Combinatorics, Probability and Computing (2013) 22, 21–54. © Cambridge University Press 2012
doi:10.1017/S0963548312000508

Applications of the Semi-Definite Method to the Turán Density Problem for 3-Graphs

VICTOR FALGAS-RAVRY[1†] and EMIL R. VAUGHAN[2‡]

[1]Institutionen för matematik och matematisk statistik, Umeå Universitet, 901 87 Umeå, Sweden
(e-mail: victor.falgas-ravry@math.umu.se)
[2]School of Electronic Engineering and Computer Science, Queen Mary, University of London,
Mile End Road, London E1 4NS, UK
(e-mail: emil79@gmail.com)

Received 13 October 2011; revised 28 September 2012

In this paper, we prove several new Turán density results for 3-graphs with independent neighbourhoods. We show:

$$\pi(K_4^-, C_5, F_{3,2}) = 12/49, \quad \pi(K_4^-, F_{3,2}) = 5/18 \ \text{ and } \ \pi(J_4, F_{3,2}) = \pi(J_5, F_{3,2}) = 3/8,$$

where J_t is the 3-graph consisting of a single vertex x together with a disjoint set A of size t and all $\binom{|A|}{2}$ 3-edges containing x. We also prove two Turán density results where we forbid certain induced subgraphs:

$$\pi(F_{3,2}, \text{ induced } K_4^-) = 3/8 \ \text{ and } \ \pi(K_5, \text{5-set spanning exactly 8 edges}) = 3/4.$$

The latter result is an analogue for K_5 of Razborov's result that

$$\pi(K_4, \text{4-set spanning exactly 1 edge}) = 5/9.$$

We give several new constructions, conjectures and bounds for Turán densities of 3-graphs which should be of interest to researchers in the area. Our main tool is 'Flagmatic', an implementation of Razborov's semi-definite method, which we are making publicly available. In a bid to make the power of Razborov's method more widely accessible, we have tried to make Flagmatic as user-friendly as possible, hoping to remove thereby the major hurdle that needs to be cleared before using the semi-definite method. Finally, we spend some time reflecting on the limitations of our approach, and in particular on which problems we may be unable to solve. Our discussion of the 'complexity barrier' for the semi-definite method may be of general interest.

AMS 2010 *Mathematics subject classification*: Primary 05D05
Secondary 05C65

† Supported by a postdoctoral grant from the Kempe foundation. Part of the work on this paper was done while the author was a PhD student at Queen Mary, University of London, supported by an EPSRC studentship.
‡ Supported by EPSRC grant EP/H016015/1.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

1. Introduction

Extremal graph and hypergraph theory have in recent years seen a string of results obtained by application of the semi-definite method of Razborov [36], a by-product of his flag algebra calculus.

With the notable exception of the Fano plane, most known Turán density results for 3-graphs have been obtained anew using his method, as well as some new results and the best known upper bounds for several other problems [37, 38, 5, 17, 6, 7]. Particularly impressive in this respect was Razborov's proof of Turán's conjecture under an (admittedly important) additional restriction [37]:

$$\pi(K_4,\ \text{induced 4-set spanning 1 edge}) = 5/9.$$

(More results in a similar vein have recently been established by Baber and Talbot [7]. In particular they showed that $\pi(K_4, H) = 5/9$ for some 3-graph H of order 6 with $\pi(H) = 3/4$.)

In this paper we use the semi-definite method to prove several new Turán density results. In Section 3.1 we develop the extremal theory of 3-graphs with independent neighbourhoods, proving:

$$\pi(K_4^-, C_5, F_{3,2}) = 12/49,$$
$$\pi(K_4^-, F_{3,2}) = 5/18,$$
$$\pi(J_4, F_{3,2}) = \pi(J_5, F_{3,2}) = 3/8,$$

where J_t is the 3-graph consisting of a vertex x together with a disjoint set A of size t and all $\binom{|A|}{2}$ 3-edges containing x. In Section 3.2, we prove two density results where we forbid certain induced subgraphs:

$$\pi(F_{3,2},\ \text{induced } K_4^-) = 3/8,$$
$$\pi(K_5,\ \text{5-set spanning 8 edges}) = 3/4.$$

The latter result is an analogue for K_5 of the aforementioned theorem of Razborov for K_4. In addition we provide a number of new bounds, constructions and conjectures.

In applying the semi-definite method, we use the publicly available Flagmatic package written by Vaughan to assist us with the calculations. The semi-definite method provides us with an efficient formalism for computing density bounds in extremal combinatorics. In the case of extremal 3-graph theory, it does this by reducing an initial problem of proving inequalities for subgraph densities to a semi-definite programming problem, which in some cases can be solved exactly with the aid of a computer – and this is where Flagmatic comes into play. We discuss the meaning of 'in some cases' in greater detail in Section 4. Let us only say for the moment that without extra ideas we cannot hope for a general extremal theory to emerge from a direct application of Razborov's semi-definite method.

However, given the difficulty of extremal 3-graph theory and the paucity of known results, the semi-definite method can still be quite helpful in providing many useful bounds and exact results as well as in guiding investigations towards attainable goals.

A major hurdle for mathematicians wishing to use the semi-definite method in their work is the need for a computer program to assist them in the calculations. Flagmatic

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

was designed with this hurdle in mind, and we have tried collaboratively to make it as user-friendly as possible.

As the flag algebra calculations involved in our proofs are very long and not terribly informative, we have produced 'certificates' rather than write them out in full. The certificates are available as ancillary files in our arXiv submission [16] as well as on the Flagmatic website, `flagmatic.org`, where the interested reader may also download a copy of Flagmatic for herself. In addition, our results have also been independently verified by Baber and Talbot [7].

Our proofs use computer assistance to enumerate certain graphs and to solve a semi-definite programming problem by finding some 'good' matrices. In practice, in some of the problems we consider we could just about do the enumeration by hand, and then pull some quite large positive semi-definite matrices out of a hat, thus removing the visible presence of the computer from the argument. This would take up many pages, however, and would not be very informative. We have therefore opted not to do so.

