3(2)$ in $\mathcal{H}$, so we are done. □

4. A technical lemma

This section is devoted to the proof of Lemma 3.2. A key ingredient is the following lemma on automorphisms of the Fano plane. Recall that the elements of the Fano plane $PG_2(2)$ can be identified with non-zero elements of $\mathbb{F}_2^3$, which we denote by 0, 1, 2, 01, 02, 12, 012, where e.g. 02 denotes the vector $(1, 0, 1)$.

Lemma 4.1. *Suppose that for each $x \in PG_2(2)$ we have a subset $S_x \subset PG_2(2)$ with the following properties*:

(i) *no S_x is equal to the whole of $PG_2(2)$,*
(ii) $\sum_{x \in PG_2(2)} |S_x| \leqslant 9$,
(iii) *at least one S_x is empty,*
(iv) *if exactly one S_x is empty then we have $S_0 = S_1 = S_{01} = \{0, 1\}$ and $S_{02} = S_{12} = S_{012} = \{01\}$,*
(v) *there are not two lines L_1, L_2 of $PG_2(2)$ such that $S_x = L_2$ for each $x \in L_1$. Then there is an automorphism $\pi \in L_3(2)$ such that $\pi(x) \notin S_x$ for all x.*

The proof of this lemma is long and uninstructive, so we relegate it to Appendix B.

Proof of Lemma 3.2. Suppose $\varepsilon > 0$. Let G be a multigraph on n vertices with maximum multiplicity 7 and minimum degree at least $(6.5 + \varepsilon)(n - 1)$. Suppose also that for each pair of vertices x, y we have a set $T_{x,y} \subset PG_2(2)$ and that $|T_{x,y}|$ is the multiplicity of the pair x, y in G. By Fact 2.1 it is enough to show that we can find 8 vertices of G and label them with the elements of $\mathbb{F}_2^3$ so for every $a \in PG_2(2)$ and $x \in \mathbb{F}_2^3$ we have $a \in T_{x,x+a}$.

For every $a \in PG_2(2)$ let M_a denote the matching of $\mathbb{F}_2^3$ in which x is paired with $x + a$. Our strategy will be to find 8 vertices and a labelling by $\mathbb{F}_2^3$ such that the sets $S_a = PG_2(2) \setminus \bigcap_{e \in M_a} T_e$ for $a \in PG_2(2)$ satisfy the conditions of Lemma 4.1. It will then follow that there is an automorphism π of $PG_2(2)$ such that $\pi(a) \notin S_a$, i.e. $\pi(a) \in \bigcap_{e \in M_a} T_e$, for all a. Now we relabel our set of 8 vertices so that the vertex with label a has now has label $\pi(a)$. This has the effect of relabelling M_a as $M_{\pi(a)}$, so now we have $a \in \bigcap_{e \in M_a} T_e$, as required.

We choose the 8 vertices as follows. Let $v_1 v_2$ be any edge of multiplicity 2. We claim that we can choose $v_3, \ldots, v_8$ so that $\sum_{i=1}^{j-1} e(v_i v_j) > 6.5(j - 1)$ for $3 \leqslant j \leqslant 8$. (Recall that in a multigraph $e(xy)$ denotes the multiplicity of the edge xy.) For suppose we cannot choose some v_j in this manner. Let $U = \{v_1, \ldots, v_{j-1}\}$. Then for every $v \notin U$ we have $d_U(v) = \sum_{i=1}^{j-1} e(v_i u) \leqslant 6.5(j - 1)$. Therefore

$$(j - 1)(6.5 + \varepsilon)(n - 1) \leqslant \sum_{u \in U} d(u) = 2e(U) + \sum_{v \notin U} d_U(v) \leqslant 14 \binom{j-1}{2} + 6.5(j - 1)(n - j + 1).$$

This implies that $2\varepsilon(n-1) \leqslant j-2$, which is a contradiction for n sufficiently large. Therefore we can choose $v_1, \ldots, v_8$ as described.

Next, we attempt to improve our choice in the following manner. Suppose there is a subset $I \subset \{1, \ldots, 8\}$ and a bijection $f : \{1, \ldots, |I|\} \to I$ such that $e(v_i v_j) \leqslant e(v_{f(i)} v_{f(j)})$ for every $1 \leqslant i < j \leqslant |I|$ and at least one inequality is strict. Then we will choose $v'_i = v_{f(i)}$ for $1 \leqslant i \leqslant |I|$ and then repeat the above argument to complete the set, choosing $v'_{|I|+1}, \ldots, v'_8$ so that $\sum_{i=1}^{j-1} e(v'_i v'_j) > 6.5(j-1)$ for $|I|+1 \leqslant j \leqslant 8$. Now we iterate this improvement procedure. Clearly there can only be finitely many iterations. [To see this formally, define a total order on 8-tuples of vertices, where $(v_1, \ldots, v_8) < (v'_1, \ldots, v'_8)$ when there are $a < b$ such that $e(v_a v_b) < e(v'_a v'_b)$ and $e(v_i v_j) = e(v'_i v'_j)$ for all (i, j) colexicographically less than (a, b), i.e. $i, j \leqslant b - 1$ or $j = b$ and $i < a$. Each iteration increases the rank of our 8-tuple in this order, and there are only finitely many possible choices for all the multiplicities $e(v_i v_j)$, so the process terminates.] Therefore we may assume that the choice $v_1, \ldots, v_8$ cannot be improved.

Write $m_e = |PG_2(2) \setminus T_e|$, i.e. $7 - m_e$ is the multiplicity of the edge e. Then we must have $\sum_{i=1}^{j-1} m_{v_i v_j} \leqslant \lfloor j/2 \rfloor - 1$ for each $j \geqslant 3$. Otherwise we would have

$$\sum_{i=1}^{j-1} e(v_i v_j) \leqslant 7(j-1) - \lfloor j/2 \rfloor \leqslant 6.5(j-1),$$

which contradicts our choice of v_j. Let J be the graph on $v_1, \ldots, v_8$ in which v_i is adjacent to v_j exactly when $e_G(v_i v_j) < 7$. We use the notation $d^*(v_j)$ to denote the number of vertices v_i adjacent in J to v_j with $i < j$. Note that

$$d^*(v_j) \leqslant \sum_{i=1}^{j-1} m_{v_i v_j} \leqslant \lfloor j/2 \rfloor - 1, \tag{2}$$

for each $j \geqslant 3$. In particular

$$e(J) = \sum_{j=3}^{8} d^*(v_j) \leqslant \sum_{j=3}^{8} (\lfloor j/2 \rfloor - 1) = 0 + 1 + 1 + 2 + 2 + 3 = 9.$$

Now we need the following claim.

Claim 4.2. *There are two disjoint four-cycles in the complement of J.*

Proof. Suppose for the sake of contradiction that the claim is not true. We will repeatedly use the observation that a graph on 4 vertices has a C_4 in its complement exactly when it has maximum degree at most 1. For $1 \leqslant i \leqslant j \leqslant 8$ we will use the notation $V_{ij} = \{v_k : i \leqslant k \leqslant j\}$.