1.1. Structure of the paper

This paper is structured as follows. After introducing a small amount of notation, Section 2 is devoted to explaining how Flagmatic works, beginning with an exposition of the semi-definite method (Section 2.2), some remarks about Flagmatic (Section 2.3), a discussion of the proof certificates it produces (Section 2.4), and some remarks on the additional structural information we can obtain from our proofs (Section 2.5).

Section 3 contains our main results. In Section 3.1 we develop an extremal theory of 3-graphs with independent neighbourhoods, proving the first set of results mentioned above and providing several new constructions and conjectures. In Section 3.2 we consider forbidding induced subgraphs, obtaining in particular a theorem related to the conjecture of Turán that $\pi(K_5) = 3/4$. In Section 3.3 we go on to discuss non-principality.

Finally in Section 4 we consider the limits inherent to our approach, in particular the 'complexity barrier' it runs into. We end with some questions and a summary of results and constructions.

2. The semi-definite method

2.1. Some notation and definitions

We begin with some notation and definitions, most of which are standard. A 3-*graph* G is a pair of sets $G = (V, E)$, with $V = V(G)$ a set of vertices, and $E = E(G)$ a collection of 3-sets from V, which are the 3-*edges* of G. A *subgraph* of G is a a 3-graph H with $V(H) \subseteq V(G)$ and $E(H) \subseteq E(G)$. An *induced subgraph* is a subgraph H with $E(H) = \{e \in E(G) : e \subseteq V(H)\}$.

Given a family of 3-graphs $\mathcal{F}$, we say that a 3-graph G is $\mathcal{F}$-*free* if G contains no member of $\mathcal{F}$ as a subgraph. We write $\mathrm{ex}(n, \mathcal{F})$ for the maximal number of 3-edges that can be present in an $\mathcal{F}$-free 3-graph. The non-negative function $\mathrm{ex}(n, \mathcal{F})$ is referred to as the *Turán number* of $\mathcal{F}$.

A standard averaging argument shows that $\mathrm{ex}(n, \mathcal{F})/\binom{n}{3}$ is non-increasing and hence tends to a limit as $n \to \infty$. This limit, denoted by $\pi(\mathcal{F})$, is the *Turán density* of $\mathcal{F}$, and

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

is the asymptotically maximal proportion of edges that can be present in an $\mathcal{F}$-free 3-graph.

The standard *Turán* (*density*) *problem* for 3-graphs is: given a family $\mathcal{F}$, determine $\pi(\mathcal{F})$. The analogous question for 2-graphs has been completely answered by the Erdős–Stone Theorem [14]; by contrast very few Turán densities of 3-graphs are known. (See the recent survey paper of Keevash [26] for details.)

Let us define some of the 3-graphs that often appear in this paper. Write $[n]$ for $\{1, 2, \ldots, n\}$ and $[n]^{(r)}$ for the collection of r-subsets of $[n]$. When enumerating 3-edges in this paper, we shall usually write xyz for $\{x, y, z\}$. When no confusion is possible, we may also use 'edge' for '3-edge' and 'graph' for '3-graph'.

The *complete* 3-graph on t vertices is the 3-graph $K_t = ([t], [t]^{(3)})$. Deleting a single 3-edge from K_t yields a copy of K_t^-, the unique (up to isomorphism) 3-graph on t vertices with $\binom{t}{3} - 1$ edges. We let C_5 denote the (*strong*) 5-*cycle* $C_5 = ([5], \{123, 234, 345, 451, 512\})$, and write $F_{3,2}$ for the 3-graph $F_{3,2} = ([5], \{123, 145, 245, 345\})$.

We say that a particular instance of the Turán problem for 3-graphs is *stable* if there is a sequence of 3-graphs

$$G_1, G_2, \ldots, G_n, \ldots$$

such that for any $\varepsilon > 0$ there exists $\delta > 0$ and $n_0 \in \mathbb{N}$ such that any $\mathcal{F}$-free 3-graph on $n \geqslant n_0$ vertices with more than $(\pi(\mathcal{F}) - \delta)\binom{n}{3}$ 3-edges can be transformed into (an isomorphic copy of) G_n by adding or deleting at most εn^3 of the 3-edges. (Intuitively, this says there is an essentially unique extremal configuration, and that any 'close to extremal' 3-graph must lie at a small *edit distance* from it: at most εn^3.)

We shall also touch on *links*.

Definition 2.1. Given a 3-graph G and $x \in V(G)$, the *link graph* (or *link*) of x in G is the 2-graph

$$G_x = \big(V \setminus \{x\}, \{ab : xab \in E(G)\}\big).$$

We shall consider the problem of forbidding the links of a 3-graph from containing a complete 2-graph on t vertices, and we define J_t to be the corresponding forbidden 3-subgraph, namely

$$J_t = \big([t+1], \{\{x, y, t+1\} : \{xy\} \in [t]^{(2)}\}\big).$$

This 3-graph J_t is a special case of a 'suspension' (namely the 3-suspension of $K_t^{(2)}$); in the more general notation due to Keevash [26] it is denoted by $S^3 K_t^2$.

Various constructions we consider in this paper involve taking a (possibly unbalanced) partition of the vertex set $V = A_1 \sqcup A_2 \sqcup \cdots \sqcup A_r$ and then adding edges according to some rule. In this setting, a 3-edge has *type* $A_i A_j A_k$ if it is of the form xyz with $x \in A_i$, $y \in A_j$, $z \in A_k$.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Definition 2.2. A *blow-up* construction is obtained by taking a 3-graph H on $V(H) = [r]$ with some possibly degenerate edges – for example '112' or '333' – and using it as a template to construct graphs of order n for every $n \in \mathbb{N}$ as follows:

- partition $[n]$ into r parts $A_1 \sqcup A_2 \sqcup \cdots \sqcup A_r$,
- add all edges of type $A_iA_jA_k$ with $ijk \in E(H)$.

An *iterated blow-up* construction is obtained, as the name suggests, by taking a blow-up construction from a template H and then repeating the construction inside (some of) the $|V(H)|$ parts of the resulting 3-graph, and then again in the resulting subparts, and so on. The partition and edges obtained by the first iteration are said to be at *level 1* of the construction, the subpartition and edges given by the second iteration are said to lie at *level 2*, and so on.