Consider first the case when $V_{15} = \{v_1, \ldots, v_5\}$ is an independent set in J. Then $e(J) = d^*(6) + d^*(7) + d^*(8) \leqslant 2 + 2 + 3 = 7$, by Eq. (2). For each $v \in V_{15}$ the restriction of J to $v v_6 v_7 v_8$ has a vertex of degree 2. Otherwise there would be a C_4 in its complement, and since $V_{15} \setminus \{v\}$ is independent it also has a C_4 in its complement, so we have a contradiction. Now $V_{68} = \{v_6, v_7, v_8\}$ cannot be an independent set, as then each $v \in V_{15}$ would have two neighbours in V_{68}, giving $e(J) \geqslant 10$, a contradiction. In fact V_{68} must contain at least two edges. For if xy is the only edge in V_{68} then each $v \in V_{15}$ is adjacent to at least one of x, y. Including the edge xy we get at least 6 edges within $V_{15} \cup \{x, y\}$. However $d^*(x) + d^*(y) \leqslant 3 + 2 \leqslant 5$, so this is a contradiction.

Since V_{68} contains at least 2 edges there is an edge in V_{68} incident to v_8, so each $x \in V_{68}$ has at most 2 neighbours in V_{15}. If V_{68} contains exactly 2 edges then there are at most 5 edges between V_{15} and V_{68}; two vertices of V_{68} each being adjacent to at most two vertices of V_{15} and the other being adjacent to at most one vertex of V_{15}. On the other hand if V_{68} is complete then there are at most 4 edges between V_{15} and V_{68}; where v_8 and v_7 each have at most one neighbour in V_{15} and v_6 has at most two neighbours in V_{15}.

For each $x \in V_{68}$ and each pair $u, v \in V_{15}$ we consider the partition into two 4-tuples $A_{xuv} = V_{15} \cup \{x\} \backslash \{u, v\}$ and $B_{xuv} = V_{68} \cup \{u, v\} \backslash \{x\}$. Observe that for each $x \in V_{68}$ the complement of J contains a C_4 on A_{xuv} unless x has exactly 2 neighbours in V_{15} and neither is u or v. Also, if the complement of J does contain a C_4 on A_{xuv} then by assumption it cannot contain a C_4 on B_{xuv}. Write $\{y, z\} = V_{68} \backslash \{x\}$. It follows that there is at least one edge between $\{u, v\}$ and $\{y, z\}$. Also, if yz is not an edge there are at least two such edges, as u, v is not an edge.

Choose $x \in V_{68}$ to have at most one neighbour in V_{15}. Then for any pair $\{u, v\}$ in V_{15} the complement of J contains a C_4 on A_{xuv}, so by the above there is at least one edge between $\{u, v\}$ and $\{y, z\} = V_{68} \backslash \{x\}$. In particular, there is at most one vertex in V_{15} with no neighbour in $\{y, z\}$, so at least 4 edges between V_{15} and $\{y, z\}$. Therefore each of y and z has two neighbours in V_{15}. This is only possible when V_{68} contains exactly two edges.

Now choose $x \in V_{68}$ so that the other vertices y, z are not adjacent. Suppose first that x has at most one neighbour in V_{15}. Then there are at least two edges between $\{u, v\}$ and $\{y, z\}$ for any pair $\{u, v\}$ in V_{15}. If any $u \in V_{15}$ is not adjacent to y or z then the other four vertices of V_{15} must all be adjacent to both y and z, which is impossible. Therefore every $u \in V_{15}$ is adjacent to y or z. However this forces one of y or z to be adjacent to 3 vertices of V_{15}, which is impossible.

It follows that x has 2 neighbours in V_{15}, call them a and b. Since $e(J) \leqslant 7$ there are at most 3 edges between $\{y, z\}$ and V_{15}. Therefore we can assume a is not adjacent to both y and z. For any other $v \in V_{15}$ the complement of J contains a C_4 on A_{xav}, and so there are at least two edges between $\{a, v\}$ and $\{y, z\}$, i.e. v is adjacent to y or z. This gives at least 4 edges between $\{y, z\}$ and V_{15}, which is a contradiction. This completes the analysis of the case when $\{v_1, \dots, v_5\}$ is an independent set in J.

Next, we will consider the case when $V_{14} = \{v_1, \dots, v_4\}$ is an independent set in J. Then $e(J) = d^*(5) + d^*(6) + d^*(7) + d^*(8) \leqslant 1 + 2 + 2 + 3 = 8$, by Eq. (2). We can assume that there is no independent set of size 5 in J. For by definition of the improvement procedure of 8-tuples it would follow that V_{15} is independent, and we have already dealt with this case. It follows that each $v \in V_{58}$ has at least one neighbour in V_{14}, so there are at most 4 edges inside V_{58}. Also, for any non-adjacent pair $u, v \in V_{58}$ there is a pair $a, b \in V_{14}$ such that u is adjacent to a and v is adjacent to b. For if only one $a \in V_{14}$ was adjacent to u or v then $V_{14} \cup \{u, v\} \backslash \{a\}$ would be independent.

For brevity we let J_{58} denote the restriction of J to V_{58}. Note that J_{58} contains a vertex of degree at least 2. Otherwise its complement has a C_4, and since V_{14} is independent it has a C_4 in the complement, giving a contradiction. Also, we claim that J_{58} cannot have an independent set of size 3. For suppose $V_{58} \backslash \{v\}$ is independent for some $v \in V_{58}$. Then v must be the vertex of degree at least 2 in J_{58}. Note that for every $u \in V_{14}$ either u has 2 neighbours in $V_{58} \backslash \{v\}$ or v has 2 neighbours in $V_{14} \backslash \{u\}$. There is at least one $u \in V_{14}$ that does not have 2 neighbours in $V_{58} \backslash \{v\}$, otherwise we would have at least 8 edges between

V_{14} and V_{58}, and at least 2 edges inside V_{58}, giving $e(J) \geqslant 10$, a contradiction. Therefore v has at least 2 neighbours in V_{14}; without loss of generality they are v_3 and v_4. Now $d^*(v) \geqslant 2$ and v has at least 2 neighbours in V_{58}, so we must have $v = v_6$, with v_6 adjacent to v_7 and v_8. Since v_6 does not have 2 neighbours in $V_{14} \backslash \{v_3\}$ we know that v_3 has at least 2 neighbours in $V_{58} \backslash \{v_6\}$. Similarly v_4 has at least 2 neighbours in $V_{58} \backslash \{v_6\}$. Therefore v_8 is adjacent to both v_3 and v_4, and without loss of generality v_5 is adjacent to v_3 and v_7 is adjacent to v_4. We have now listed 8 edges of J, so there are no more. Observe now that $v_3 v_7 v_5 v_4$ and $v_1 v_6 v_2 v_8$ give two disjoint four-cycles in the complement of J. This contradiction shows that J_{58} cannot have an independent set of size 3.