Finally and most importantly, we define the notion of *induced subgraph density*, which is central to the theory of flag algebras.

Definition 2.3. Given a 3-graph G of order $|V(G)| = n$ and a 3-graph H of order $m \leqslant n$, the (*induced*) *subgraph density* of H in G, denoted by $d_H(G)$, is the probability that an m-subset of $V(G)$ chosen uniformly at random induces a copy of H in G, *i.e.*, that the resulting random subgraph of G is isomorphic to H. When H is the 3-edge $([3], \{123\})$, we write $d(G)$ for $d_H(G)$ and call it the (*edge*) *density* of G.

2.2. Mantel's theorem via the semi-definite method

In this subsection, we give an overview of the semi-definite method. As mentioned previously, this method is a by-product of the flag algebra calculus of Razborov [36]. It consists of an efficient formalism for converting the problem of proving certain inequalities between subgraph densities into a semi-definite programming problem, which can then be solved with the aid of a computer. Excellent expositions of this method from an extremal combinatorics perspective have already appeared in the literature; our presentation draws in particular on Section 2.1 of [6] and Section 7 of [26].

For ease of notation and the sake of clarity, we shall consider 2-graphs rather than 3-graphs for our exposition, in contrast to [6, 26]. Razborov [36] in fact defined his flag algebra calculus in a much more general setting which includes 2-graphs and 3-graphs as special cases; we feel that the 2-graph case gives all the intuition necessary, while keeping calculations to a minimum.

Let $K_3^{(2)}$ denote the complete 2-graph on 3 vertices, otherwise known as the triangle. To illustrate our discussion, we shall use the following weak form of Mantel's theorem as a running example.

Theorem 2.4.

$$\pi(K_3^{(2)}) = 1/2.$$

What would be the crudest possible way of finding a non-trivial upper bound on $\pi(K_3^{(2)})$? We could observe that a triangle-free graph G on n vertices is at most as dense as

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

the most dense subgraph of order $m \leqslant n$ that it contains. Note that as G is triangle-free, so are its subgraphs. Say therefore that a subgraph is *admissible* if it is triangle-free and so could occur as a subgraph of G. Pick some integer m, and let $\mathcal{H}$ denote the collection of all *admissible subgraphs* of order m up to isomorphism. We then have

$$d(G) = \sum_{H \in \mathcal{H}} d_H(G) d(H) \tag{2.1}$$

with $\sum_{H \in \mathcal{H}} d_H(G) = 1$, and thus

$$d(G) \leqslant \max_{H \in \mathcal{H}} d(H). \tag{2.2}$$

This is fairly obviously a poor way to go about bounding $\pi(K_3^{(2)})$. Indeed pick for example $m = 3$. The family $\mathcal{H}$ then consists of three graphs H_0, H_1, H_2, with H_i being the unique (up to isomorphism) graph on 3 vertices with exactly i edges. Thus (2.2) shows $\pi(K_3^{(2)}) \leqslant 2/3$, but this could only be sharp if *all* induced subgraphs of order 3 were isomorphic to H_2. This is impossible for $n \geqslant 5$. Indeed, suppose we have x, y with xy a non-edge, and a, b, c such that $\{xya\}, \{xyb\}$ and $\{xyc\}$ all induce copies of H_2 in G. Then as G is triangle-free, $\{abc\}$ must induce a copy of H_0. We therefore expect the density of H_0 in G to be bounded below by some function of the density of H_2 (the density of H_1 being determined by the fact that $\sum_i d_{H_i}(G) = 1$). Thus one way we could try to refine inequality (2.2) would be to take such a relationship into account and exploit it to improve our bound.

The simplest relationship of this kind we could hope for is a linear inequality for subgraph densities of the form

$$\sum_{H \in \mathcal{H}} d_H(G) a_H \geqslant 0.$$

Given such an inequality, it follows from (2.1) that

$$\begin{aligned} d(G) &\leqslant \sum_{H \in \mathcal{H}} d_H(G)(d(H) + a_H) \\ &\leqslant \max_{H \in \mathcal{H}} (d(H) + a_H). \end{aligned}$$

Provided our linear inequality is 'good', the a_H 'even out' the coefficients $d(H) + a_H$ by transferring weight from dense subgraphs to sparser ones, improving on (2.2).

Following this line of thought, we then ask ourselves: How can we produce (good) linear inequalities for subgraph densities? Our remark on the fact that we cannot pack a graph full of induced copies of H_2 suggests a possible answer: We can consider the ways in which different kinds of subgraphs can intersect, and from this information derive bounds on subgraph densities. What Razborov's flag algebra calculus gives us is an efficient formalism for doing just that, which we now present.

Suppose we work in the general framework of $\mathcal{F}$-free graphs. (Our example had $\mathcal{F} = \{K_3^{(2)}\}$.) Let m be an integer, which we shall fix later on, and let $\mathcal{H}$ denote as before the set of all (up to isomorphism) admissible subgraphs of order m.

An *intersection type* is a graph on a labelled vertex set, with every vertex having a distinct label. Given an intersection type σ, a *σ-flag* is an admissible graph F on a partially labelled

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

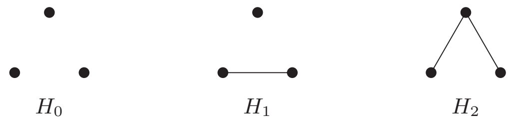

Figure 1. The admissible graphs.

vertex set such that the subgraph induced by the labelled vertices is a copy of σ (with identical labels for the vertices.) For example, let us consider the intersection type σ consisting of a single vertex labelled '1'. Then there are (up to isomorphism) two σ-flags of order 2, namely F_0 consisting of a non-edge with one end labelled '1', and F_1 consisting of an edge with one end labelled '1' (see Figure 2). We shall write $\mathcal{F}_\sigma^l$ for the collection of all (up to isomorphism) σ-flags of order l.