Now we know that J_{58} contains a vertex of degree at least 2 and no independent set of size 3. This easily implies that it must contain a triangle or a path of length three. In particular J_{58} has at least 3 edges, so there are at most 5 edges between V_{58} and V_{14}.

Suppose there is $v \in V_{58}$ so that $V_{58} \backslash \{v\}$ is a triangle in J. For any pair $x, y \in V_{58} \backslash \{v\}$ we claim that there are at most 3 edges between $\{x, y\}$ and V_{14}. If neither of $\{x, y\}$ is v_8 this holds because there are at most $d^*(x) + d^*(y) \leqslant 4$ edges within $V_{14} \cup \{x, y\}$ and one of them joins x to y. On the other hand, if say $x = v_8$ then since it has two neighbours in V_{58} it has at most one neighbour in V_{14}. Also $d^*(y) \leqslant 2$, so again there are at most 3 edges between $\{x, y\}$ and V_{14}. Recall that $e(J_{58}) \leqslant 4$, so we can choose $x, y \in V_{58} \backslash \{v\}$ not adjacent to v. Let z be the fourth element of V_{58}. For each $a \in V_{14}$ consider the 4-tuples $axyv$ and $V_{14} \cup \{z\} \backslash \{a\}$. One of them contains a vertex of degree at least 2, so either a is adjacent to one of x, y or z has at least 2 neighbours in $V_{14} \backslash \{a\}$. Since there are at most 3 edges between x, y and V_{14} we see that z has at least 2 neighbours in V_{14}. Since z also has 2 neighbours in V_{58} we must have $z = v_6$ and x, y equal to v_7, v_8 (in some order), so $v = v_5$. In particular, we see that v_5 and v_6 are not adjacent, i.e. the only edges of J_{58} form a triangle on V_{68}. Now repeat the argument choosing $x = v_6$ and $y = v_7$ (say). We conclude that $z = v_8$ has at least 2 neighbours in V_{14}, which is impossible. It follows that J_{58} does not contain a triangle.

Since $e(J_{58}) \leqslant 4$ we know that J_{58} is either a path of length 3 or a 4-cycle. In particular, we can label the vertices of V_{58} as $x_{00}, x_{01}, x_{10}, x_{11}$ so that $x_{i0} x_{i1}$ is not an edge of J for $i = 0, 1$. Consider any labelling of V_{14} as $y_{00}, y_{01}, y_{10}, y_{11}$. Note that $x_{i0} x_{i1} y_{j0} y_{j1}$ spans a bipartite subgraph of J for each of the 4 choices of $i, j \in \{0, 1\}$. These fall into two complementary pairs, and in each pair at least one of the subgraphs has a vertex of degree 2.

Recall that there at most 5 edges between V_{14} and V_{58}. Suppose that there are three vertices (say v_1, v_2, v_3) of V_{14} with at most one neighbour in V_{58}. They do not have a common neighbour in V_{58}, as this would have to be v_8, but v_8 has at least one neighbour in V_{58} already. Now not every pair among v_1, v_2, v_3 can have a common neighbour in V_{58}, as this would give at least 6 edges between V_{14} and V_{58}. Therefore we can assume that v_1 and v_2 do not have a common neighbour in V_{58}. Then the 4-tuple $x_{i0} x_{i1} v_1 v_2$ does not have a vertex of degree 2 for $i = 0, 1$ so $x_{i0} x_{i1} v_3 v_4$ must have a vertex of degree 2 for $i = 0, 1$. This gives at least 4 edges incident with $\{v_3, v_4\}$. Then we may suppose that v_1 has no neighbours in V_{58}. Note that v_4 cannot be adjacent to all of V_{58}. For recall we noted at the beginning of this case that for any non-adjacent pair $u, v \in V_{58}$ there is a pair $a, b \in V_{14}$ such that u is adjacent to a and v is adjacent to b. This implies that one of x_{i0}, x_{i1} has a neighbour in $V_{14} \backslash \{v_4\}$ for $i = 0, 1$, giving at least 6 edges between V_{14} and

V_{58}, a contradiction. Therefore we may suppose that v_4 is not adjacent to x_{00}, say. Since v_1 has no neighbours in V_{58} there is at most one edge in $x_{10}x_{11}v_1v_2$, so there is a degree 2 vertex in $x_{00}x_{01}v_3v_4$. Since v_4 is not adjacent to x_{00} we see that v_3 is adjacent to x_{01}. Now considering $x_{10}x_{11}v_3v_4$ shows that v_4 is adjacent to x_{10} and x_{11}. However, x_{00} has a neighbour in V_{14} and one of x_{10}, x_{11} has a neighbour in $V_{14}\backslash\{v_4\}$, giving at least 6 edges between V_{14} and V_{58}, a contradiction.

It follows that there cannot be 3 vertices of V_{14} with at most one neighbour in V_{58}. Then two vertices, say v_3, v_4, have at least two neighbours in V_{58}, and we may assume that v_4 has at least as many as v_3. There is at most one edge between $\{v_1, v_2\}$ and V_{58} so we can suppose that v_1 is isolated and v_2 has degree at most 1. We can suppose that v_4 is not adjacent to x_{00}. Considering $v_3v_4x_{00}x_{01}$ we see that v_3 is adjacent to x_{01}. Also $v_1v_4x_{00}x_{01}$ has at most one edge, so $v_2v_3x_{10}x_{11}$ has a vertex of degree 2. This vertex cannot be v_2 or v_3 (that would give 3 neighbours to v_3, and v_4 has at least as many). Then we can suppose that x_{10} is adjacent to v_2 and v_3. Now v_2 is adjacent to x_{10}, v_3 to x_{01}, x_{10}, and v_4 has degree 2. Then $v_2v_3x_{00}x_{01}$ has one edge, so $v_1v_4x_{10}x_{11}$ has a vertex of degree 2, which must be v_4. Now $v_1v_3x_{10}x_{11}$ and $v_2v_4x_{00}x_{01}$ each have at most one edge, which is a contradiction. This completes the analysis of the case when $\{v_1, \ldots, v_4\}$ is an independent set in J.

Now we can assume that J does not have an independent set of size 4, by definition of our improvement procedure for $v_1, \ldots, v_8$. Then V_{17} is not bipartite, as then it would have an independent set of size 4. It cannot be connected, as it has at most 6 edges, and a tree is bipartite.

Consider the case when the components are a 5-cycle and an isolated edge. Label the 5-cycle $x_1 \cdots x_5$ consecutively and the edge y_1y_2. Note that $x_2x_5x_3y_1$ is a 4-cycle in the complement of J, so v_8 has at least 2 neighbours in $\{x_1, x_4, y_2\}$. Similarly, for any non-adjacent a, b in the 5-cycle and y_i in the isolated edge we see that v_8 has at least 2 neighbours in $\{a, b, y_i\}$. It follows that v_8 is adjacent to y_1 and y_2, or it would be joined to the entire 5-cycle, which is impossible as it has at most 3 neighbours. Then v_8 has at most one neighbour on the 5-cycle, so we can choose a, b in the 5-cycle so that $\{a, b, v_8\}$ is independent. Then v_8 only has one neighbour in $\{a, b, y_1\}$, a contradiction.