Let us now define some flag densities. Fix an intersection type σ of order $|V(\sigma)| = s$ and an integer $l \geqslant s$. Given a graph G, select a partial labelling of $V(G)$ with the labels from σ, chosen uniformly at random (by which we mean: randomly select $|V(\sigma)|$ vertices and assign them distinct labels from σ). This makes G into a potential σ-flag. Note that the labelled vertices could fail to induce a copy of σ, and that we allow this. Now select a set S_1 of $l - s$ other vertices (necessarily unlabelled) uniformly at random. Taken together with the labelled vertices, S_1 gives us a potential σ-flag of order $(l - s) + s = l$; so, given $F \in \mathcal{F}_\sigma^l$, write $d_F(G)$ for the probability that this potential σ-flag is a copy of F. We call this the *flag density* of F in G.

Having selected S_1, pick a disjoint set S_2 of $l - s$ unlabelled vertices uniformly at random. Taken together with the partially labelled vertices, S_1 and S_2 give us two potential σ-flags of order l. Then, given $F, F' \in \mathcal{F}_\sigma^l$, let $d_{F,F'}(G)$ be the probability that S_1 and S_2 induce copies of F and F' respectively when taken together with the labelled vertices. We call $d_{F,F'}(G)$ the *flag pair density* of (F, F') in G.

Finally, for a fixed partial labelling θ of $V(G)$ with labels from σ, select an $(l - s)$-set S_1 from the unlabelled vertices of G uniformly at random, and write $d_F^\theta(G)$ for the probability that S_1 together with the vertices labelled by θ induces a copy of the the σ-flag F. Then select a disjoint $(l - s)$-set S_2 uniformly at random from the remaining unlabelled vertices and write $d_{F,F'}^\theta(G)$ for the probability that S_1 and S_2 induce copies of F and F' respectively when taken together with the vertices labelled by θ.

In our running example with σ consisting of a single vertex labelled '1', $d_{F_1}(G)$ measures the probability that if we randomly label a vertex x in $V(G)$ and randomly select $y \in V(G) \setminus \{x\}$ then $xy \in E(G)$: in other words, $d_{F_1}(G)$ is exactly the edge density of G. On the other hand, $d_{F_1,F_0}(G)$ measures something slightly more complicated. Letting $n = |V(G)|$ and writing $d(v)$ for the degree of v in G, we have

$$d_{F_1,F_0}(G) = \sum_{v \in V(G)} \frac{1}{n} \left(\frac{d(v)}{n-1}\right) \left(\frac{n-1-d(v)}{n-2}\right).$$

This is in fact exactly $(d_{H_2}(G) + d_{H_1}(G))/3$. Similarly interesting from a combinatorial perspective is

$$d_{F_1,F_1}(G) = d_{H_2}(G)/3 + d_{K_3^{(2)}}(G),$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

which in a triangle-free graph measures the H_2 density (divided by 3).

Now, let us fix σ, l and make two easy observations. Firstly, if $n = |V(G)|$ is sufficiently large, then picking two random extensions of order $l-s$ for a randomly labelled set of s vertices is essentially the same as picking a random pair of *disjoint* extensions: indeed, the probability that two randomly chosen $(l-s)$-sets from $V(G)$ intersect is $O(1/n)$.

Observation 2.5. For all $F, F' \in \mathcal{F}^l_\sigma$, for all possible partial labellings θ of $V(G)$ with labels from σ,

$$d^\theta_F(G) d^\theta_{F'}(G) = d^\theta_{F,F'}(G) + O(1/n).$$

In particular, taking expectations over θ on both sides, we have

$$\mathbb{E}_\theta\, d^\theta_F(G) d^\theta_{F'}(G) = d_{F,F'}(G) + O(1/n).$$

Secondly, we can take averages.

Observation 2.6. Let m be any integer with $m \geqslant 2l - s$, and let $\mathcal{H}$ be the family of all (up to isomorphism) admissible subgraphs of order m defined earlier. Then, for all $F, F' \in \mathcal{F}^l_\sigma$,

$$d_{F,F'}(G) = \sum_{H \in \mathcal{H}} d_H(G) d_{F,F'}(H).$$

The appearance of the $d_H(G)$ terms in Observation 2.6 suggests that we are close to achieving our goal. And indeed, let Q be any fixed positive semi-definite $|\mathcal{F}^l_\sigma| \times |\mathcal{F}^l_\sigma|$ matrix with entries indexed by $\mathcal{F}^l_\sigma$. Then

$$\begin{aligned}
0 &\leqslant \mathbb{E}_\theta \sum_{F,F' \in \mathcal{F}^l_\sigma} Q_{F,F'} d^\theta_F(G) d^\theta_{F'}(G) && \text{(by positive semi-definiteness)} \\
&= \sum_{F,F' \in \mathcal{F}^l_\sigma} Q_{F,F'} d_{F,F'}(G) + O(1/n) && \text{(by Observation 2.5)} \\
&= \sum_{F,F' \in \mathcal{F}^l_\sigma} Q_{F,F'} \sum_{H \in \mathcal{H}} d_H(G) d_{F,F'}(H) + O(1/n) && \text{(by Observation 2.6)} \\
&= \sum_{H \in \mathcal{H}} d_H(G) \Bigg(\sum_{F,F' \in \mathcal{F}^l_\sigma} Q_{F,F'} d_{F,F'}(H) \Bigg) + O(1/n), && (2.3)
\end{aligned}$$

by changing order of summation again in the last line.

This is of the desired form $0 \leqslant \sum_{H \in \mathcal{H}} d_H(G) \lambda_H + O(1/n)$ (the $O(1/n)$ error term being irrelevant when bounding the Turán density). Thus for a fixed m, every choice of σ and l such that $2l - |V(\sigma)| \leqslant m$, and positive semi-definite matrix Q, gives us some linear inequality between subgraph densities for admissible subgraphs of order m. We can then sum these inequalities together. For example, if we have r choices,

$$(\sigma_1, l_1, Q_1), (\sigma_2, l_2, Q_2), \ldots, (\sigma_r, l_r, Q_r),$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

we can add the corresponding inequalities (2.3) to get

$$0 \leqslant \sum_{H \in \mathcal{H}} d_H(G) \Bigg(\sum_{i=1}^{r} \sum_{F,F' \in \mathcal{F}_{\sigma_i}^{l_i}} (Q_i)_{F,F'} d_{F,F'}(H) \Bigg) + O(1/n).$$