Now every component of V_{17} must have size at most 4, so we can partition $V_{17} = L \cup R$ so that $|L| = 3$, $|R| = 4$ and there are no edges between L and R. Then $acbd$ is a 4-cycle in the complement of J for each pair $a, b \in L$ and pair $c, d \in R$. Therefore $abcv_8$ has a vertex of degree 2 for every $a \in L$ and $b, c \in R$. In particular v_8 is adjacent to one of b, c for each $b, c \in R$, so is adjacent to at least 3 vertices of R. Therefore v_8 is adjacent to exactly 3 vertices of R and to no vertices of L. Say bv_8 is not an edge, $b \in R$. If there is c such that bc is not an edge then $abcv_8$ does not have a vertex of degree 2 for any $a \in L$, contradiction. Therefore b has degree 3. Then $V_{17}\backslash\{b\}$ is a graph on 6 vertices with at most 3 edges, and it does not consist of 3 disjoint edges. Such a graph has an independent set of size 4. (This follows from Turán's theorem, but is also easy to see directly.) This contradiction completes the proof of Claim 4.2. □

Now we return to the proof of Lemma 3.2. Recall that it suffices to label $v_1, \ldots, v_8$ by $\mathbb{F}_2^3$ so that the sets $S_a = PG_2(2)\backslash \bigcap_{e\in M_a} T_e$ for $a \in PG_2(2)$ satisfy the conditions of Lemma 4.1. Recall also that $m_e = |PG_2(2)\backslash T_e|$. We chose $v_1, \ldots, v_8$ so that $\sum_{i<j} m_{v_iv_j} \leqslant 9$, so any labelling will satisfy condition (ii). By Claim 4.2 we know that the complement of J

has two disjoint four-cycles. Consider a labelling of $v_1, \dots, v_8$ by $\mathbb{F}_2^3$ so that one 4-cycle is labelled in cyclic order $(0, 0, 0)$, $(1, 0, 0)$, $(1, 1, 0)$, $(0, 1, 0)$ and the other is labelled $(0, 0, 1)$, $(1, 0, 1)$, $(1, 1, 1)$, $(0, 1, 1)$. Then the matchings $M_{(1,0,0)}$ and $M_{(0,1,0)}$ belong to the complement of J, so by definition $S_{(1,0,0)} = S_{(0,1,0)} = \emptyset$. Therefore this labelling satisfies conditions (iii) and (iv).

Suppose that the labelling does not satisfy condition (i). Then there is $a \in PG_2(2)$ such that $S_a = PG_2(2)$, i.e. $\bigcap_{e \in M_a} T_e = \emptyset$. Then $\sum_{e \in M_a} m_e \geqslant 7$ so there are at most 2 other edges e with $m_e > 0$; call them e_1, e_2. We chose $v_1, \dots, v_8$ so that $\sum_{i=1}^{j-1} m_{v_i v_j} \leqslant \lfloor j/2 \rfloor - 1$. In particular, $m_e \leqslant 3$ for all e, at most one e has $m_e = 3$ and at most three e have $m_e \geqslant 2$. It follows that we can choose x, y so that $\{a, x, y\}$ is a basis of $\mathbb{F}_2^3$, $m_{(0,a)} + m_{(x,x+a)} \leqslant 4$ and $m_{(y,y+a)} + m_{(x+y,x+y+a)} \leqslant 4$. Now we relabel our vertex set by interchanging the labels 0 and $x + a$. Note that the edges labelled $\{0, a\}$, $\{x, x + a\}$ are now labelled $\{0, x\}$, $\{a, a + x\}$. Even if e_1, e_2 are now both in M_a we still have $\sum_{e \in M_a} m_e \leqslant 6$. For the same reason $\sum_{e \in M_x} m_e \leqslant 6$. Also at least three of the S_i are empty. Therefore conditions (i)–(iv) are satisfied by this labelling.

Finally, suppose that the labelling does not satisfy condition (v), i.e. there are lines L_1, L_2 so that $S_x = L_2$ for each $x \in L_1$. Suppose that $L_i = \{x_i, y_i, x_i + y_i\}$. Let $A = \{0, x_1, y_1, x_1 + y_1\}$ and $B = \mathbb{F}_2^3 \backslash A$. Since $\sum_e m_e \leqslant 9$ we must have $\sum_e m_e = 9$, $\sum_{i=1}^{j-1} m_{v_i v_j} = \lfloor j/2 \rfloor - 1$ for $3 \leqslant j \leqslant 8$, and all edges e with $m_e > 0$ lie within A or within B. Note that all edges of $M_{x_1}, M_{y_1}, M_{x_1+y_1}$ are within A or within B. We may assume that however we permute the labels on A we still have $S_x = L_2$ for each $x \in L_1$. Otherwise all conditions of Lemma 4.1 are satisfied, and we are done. By possibly renaming A,B and the elements of L_2 we deduce that the sets T_e for edges e within A or within B have the following properties.

(1) Each of the three matchings of size two within A has an edge e on which $PG_2(2) \backslash T_e = \{x_2\}$.
(2) For each matching $\{e, f\}$ within B we have $PG_2(2) \backslash (T_e \cup T_f) = \{y_2, x_2 + y_2\}$

There must be 3 elements of $V_{48} = \{v_4, \dots, v_8\}$ in B and 2 elements of V_{48} in A, as if either set contained 4 then the one with smallest index would be incident to an edge e with $m_e > 0$ crossing between A and B. Since $\sum_{i=1}^{7} m_{v_i v_8} = 3$, it follows that $v_8 \in B$. Then the edges within A with $m_e > 0$ must form a triangle in which each edge e has $m_e = 1$. Let $a \in A$ be the vertex that is not in this triangle.

There are two possibilities for the edges in B. Consider first the case when there are $x, y \in B$ so that $m_{xv_8} = 2$, $m_{yv_8} = 1$. Let z be the other vertex of B. Then $m_{xz} = 1$ and $m_{xy} = 2$. The other case is when $m_e = 1$ for each edge in B. Then for each matching $\{e, f\}$ within B we have $PG_2(2) \backslash T_e = \{y_2\}$ and $PG_2(2) \backslash T_f = \{x_2 + y_2\}$, or vice versa. It follows that there is a vertex z such that $PG_2(2) \backslash T_e$ is the same set (say $\{y_2\}$) for all edges within B incident to z. In either case we consider the new labelling obtained by interchanging the labels a and z. In the first case the only effect on the sets S_i is that the edge formerly labelled xz now contributes to a different S_i. This labelling clearly satisfies the conditions of Lemma 4.1. In the second case we have a situation isomorphic to that described in condition (iv) of Lemma 4.1. In all cases we deduce that there is an automorphism $\pi \in L_3(2)$ such that $\pi(a) \notin S_a$, i.e. $\pi(a) \in \bigcap_{e \in M_a} T_e$, for all $a \in PG_2(2)$. Now we relabel our set of 8 vertices so that the vertex with label a has now has label $\pi(a)$. This has the effect of relabelling M_a as $M_{\pi(a)}$, so now we have $a \in \bigcap_{e\in M_a} T_e$ for all a. This completes the proof of the lemma. □

5. Concluding remarks and open problems

The construction for the lower bound in Theorems 1.1 and 1.2 seems very naïve, but we suspect it may be optimal for sufficiently large m. For $q = 2$ we have seen that it is not optimal for $m = 2$ or $m = 3$. The constructions that beat it are based on colouring properties of the projective geometries. If $t = \chi(PG_m(q))$ denotes the chromatic number of $PG_m(q)$ then a natural competing construction is to divide a set of vertices into $t - 1$ parts and take as edges all $(q + 1)$-tuples that do not fall entirely within one of the classes. This gives a lower bound on the Turán density of $1 - (t - 1)^{-q}$.

However, it is not hard to see that $\chi(PG_m(q)) \leqslant m$ for all $m, q \geqslant 2$ (except when $m = q = 2$ when we have $\chi(PG_2(2)) = 3$). Indeed, $\chi(PG_{m+1}(q)) \leqslant \chi(PG_m(q)) + 1$, as given $PG_{m+1}(q)$ we can colour a copy of $PG_m(q)$ inside it and then assign one new colour to all the other vertices. Pelikán [15] shows that $\chi(PG_3(2)) = 3$ and $\chi(PG_2(q)) = 2$ for all $q > 2$, so this verifies the observation. We remark in passing that the determination of $\chi(PG_m(q))$ in general is open. In addition to the results mentioned it is known that $\chi(PG_4(2)) = 4$ (see [15]) and $\chi(PG_5(2)) = 5$ (see [3]). Haddad [9] showed that $\lim_{m\to\infty} \chi(PG_m(2)) = \infty$, and conjectured that in fact $\chi(PG_m(2)) = m$ for $m \geqslant 3$. (We note that his proof actually shows that $\lim_{m\to\infty} \chi(PG_m(q)) = \infty$, via the result of [13] on the non-existence of blocking sets in $PG_m(q)$.)

In any case, the construction described above cannot give a lower bound better than $1 - (m - 1)^{-q}$ for the Turán density of $PG_m(q)$. For large m this is much worse than the lower bound from Theorem 1.1, which is $1 - O(q^{2-m})$. It seems unlikely that variations along the lines of [1] and our construction in Theorem 3.1 can close this gap, which provides some grounds for thinking that the naïve construction may eventually be optimal. Also, we note that the Lagrangian method (see Chapter 6 of [4]) easily shows that any $(q+1)$-uniform hypergraph with density at least $\prod_{i=1}^{q}(1 - i/t)$ contains some subhypergraph on at least t vertices that is a 2-cover, i.e. any pair of vertices is contained in some edge. Since $PG_m(q)$ is a 2-cover, this could be viewed as giving further support to this suggestion.

It would be very interesting to determine the Turán density of $PG_m(q)$ for any pair $m, q \geqslant 2$ other than the Fano plane. We have focussed particularly on the case $q = 2$, $m = 3$ in the latter part of this paper. Our methods suggest that further improvements on our upper bound should be possible, although it would be preferable to achieve this without magnifying the amount of case analysis required. An obvious starting point is to try to reduce the constant 6.5 in the statement of Lemma 3.2. Considering the complete balanced 7-partite graph in which every edge has multiplicity 7, we see that this constant cannot be less than 6, so the best possible upper bound that could be proved by our methods is 6/7. Note that $6/7 \sim 0.857143$ is fairly close to our lower bound of $3\sqrt{3}+2\sqrt{2(9 - 5\sqrt{3})}-6 \sim 0.844778$. It is hard to make a plausible conjecture on the true value of $\pi(PG_3(2))$. The construction seems simple enough that it is hard to imagine beating it, but the same could have been said of that in [1]. Also, it would be somewhat surprising if the density was irrational.

Our proof of the auxiliary Lemma 4.1 (given in Appendix B) is somewhat ad hoc, and our only concern is to obtain the minimum necessary for the proof of Lemma 3.2. However, we believe that the following general problem may be interesting.

Problem 5.1 (*Group marriage problem*). *Let G be a group acting on a set X. Suppose we have a subset $A_x \subset X$ for each $x \in X$. When is there an element $g \in G$ so that $gx \in A_x$ for each $x \in X$?*

Note that in the case when G is the entire symmetric group the problem asks for a system of distinct representatives of the sets $\{A_x : x \in X\}$. This is an equivalent formulation of Hall's marriage problem, and the necessary and sufficient condition is that $|\bigcup_{y\in Y} A_y| \geqslant |Y|$ for each $Y \subset X$. For general G, an obvious necessary condition is the *orbit condition* that $\bigcup_{y\in Y} A_y$ must contain an orbit of Y for each $Y \subset X$. However, this is not sufficient. For example, suppose G is the cyclic group of order 3 acting on $\{1, 2, 3\}$ via the permutation (123) and let $A_1 = \{2\}$, $A_2 = \{1\}$, $A_3 = \{3\}$. It is easy to check that $\bigcup_{y\in Y} A_y$ contains an orbit of Y for each $Y \subset \{1, 2, 3\}$, but the only possible marriage is the permutation (12), which is not in the group. It would be interesting to classify the groups G for which the orbit condition is sufficient for the group marriage problem. It seems likely that 'sufficiently transitive' groups should have this property, and in particular that $L_3(2)$ should. A nice proof of this would make the proof of Lemma 3.2 much cleaner.

Appendix A. Proof of Lemma 2.2

Note that if s is an integer we have $\sum_{i=1}^{k-1} \lfloor 1 + i(s+t) \rfloor = s\binom{k}{2} + \sum_{i=1}^{k-1} \lfloor 1 + it \rfloor$. Since $(2^m - 1)\left(\binom{2^m}{2} - 1\right) = (2^m - 2)\binom{2^m}{2} + \left(\binom{2^m}{2} - 2^m + 1\right)$ we have $f\left(2^m, (2^m - 1)\left(\binom{2^m}{2} - 1\right)\right) = 2^m - 2 + f\left(2^m, \binom{2^m}{2} - 2^m + 1\right)$. Therefore it suffices to compute $f\left(2^m, \binom{2^m}{2} - 2^m + 1\right)$.