With a view to getting the best possible improvement of (2.2), we can, for a fixed choice of $(\sigma_1, l_1), (\sigma_2, l_2), \ldots, (\sigma_r, l_r)$, optimize the choice of the matrices $Q_1, Q_2, \ldots, Q_r$ to obtain a 'best inequality possible':

$$0 \leqslant \sum_{H \in \mathcal{H}} d_H(G) a_H + O(1/n),$$

where

$$a_H = \sum_{i=1}^{r} \sum_{F,F' \in \mathcal{F}_{\sigma_i}^{l_i}} (Q_i)_{F,F'} d_{F,F'}(H).$$

We can add this to (2.1) to get

$$d(G_n) \leqslant \sum_{H \in \mathcal{H}} d_H(G_n)(d(H) + a_H) + O(1/n), \tag{2.4}$$

and thus obtain a bound on the Turán density of our family $\mathcal{F}$ of forbidden subgraphs,

$$\pi(\mathcal{F}) \leqslant \max_{H \in \mathcal{H}} (d(H) + a_H). \tag{2.5}$$

We refer to (2.5) as the *flag algebra bound*, and for each $H \in \mathcal{H}$ we call $d(H) + a_H$ the *flag algebra coefficient* of H in the bound.

At this point, let us make two important observations.

Lemma 2.7. *Suppose the flag algebra bound is tight, i.e.,*

$$\pi(\mathcal{F}) = \max_{H} (d(G) + a_H),$$

and there is an admissible subgraph H' whose flag algebra coefficient is ρ, where $\rho < \pi(\mathcal{F})$. Then, for any sequence of $\mathcal{F}$-free graphs $(G_n)_{n \in \mathbb{N}}$ with $|V(G_n)| = n$ and $e(G_n) = (\pi(\mathcal{F}) + o(1))\binom{n}{2}$, we have

$$\limsup_{n \to \infty} d_{H'}(G_n) = 0.$$

Proof. Indeed, suppose $\limsup_{n \to \infty} d_{H'}(G_n) > \varepsilon$ for some fixed $\varepsilon > 0$. Then, by (2.4), we have for infinitely many n,

$$d(G_n) < \varepsilon\rho + (1 - \varepsilon)\pi(\mathcal{F}) + O(1/n),$$

which is bounded away from $\pi(\mathcal{F})$ for n large enough, a contradiction. □

Similarly, a consequence of requiring the flag algebra bound to be tight is that

$$\sum_{F,F' \in \mathcal{F}_{\sigma}^{l}} \mathbb{E}_\theta \, Q_{F,F'} d_F^\theta(G) d_{F'}^\theta(G) = O(1/n) \tag{2.6}$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

for all our optimized choices of (σ, l, Q) and graphs G that are 'close' to being extremal.

Further, if the problem has a blow-up construction as a stable extremum, consider the 'limit' of a sequence of extremal configurations G_n as $n \to \infty$. For all $F \in \mathcal{F}_\sigma^l$, the quantity $d_F^\theta(G_n)$ is determined (up to $o(1)$) by the parts of the blow-up construction in which we set the labelled vertices; in particular, if θ and θ' place the same labels in the same parts, then $d_F^\theta(G_n) = d_F^{\theta'}(G_n) + o(1)$ for all choices of F, and we may treat θ and θ' as being 'equivalent'. We can reduce in this way the set of all partial labellings into a finite set of 'equivalence' classes.

To illustrate this informal discussion with an example, suppose the extremal configurations G_n consist of complete balanced bipartite graphs and that $|V(\sigma)| = 2$. Then there are two 'equivalence' classes of partial labellings: one in which both labelled vertices are put in the same part of G_n, and one in which the labelled vertices are assigned to different parts of G_n.

Now for each 'equivalence' class, choose a sequence of distinct representatives, *i.e.*, a sequence of partial labellings of n-vertex extremal configurations, and write U for the finite set of sequences thus defined. Since Q is positive semi-definite, we have the following.

Remark 2.8 (Baber [5]). Suppose the flag algebra bound is tight, and that the problem has a blow-up construction as a stable extremum.

Let (σ, l, Q) be one of our optimized choices of intersection type, flag order and matrix. Then, for any $(\theta_n)_{n\in\mathbb{N}} \in U$, where U is the set of sequences of partial labellings informally defined above,

$$\lim_{n\to\infty} \sum_{F,F'\in\mathcal{F}_\sigma^l} Q_{F,F'} d_F^{\theta_n}(G_n) d_{F'}^{\theta_n}(G_n) = 0.$$

In other words, the vectors of flag densities associated with a fixed embedding of σ in a large extremal configuration accumulate around the set consisting of the zero vector and of the zero eigenvectors of the positive semi-definite matrix Q. This remark was first made in a more formal infinitary setting by Baber (Lemma 2.4.4 in [5]), to whom we refer the reader for a rigorous proof.

Having made these two observations, let us return to our running example as an illustration of how Razborov's method is used to provide upper bounds for Turán densities. Recall that we are trying to show $\pi(K_3^{(2)}) \leqslant 1/2$ using the semi-definite method. In this case consideration of one intersection type suffices, namely the type σ consisting of a single labelled vertex. We have two σ-flags of order 2, F_0 and F_1, and three admissible subgraphs of order 3, H_0, H_1 and H_2 (see Figures 1 and 2). Let us compute $d_{F,F'}(H)$ for all possible choices of F, F' and H.

Since our intersection type σ consists of a single vertex, our random labelling and our two random extensions always give us an ordered pair of σ-flags, so that $\sum_{F,F'} d_{F,F'}(H) = 1$. Now it is easy to see that $d_{F_0,F_0}(H_0) = 1$, and that $d_{F,F'}(H_0) = 0$ for all other choices of F, F'. Next, we see that $d_{F_0,F_0}(H_1) = 1/3$, as the only way of getting two copies of F_0 is to label the unique degree zero vertex in H_1 '1' (which happens a third of the

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Table 1.