First suppose $m \geqslant 3$ is odd. Let $p = (2^m + 1)/3$ (an integer). Note that if $i = xp + r$, with $1 \leqslant r \leqslant p$ then

$$\lfloor 1 + i(1 - 1/p) \rfloor = x(p-1) + r + \lfloor 1 - r/p \rfloor = i - x.$$

Therefore

$$\begin{aligned} \sum_{i=1}^{2^m-1} \lfloor 1 + i(1 - 1/p) \rfloor &= \sum_{i=1}^{p} i + \sum_{i=p+1}^{2p} (i-1) + \sum_{i=2p+1}^{2^m-1} (i-2) \\ &= \left(\sum_{i=1}^{2^m-1} i \right) - p - 2(2^m - 1 - 2p) \\ &= \binom{2^m}{2} + 3p - 2^{m+1} + 2 = \binom{2^m}{2} - 2^m + 3. \end{aligned}$$

On the other hand, if $t < 1-1/p$ then $\lfloor 1+pt\rfloor \leqslant p-1$ and $\lfloor 1+2pt\rfloor \leqslant 2p-2$, so $\sum_{i=1}^{2^m-1}\lfloor 1+it\rfloor \leqslant \binom{2^m}{2}-2^m+1$. By definition it follows that $f\left(2^m,\binom{2^m}{2}-2^m+1\right) = 1-1/p$ and so

$$f\left(2^m,(2^m-1)\left(\binom{2^m}{2}-1\right)\right) = 2^m-2+1-1/p = 2^m-1-\frac{3}{2^m+1}.$$

Now suppose m is even. Let $p=(2^m-1)/3$ (an integer). We have

$$\sum_{i=1}^{2^m-1}\lfloor 1+i(1-1/p)\rfloor = \sum_{i=1}^{p} i + \sum_{i=p+1}^{2p}(i-1) + \sum_{i=2p+1}^{2^m-1}(i-2) = \sum_{i=1}^{2^m-1} i - 3p$$
$$= \binom{2^m}{2} - 2^m + 1.$$

Let $\alpha = \frac{1}{p(2p+1)}$. Note that for $0 \leqslant \beta < \alpha$ we have $\lfloor 1+i(1-1/p+\beta)\rfloor = \lfloor 1+i(1-1/p)\rfloor$. For when $i \leqslant 2p+1$ we have $i\beta < 1/p$ and for $2p+2 \leqslant i \leqslant 2^m-1$ we have $i\beta < 2/p$; in neither case do we add enough to exceed $\lfloor 1+i(1-1/p)\rfloor + 1$. On the other hand, $\lfloor 1+i(1-1/p+\alpha)\rfloor$ is equal to $\lfloor 1+i(1-1/p)\rfloor$, except when $i=2p+1$ when we have $\lfloor 1+(2p+1)(1-1/p+\alpha)\rfloor = 2p = \lfloor 1+(2p+1)(1-1/p)\rfloor + 1$. Therefore $\sum_{i=1}^{2^m-1}\lfloor 1+i(1-1/p+\alpha)\rfloor = \binom{2^m}{2}-2^m+2$. We deduce that $f\left(2^m,\binom{2^m}{2}-2^m+1\right) = 1-1/p+\alpha = 1-2/(2p+1)$ and so

$$f\left(2^m,(2^m-1)\left(\binom{2^m}{2}-1\right)\right) = 2^m-2+1-2/(2p+1)$$
$$= 2^m-1-\frac{6}{2^{m+1}+1}.$$

This completes the proof of the lemma.

Appendix B. Proof of Lemma 4.1

We divide into cases according to the number of non-empty sets S_x. First we deal with the case when exactly one S_x is empty. Here condition (iv) determines the sets S_x precisely. We choose $\pi(0)=01$, $\pi(1)=02$, $\pi(2)=0$ and extend by linearity. This gives $\pi = (0,01,12,2)(1,02)(012)$ (using cycle notation), which satisfies the condition $\pi(x)\notin S_x$ for all x. Next we consider the case when exactly one S_x is non-empty, say S_0. By condition (i) we can pick $u \notin S_0$, and then any π with $\pi(0)=u$ will do.

Now suppose there are exactly two non-empty sets, say S_x and S_y. From condition (ii) we can suppose without loss of generality that $|S_y| \leqslant 5$. By condition (i) we can pick $u \in PG_2(2)\backslash S_x$ and then $v \in PG_2(2)\backslash(S_y \cup \{u\})$. Now $L_3(2)$ is transitive on pairs of elements in $PG_2(2)$, so we can choose π so that $\pi(x)=u$ and $\pi(y)=v$, as required.

Now consider the case when there are exactly three non-empty S_x. There are two possibilities, according to whether the three indexing elements are collinear or not. Suppose first that they are collinear. Since $L_3(2)$ is transitive on lines and can permute the points within a given line arbitrarily, we can assume that S_0, S_1, S_{01} are non-empty, with $|S_0| \leqslant |S_1| \leqslant |S_{01}|$. We need to find elements $a \neq b$ so that $a \notin S_0$, $b \notin S_1$ and $a+b \notin S_{01}$. Then we can define

$\pi(0) = a$, $\pi(1) = b$ and an arbitarily linear extension gives us the required automorphism. Let $T = \{a + b : a \notin S_0, b \notin S_1, a \neq b\}$. Suppose for the sake of contradiction that $T \subset S_{01}$. Since $|S_{01}| < 7$ there is $c \notin T$. Then for each $a \notin S_0$, $a \neq c$ we have $a + c \in S_1$, and so $|S_1| \geqslant 6 - |S_0|$. Therefore $9 \geqslant |S_0| + |S_1| + |S_{01}| \geqslant 6 + |S_{01}|$, so $|S_{01}| \leqslant 3$. It follows that $|S_0| = |S_1| = |S_{01}| = 3$. Note that if $b \notin S_1$ then we must have $b \notin S_0$ also, otherwise $|T| \geqslant |\{a + b : a \notin S_0\}| = 4$, which contradicts $T \subset S_{01}$. This shows that $S_0 \subset S_1$. Similarly we have $S_1 \subset S_0$, i.e. $S_0 = S_1$. Note that the complement of S_0 cannot contain a line $a, b, a + b$. For letting d denote its other element we see that T contains the distinct elements $a, b, a + b, a + d, b + d, a + b + d$, which contradicts $T \subset S_{01}$. It follows that $S_0 = S_1 = S_{01}$ is a line. This contradicts condition (v). We conclude that T is not contained in S_{01}, so the required automorphism π exists.

We also have the possibility that there are exactly three non-empty S_x and the indexing set is not collinear. Since $L_3(2)$ is transitive on non-collinear triples and can permute the points within a given non-collinear triple arbitarily, we can assume that S_0, S_1, S_2 are non-empty, with $|S_0| \leqslant |S_1| \leqslant |S_2|$. We need to find a non-collinear triple a, b, c such that $a \notin S_0$, $b \notin S_1$, $c \notin S_2$. Then we can define $\pi(0) = a$, $\pi(1) = b$, $\pi(2) = c$ and extend linearly to obtain the required automorphism. Suppose this is not possible. Then for each $a \notin S_0$, $b \notin S_1$, $a \neq b$ we have $S_2 \subset \{a, b, a + b\}$. Since $|S_0| + |S_1| + |S_2| \leqslant 9$ we must have $|S_0| = |S_1| = |S_2| = 3$. Then $S_2 = \{a, b, a + b\}$ for any $a \notin S_0$, $b \notin S_1$, $a \neq b$. In particular the 4 elements $a \notin S_0$ all belong to S_2, which is a contradiction. We conclude that the required automorphism exists, which finishes the case when there are exactly three non-empty S_x.

The last case, when the number of non-empty S_x is 4 or 5, is the most complicated. We may assume that S_{12} and S_{012} are empty. Pick any $a \notin S_0$ and extend to a basis $\{a, b, c\}$ of $\mathbb{F}_2^3$. Consider the following bipartite graph B_1. The left vertex class is $\{1, 01, 2, 02\}$, the right vertex class is $\{b, a+b, c, a+c, b+c, a+b+c\}$, and an element i in the left class is joined to an element j in the right class exactly when $j \in S_i$. We want to find a matching of the left class into the right class in the complement graph of B_1 in such a way that the pairs $\{1, 01\}$ and $\{2, 02\}$ are each matched with one of the pairs $\{b, a+b\}$, $\{c, a+c\}$, $\{b+c, a+b+c\}$. With such a matching we can construct the automorphism π by mapping $\{1, 01, 2, 02\}$ to their match, 0 to a, 12 to $\pi(1) + \pi(2)$ and 012 to $\pi(01) + \pi(2)$. Suppose for the sake of contradiction that no such matching exists.

Note that we can match $\{1, 01\}$ to $\{b, a + b\}$ in the complement exactly when B_1 restricted to these four vertices has maximum degree at most 1. Consider the following auxiliary bipartite graph B_2. The left vertex class consists of the pairs $\{1, 01\}$,$\{2, 02\}$, the right vertex class consists of the pairs $\{b, a + b\}$, $\{c, a + c\}$, $\{b + c, a + b + c\}$, and a pair on the left is joined to a pair on the right exactly when B_1 restricted to these four vertices has a vertex of degree 2. Note that a matching of the left class into the right class in the complement of B_2 will give us the matching in the complement of B_1 we require. By Hall's theorem (or direct analysis) we can do this unless one vertex on the left side is joined to all vertices on the right side, or some two vertices on the right side are joined to both vertices on the left side.

Consider the case when some two vertices on the right side, say $\{b, a+b\}$ and $\{c, a+c\}$, are joined to both $\{1, 01\}$ and $\{2, 02\}$. Then for each pair $\{x, a + x\}$ with $x = b, c$ the restrictions of B_1 to $\{1, 01, x, a + x\}$ and $\{2, 02, x, a + x\}$ each contain a vertex of degree

2. Therefore $|S_1|+|S_{01}|\geqslant 4$ and $|S_2|+|S_{02}|\geqslant 4$, so $|S_1|+|S_{01}|+|S_2|+|S_{02}|\geqslant 8$. Then by condition (ii) we have $|S_0|\leqslant 1$, so we can choose a vertex a' among $\{b, a+b, c, a+c\}$ with $a' \notin S_0$. Note that a' has degree at least 2 in B_1, as it is joined to at least one of $\{1, 01\}$ and at least one of $\{2, 02\}$. Now repeat the above construction using a' instead of a. Specifically, we let $\{a', b', c'\}$ be a basis of $\mathbb{F}_2^3$, from which we construct B_1' as B_1 was constructed from $\{a, b, c\}$, and construct B_2' from B_1' as B_2 was constructed from B_1. As above we conclude that in B_2' either one of the pairs $\{1, 01\}$ and $\{2, 02\}$ is joined to all of the pairs $\{b', a'+b'\}$, $\{c', a'+c'\}$, $\{b'+c', a'+b'+c'\}$, or some two of the pairs $\{b', a'+b'\}$, $\{c', a'+c'\}$, $\{b'+c', a'+b'+c'\}$, are joined to both of the pairs $\{1, 01\}$, $\{2, 02\}$. The first possibility cannot occur, as if say $\{1, 01\}$ is joined to all of $\{b', a'+b'\}$, $\{c', a'+c'\}$, $\{b'+c', a'+b'+c'\}$ then $|S_1|+|S_{01}|\geqslant 6$, which gives $|S_1|+|S_{01}|+|S_2|+|S_{02}|\geqslant 10$, a contradiction. Therefore some two of the pairs $\{b', a'+b'\}$, $\{c', a'+c'\}$, $\{b'+c', a'+b'+c'\}$, are joined to both of the pairs $\{1, 01\}$, $\{2, 02\}$. Recall however that a' is also joined to at least 2 of $\{1, 01, 2, 02\}$. This gives $|S_1|+|S_{01}|+|S_2|+|S_{02}|\geqslant 10$, a contradiction.

Therefore we are reduced to the case when one vertex of B_2, say $\{1, 01\}$, is joined to each of the vertices $\{b, a+b\}$, $\{c, a+c\}$ and $\{b+c, a+b+c\}$. Then for each pair $\{x, a+x\}$ there is a vertex of degree 2 in the restriction of B_1 to $\{1, 01, x, a+x\}$. Therefore $|S_1|+|S_{01}|\geqslant 6$. By assumption at least two other S_x are non-empty, so by condition (ii) $|S_1|+|S_{01}|$ is equal to 6 or 7, and each other S_x has at most 2 elements. Without loss of generality suppose that $|S_1|\geqslant|S_{01}|$. We claim that we can pick $a' \in S_1\backslash S_0$ with $a' \neq a$. For if $|S_1|+|S_{01}|=6$ then a does not belong to S_1 or S_{01}, $|S_1|\geqslant 3$ and $|S_0|\leqslant 2$, so we can pick $a' \in S_1\backslash S_0$. Also, if $|S_1|+|S_{01}|=7$ then $|S_1|\geqslant 4$ and $|S_0|\leqslant 1$ so we can choose $a' \in S_1\backslash S_0$ with $a' \neq a$. Again we repeat the construction using a' instead of a, letting $\{a', b', c'\}$ be a basis of $\mathbb{F}_2^3$ and constructing the bipartite graphs B_1' and B_2'. As before we conclude that in B_2' either one of the pairs $\{1, 01\}$ and $\{2, 02\}$ is joined to all of the pairs $\{b', a'+b'\}$, $\{c', a'+c'\}$, $\{b'+c', a'+b'+c'\}$, or some two of the pairs $\{b', a'+b'\}$, $\{c', a'+c'\}$, $\{b'+c', a'+b'+c'\}$, are joined to both of the pairs $\{1, 01\}$, $\{2, 02\}$. However, condition (ii) gives $|S_2|+|S_{02}|\leqslant 9-|S_1|-|S_{01}|\leqslant 3$, so the only possibility is that the pair $\{1, 01\}$ is joined to all of the pairs $\{b', a'+b'\}$, $\{c', a'+c'\}$, $\{b'+c', a'+b'+c'\}$ in B_2'.