	$d_{F_0,F_0}(H)$	$d_{F_0,F_1}(H) = d_{F_1,F_0}(H)$	$d_{F_1,F_1}(H)$
H_0	1	0	0
H_1	1/3	1/3	0
H_2	0	1/3	1/3

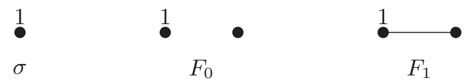

Figure 2. The intersection type σ, and σ-flags F_0 and F_1.

time), and that with this labelling we always get two copies of F_0 in the randomly chosen extensions. Also $d_{F_1,F_1}(H_1) = 0$ as H_1 contains only one edge, so that we have by symmetry $d_{F_0,F_1}(H_1) = d_{F_1,F_0}(H_1) = 1/3$. We then get the $d_{F,F'}(H_2)$ for free by noting that H_2 is the complement of H_1 and F_0 is the complement of F_1, so that $d_{F_\epsilon,F_\eta}(H_2) = d_{F_{1-\epsilon},F_{1-\eta}}(H_1)$. We give a summary in Table 1.

Now let

$$Q = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

be a positive semi-definite matrix. (In other words a, b, c, d satisfy $a, d \geqslant 0,\ ad - bc \geqslant 0$.) Then, in any triangle-free graph G of order n,

$$d(G) \leqslant d_{H_0}(G)(0 + a_{H_0}) + d_{H_1}(G)\left(\frac{1}{3} + a_{H_1}\right) + d_{H_2}(G)\left(\frac{2}{3} + a_{H_2}\right) + O(1/n),$$

where the a_{H_i} are the coefficients introduced earlier, given by

$$\begin{aligned} a_{H_0} &= a, \\ a_{H_1} &= a/3 + b/3 + c/3, \\ a_{H_2} &= b/3 + c/3 + d/3. \end{aligned}$$

We now optimize the choice of Q. Guessing that extremal triangle-free graphs are complete bipartite, we expect by Lemma 2.7 that both H_0 and H_2 must both have flag algebra coefficients equal to $1/2$ in a tight flag algebra bound. It is then a straightforward exercise in calculus to work out that

$$Q = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix}$$

is an optimal choice of matrix.

Our optimal inequality is then

$$0 \leqslant \frac{d_{H_0}(G)}{2} - \frac{d_{H_1}(G)}{6} - \frac{d_{H_2}(G)}{6} + O(1/n),$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

giving

$$d(G) \leqslant \frac{1}{2}(d_{H_0}(G) + d_{H_2}(G)) + \frac{1}{6}d_{H_1}(G) + O(1/n).$$

Taking the limit as $n \to \infty$, we deduce that $\pi(K_3^{(2)}) \leqslant 1/2$. Since a complete balanced bipartite graph achieves density $1/2 + o(1)$, we must in fact have equality. We have thus proved Theorem 2.4. (In fact we have proved a little more: our inequality tells us exactly which subgraphs can have positive density in an extremal example, and what those positive densities are, namely $d_{H_0}(G) = 1/4 + o(1)$ and $d_{H_2}(G) = 3/4 + o(1)$. This information can then be used to show that 'close' to extremal triangle-free graphs are 'close' to complete bipartite. See Section 2.5 for details.)

In general it is not practical to do the optimization above by hand (or indeed to perform manually all of the earlier calculations required to determine $\mathcal{H}, \mathcal{F}_\sigma^l$ and the $d_{F,F'}(H)$ terms), and this is where Flagmatic comes in: taking as input a set of forbidden configurations $\mathcal{F}$ and an integer m, it performs all the required computations, feeds the problem in an appropriate form into a semi-definite problem solver (SDP solver), then converts the SDP solver output into a bound on $\pi(\mathcal{F})$ and produces a 'certificate' of the flag algebra calculation. We discuss all this in detail in the following subsections.

2.3. Flagmatic

All the upper bounds on Turán densities that we give in this paper have been obtained by flag algebra calculations assisted by the Flagmatic package written by Vaughan to implement the semi-definite method. In this subsection we make some remarks concerning Flagmatic, and, in particular, how it obtains exact solutions. Note that in the remainder of the paper, starting from this section, we shall write 'graph' for '3-graph'.

Flagmatic takes as input a family of forbidden graphs $\mathcal{F}$, and an integer m. It then determines $\mathcal{H}$, the family of all admissible ($\mathcal{F}$-free) graphs of order m, up to isomorphism, and generates a set of intersection types and flags to use. By default, Flagmatic will use all intersection types σ whose order is congruent to m modulo 2. For each σ, Flagmatic takes $\mathcal{F}_\sigma^l$ with $l = (m - |V(\sigma)|)/2$ as its family of σ-flags. Flagmatic then computes the densities $d(H)$, for each $H \in \mathcal{H}$, and all the flag pair densities $d_{F,F'}(H)$ for all $H \in \mathcal{H}$ and all pairs $F, F' \in \mathcal{F}_\sigma^l$.

(It is not hard to show that if we use a type σ of order s, where s is not congruent to m modulo 2, then we can achieve at least as good a bound by replacing σ with all the types of order $s+1$ that contain σ as a labelled subgraph. For this reason, if we include all types whose order is congruent to m modulo 2, then the bound we get will be no worse than if we use all the types.)

Flagmatic uses the semi-definite program (SDP) solver 'CSDP' [11] to find symmetric matrices $Q_1, Q_2, \ldots, Q_r$ that optimize the flag algebra bound (2.5). (Note that in the search for optimal matrices, we may assume that each Q_i is symmetric, for otherwise we could replace Q_i by $(Q_i + Q_i^T)/2$ without changing a_H.) As is standard for this kind of software, CSDP uses floating-point arithmetic, which presents us with a number of issues (see, *e.g.*, [23]). Foremost of these is the fact that the (floating-point) bound thus obtained is neither exact nor entirely rigorous. Flagmatic offers two ways around this difficulty.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

If the floating-point bound is not thought to be tight, then the simplest of the two ways is also the most appropriate: Flagmatic can perform a Cholesky decomposition of the matrices, and then round off each entry to the nearest rational, with denominators bounded by a suitable integer q ($q = 10^8$ is the default, if the user does not supply a preference). In this way, a rational bound on the Turán density $\pi(\mathcal{F})$ can be obtained rigorously. The said bound may appear to be slightly worse than the floating-point bound initially reported by Flagmatic, but in practice we may keep this discrepancy below 10^{-6} by choosing q large enough.