By definition, for each pair $\{x', a'+x'\}$ there is a vertex of degree 2 in the restriction of B_1' to $\{1, 01, x', a'+x'\}$. In particular B_1' has at least 6 edges. Also a' is in S_1, which is an extra element not counted by B_1', so $|S_1|+|S_{01}|=7$. In particular $|S_1|\geqslant 4$, as $|S_1|\geqslant|S_{01}|$. Also, the two other non-empty S_x must each contain one element. In particular $|S_0|\leqslant 1$. Now we will divide into cases according to the structure of B_1. We define three possible types for the subgraph induced by $\{1, 01, x, a+x\}$; we say it has *type* 1 if 1 has degree 2, *type* 01 if 01 has degree 2, and *type* C if one of x or $a+x$ has degree 2. Note that a subgraph has at least one type, and possibly more than one. We say that B_1 has *type* (α, β, γ) if the three subgraphs $\{1, 01, x, a+x\}$ have types α, β and γ (in some order). Again, B_1 can have more than one type. Since $|S_1|\geqslant|S_{01}|$ and $|S_1|\geqslant 4$, it follows that B_1 has at least one type among $(1, 1, 1)$, $(1, 1, C)$, $(1, 1, 01)$ and $(1, C, C)$.

Suppose first that one type of B_1 is $(1, 1, 1)$, i.e. 1 is joined to all vertices except for a. Then $a \notin S_1$, by condition (i). Since $|S_1|+|S_{01}|=7$ we have $|S_{01}|=1$. By changing basis we can assume that $b \notin S_0$ and $c \notin S_0$. We can choose $a'=b$ in the construction of B_1'. Consider the induced subgraph of B_1' on $\{1, 01, a, a+b\}$. Since $1a$ is not an edge and $|S_{01}|=1$, the vertex of degree 2 must be $a+b$. In particular $a+b \in S_{01}$. The same argument applied to the choice $a' = c$ shows that $a + c \in S_{01}$. However this contradicts the fact that $|S_{01}| = 1$, so B_1 cannot have $(1, 1, 1)$ as a type.

Next suppose that B_1 has $(1, 1, C)$ as a type. We can suppose that $b, b + a, c, c + a$ belong to S_1 and that $b + c$ belongs to S_1 and S_{01}. From the previous paragraph we know that $a + b + c \notin S_1$. We can also suppose that $b \notin S_0$. Choose $a' = b$ and consider the subgraphs induced by B_1' on $\{1, 01, a, a + b\}$ and $\{1, 01, a + c, a + b + c\}$. Then 1 has degree 1 in each, so 01 has degree at least 1 in each. Together with $b + c \in S_{01}$ we get $|S_{01}| \geqslant 3$, i.e. $|S_1| + |S_{01}| \geqslant 8$, which is a contradiction. Therefore B_1 cannot have $(1, 1, C)$ as a type.

Now suppose that B_1 has $(1, 1, 01)$ as a type. We can suppose that $b, b + a, c, c + a$ belong to S_1 and that $b + c, a + b + c$ belong to S_{01}. Also, $b + c, a + b + c$ do not belong to S_1, as that would give type $(1, 1, C)$. We can suppose $b \notin S_0$ and choose $a' = b$. Consider the subgraph induced by B_1' on $\{1, 01, c, b + c\}$. Since 1 is not joined to $b + c$ the degree 2 vertex must be 01 or c. In particular $c \in S_{01}$. Similarly, considering $\{1, 01, a + c, a + b + c\}$ shows that $a + c \in S_{01}$. This gives $|S_{01}| \geqslant 4$, i.e. $|S_1| + |S_{01}| \geqslant 8$, which is a contradiction. Therefore B_1 cannot have $(1, 1, 01)$ as a type.

Finally, suppose that B_1 has $(1, C, C)$ as a type. Then $|S_1 \cap S_{01}| \geqslant 2$, so we can choose $a' \in (S_1 \cap S_{01}) \setminus S_0$. Now B_1' has at least 6 edges, which does not count the contribution of 2 that a' makes to $|S_1| + |S_{01}|$, so again we get the contradiction $|S_1| + |S_{01}| \geqslant 8$. We have shown that if there is no automorphism π with $\pi(x) \notin S_x$ for all x then B_1 must have a type that leads to a contradiction. Therefore the required automorphism exists, which completes the proof of the lemma.

References

[1] S.M. Cioabă, Bounds on the Turán density of $PG_3(2)$, Electronic J. Combin. 11 (2004) 1–7.

[2] D. de Caen, Z. Füredi, The maximum size of 3-uniform hypergraphs not containing a Fano plane, J. Combin. Theory B 78 (2000) 274–276.

[3] J. Fugère, L. Haddad, D. Wehlau, 5-chromatic Steiner triple systems, J. Combin. Des. 2 (1994) 287–299.

[4] Z. Füredi, Turán type problems, in: Surveys in Combinatorics, London Mathematical Society Lecture Note Series, vol. 166, Cambridge University Press, Cambridge, 1991, pp. 253–300.

[5] Z. Füredi, A. Kündgen, Turán problems for weighted graphs, J. Graph Theory 40 (2002) 195–225.

[6] Z. Füredi, O. Pikhurko, M. Simonovits, The Turán density of the hypergraph $\{abc, ade, bde, cde\}$, Electronic J. Combin. 10 (2003) 7pp.

[7] Z. Füredi, O. Pikhurko, M. Simonovits, On triple systems with independent neighborhoods, Combin. Probab. Comput., to appear.

[8] Z. Füredi, M. Simonovits, Triple systems not containing a Fano Configuration, Combin., Probab. Comput., to appear.

[9] L. Haddad, On the chromatic numbers of Steiner triple systems, J. Combin. Des. 7 (1999) 1–10.

[10] J.F. Humphreys, A Course in Group Theory, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1996.

[11] P. Keevash, B. Sudakov, The Turán number of the Fano plane, Combinatorica, to appear.

[12] P. Keevash, B. Sudakov, On a hypergraph Turán problem of Frankl, Combinatorica, to appear.

[13] F. Mazzocca, G. Tallini, On the nonexistence of blocking-sets in $\mathrm{PG}(n, q)$ and $\mathrm{AG}(n, q)$, for all large enough n, Simon Stevin 59 (1985) 43–50.

[14] D. Mubayi, V. Rödl, On the Turán number of Triple Systems, J. Combin. Theory Ser. A 100 (2002) 136–152.

[15] J. Pelikán, Properties of balanced incomplete block designs, Combinatorial Theory and its Applications, Balatonfüred, Colloquia Mathematica Societatics János Bolyai, vol. 4, North-Holland, Amsterdam, 1969, pp. 869–889.