On the other hand, if the floating-point bound first reported by Flagmatic is thought to be tight, and if we know a matching lower-bound construction, then we can do better. Given a lower-bound construction, Flagmatic will use it to construct zero eigenvectors of the positive semi-definite matrices found by the SDP solver. This is done by using Remark 2.8.

So for each positive semi-definite matrix Q, assuming that all the zero eigenvectors can be obtained in this way, we can factor out the zero eigenspace and write Q as a product,

$$Q = R\,Q'\,R^T,$$

where Q' is positive definite. Moreover, because the R matrix can be constructed by considering, loosely speaking, 'flag densities in the limit of an extremal configuration', it can be constructed with rational entries. Flagmatic then rounds the entries of Q' to nearby rationals, its choices being guided in a few cases by the conjectured value of $\pi(\mathcal{F})$. (The rounding procedure used by Flagmatic is somewhat unsophisticated, but we have found it to be sufficient for our purposes. More complicated methods of rounding are possible: for example one could try to minimize the Euclidean distance between the original floating-point matrix and the rounded matrix, as proposed in Section 2.4.2 of [5].)

Since the floating-point matrix Q' is positive definite, the 'rounded off' matrix will also be positive definite, provided our approximation is sufficiently fine. (Indeed if the perturbation of the entries of Q' introduced in the rounding-off process is too great, Flagmatic will report an error and ask to use larger denominators q.) Finally, to ensure that it is beyond doubt that the 'rounded off' Q' is positive definite, Flagmatic uses a change of basis (via Gaussian elimination) to put it in diagonal form. (The R matrix is modified so that $Q = R\ Q'\ R^T$ is unchanged.)

Finally, Flagmatic will produce a certificate of the rigorous flag algebra bound (2.5), of which more will be said in the next subsection. For more information about using Flagmatic, we invite the reader to consult the User's Guide [42].

2.4. Certificates

One of the drawbacks of the flag algebra method is that computations rapidly become very involved. The number of distinct 3-graphs on n vertices, up to isomorphism, for $n = 1, 2, \dots$ grows very rapidly:

$$1,\ 1,\ 2,\ 5,\ 34,\ 2136,\ 7013320,\ \dots$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

(sequence A000665 of [2]), and the size of the family of admissible graphs increases at a comparable pace in most problems. In practical terms, this means that we cannot perform any flag algebra calculations with admissible graphs of order $m > 7$, and that even for $m = 6$ and $m = 7$, many flag algebra calculations involve too many graphs to be easily verifiable by hand.

Different authors have used different ways of addressing this issue: some [24, 25, 37] include lists of admissible graphs, intersection types, flags and large positive semi-definite matrices in the body of their papers; others [6, 7] worked with matrices that were too large and admissible graphs that were too numerous for this to be a practical solution, and omitted them from their papers. Our calculations by and large fall in the latter category, and we will similarly omit long lists of data.

Instead, we have used Flagmatic to produce certificates for all the flag algebra calculations we perform. These certificates are available on the Flagmatic website, `flagmatic.org`, as well as in the ancillary files associated with the arXiv version of this paper. The certificates are in the JSON format [1], which is designed to be human-readable. Let us give details of what they contain, and of how this may be used to verify our calculations.

Flagmatic uses the following notation for 3-graphs. First the order n is given, followed by a colon and a (possibly empty) list of 3-edges, given as a string of numbers $x_1y_1z_1x_2y_2z_2\dots$. For example, '`3:`' represents the 3-graph on 3-vertices with no edges, whilst '`4:123124134`' and '`4:213214234`' both represent K_4^-. (Note that this notation does not at present allow us to represent graphs on more than 10 vertices. However, this does not turn out to be much of a restriction, as 3-graphs of order 8 are already computationally intractable for our method.)

All numbers in the certificates are rational, and are either provided as fractions 'p/q', or as integers. Symmetric matrices are given by the entries in their upper triangle, so that

```
[[1,0,0], [1,0], [1]]
```

is the 3×3 identity matrix. Matrices that are not necessarily symmetric are given by their rows, with

```
[[1,-2],[-5,3]]
```

standing for the matrix

$$\begin{pmatrix} 1 & -2 \\ -5 & 3 \end{pmatrix}.$$

The certificates produced by Flagmatic contain the following information.

(1) A description of the problem, specifying which r-graphs we are working with (in all our applications, $r = 3$); what we are trying to maximize (in this paper, the density of 3-edges, referred to as '`3:123`' in the certificate); and which configurations we are forbidding.

(2) The bound obtained (a rational number).

(3) The order m of the admissible graphs we are working with; the number of admissible graphs of order m (up to isomorphism); and a list of the admissible graphs in the Flagmatic notation.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

(4) The number of intersection types used; and a list of the intersection types in the Flagmatic notation.
(5) A list of the number of flags for each intersection type (the first number in the list corresponding to the first intersection type listed, the second number corresponding to the second intersection type, and so on); and a list of the σ-flags for each type σ (in Flagmatic notation, ordered by type as above).
(6) A list of Q' matrices (called '`qdash_matrices`' in the certificate), one for each intersection type.
(7) A list of R matrices (called '`r_matrices`' in the certificate), one for each intersection type.

At this stage the reader may wonder why we are giving two matrices for each intersection type, rather than just one. Recall that for each intersection type σ we must provide a positive semi-definite matrix Q to use in inequality (2.3). To ensure that there can be no doubt as to the positive semi-definiteness of the matrices it provides, Flagmatic gives two matrices R and Q', where Q' is a positive definite *diagonal* matrix and R is a rectangular matrix. The matrix Q is then computed as

$$Q = R\,Q'\,R^T.$$

Given all this information, what does one need to do to verify that the flag algebra calculation is indeed correct? There are four stages.

(1) First of all, one needs to check that the family of admissible 3-graphs given in the certificate is indeed the family of *all* admissible 3-graphs of order m.
(2) For all admissible graphs H and all intersection types σ, one then needs to compute the densities $d(H)$ and the flag pair densities $d_{F,F'}(H)$ for each pair of σ-flags (F, F').
(3) Next, the Q matrices must be computed from the Q' and R matrices.
(4) Finally, one needs to substitute all these terms into inequality (2.5) and check that the claimed bound is achieved.

To assist with these tasks, Emil Vaughan provides a separate checker program, available from the Flagmatic website, called '`inspect_certificate.py`'. This program is independent of Flagmatic, and only requires Python 2.6 or 2.7 to run. Given a certificate as input, it can do any of the following:

- display the list of admissible graphs,
- display the types and flags,
- display the Q' and R matrices,
- compute and display the Q matrices,
- compute and display the admissible graph densities,
- compute and display the flag pair densities,
- compute and display the flag algebra coefficients for each admissible graph,
- compute and display which admissible graphs have a flag algebra coefficient equal to the bound.

As mentioned earlier, the certificates for our results are available on the Flagmatic website, and in the data set included in our arXiv submission. Each certificate has a unique file name, which is given in Table 2.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Table 2.

Result	Certificate filename(s)
Theorem 3.2	`k4-f32c5.js`
Theorem 3.3	`k4-f32.js`
Theorem 3.4	`38.js`
Theorem 3.5	`638.js`
Theorem 3.19	`k4-15.js` and `k4-f3215.js`
Theorem 3.23	`k58i.js`
Theorem 3.26	`43if32.js`
Theorem 3.27	`k4j4.js`
Proposition 4.2	`k4-c5.js`
Proposition 4.3	`k4-.js`, `c5.js` and `blm.js`

2.5. Stability

When we do get a tight bound from Flagmatic we in fact get a little more than just a proof of the Turán density: we have some information on the structure of extremal (hyper)graphs as well. This was observed by Baber [5] and Pikhurko [34].

Let $\mathcal{F}$ be a family of (hyper)graphs, and let $\rho > 0$. Suppose we have a construction of blow-up type showing $\pi(\mathcal{F}) \geqslant \rho$ and which we can show using the semi-definite method with admissible (hyper)graphs of order m that $\pi(\mathcal{F}) \leqslant \rho$. Then not only do we know that $\pi(\mathcal{F}) = \rho$, but we know which subgraphs can appear with strictly positive upper density in near-extremal graph sequences: by Lemma 2.7, for all admissible (hyper)graphs H of order m for which the flag algebra coefficient is not tight, and all $\mathcal{F}$-free graph sequences $(G_n)_{n\in\mathbb{N}}$ with $|V(G_n)| = n$ and edge density $d(G_n) = (\pi(\mathcal{F}) + o(1))$, we have $d_H(G_n) \to 0$.

Then let $\mathcal{F}'$ denote the collection of admissible graphs of order m whose flag algebra coefficient is not tight, and call the rest of the admissible graphs of order m the *sharp graphs*. Since $\mathcal{F}'$ is finite, it follows that in all near-extremal sequences $(G_n)_{n\in\mathbb{N}}$ only a $o(1)$ proportion of m-tuples of vertices induce a copy of a member of $\mathcal{F}'$.

We can thus use a (hyper)graph removal lemma and change a $o(1)$ proportion of the edges of $(G_n)_{n\in\mathbb{N}}$ to obtain a new sequence $(G'_n)_{n\in\mathbb{N}}$ which is $\mathcal{F}$-free *and* contains no induced copy of a member of $\mathcal{F}'$. (For graphs and directed graphs, we can use the removal lemmas of Alon and Shapira [3, 4]; for uniform hypergraphs, this can be done instead by the hypergraph removal lemma of Rödl and Schacht [39].)

In this new sequence, we now know exactly what the subgraphs of order m are: the sharp graphs. It is then usually a simple matter of fitting the pieces of the puzzle together to show that the modified sequence is of the appropriate blow-up type.

Let us give an example. We know from our proof of the density version of Mantel's theorem that the sharp graphs are H_0 and H_2, *i.e.*, the graphs on three vertices spanning 0 and 2 edges respectively. Suppose we have a triangle-free graph sequence $(G_n)_{n\in\mathbb{N}}$ with $e(G_n) = (1/2 + o(1))\binom{n}{2}$ edges. By applying the removal lemma of Alon and Shapira [4],

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

we can change $c_n = o(n^2)$ edges in G_n to obtain G'_n such that all 3-sets of G'_n induce a copy of H_0 or H_2. For n large enough, G'_n must contain some edge (since $1/2 - o(1) > 0$ for n large enough). Thus there must be at least one copy of H_2 in G'_n, say $([3], \{12, 13\})$.

Consider any vertex $x \in V(G'_n) \setminus [3]$. Since $13x$ spans at least one edge, it must span a copy of H_2, and similarly for $12x$. Now if $2x$ is an edge, then $23x$ spans a copy of H_2 and thus $3x$ must be an edge and $1x$ must be a non-edge. Then we may identify 1 and x. On the other hand if $1x$ is an edge then $2x, 3x$ are non-edges and we may identify 2 and x.

Proceeding in this way until we run out of vertices, and then identifying 2 and 3, we see that our graph G'_n is in fact a blow-up of the edge $([2], \{12\})$, in other words a complete bipartite graph. Since the edge density of G'_n is $1/2 + o(1)$, it follows that both parts have size $(1 + o(1))n/2$ and that G'_n can be made into a complete balanced bipartite graph by changing $c'_n = o(n^2)$ edges. Thus the original graph G_n is at an edit distance of $c_n + c'_n = o(n^2)$ from complete balanced bipartite. This shows that the Turán problem for the triangle is *stable*.

Stability is a very useful property, and can sometimes be exploited to compute Turán numbers exactly: this is known as the *stability method*. See [33, 41] for a detailed discussion of this technique and [34] for an example. While we shall not deal explicitly with Turán numbers or stability in this paper, let us mention the recent work of Baber and Talbot [7] on stability along the lines suggested above, and the upcoming note of the present authors [15], in which we show that the tight results obtained in this paper are stable.

3. Results

3.1. On the extremal theory of 3-graphs with independent neighbourhoods

A 3-graph G is said to have *independent neighbourhoods* if for every $x, y \in V(G)$ the *joint neighbourhood*

$$\Gamma(x, y) = \{z : xyz \in E(G)\}$$

of x and y is an edge-free set in G. This is equivalent to saying that G contains no copy of $F_{3,2}$ as a subgraph, where $F_{3,2} = ([5], \{123, 145, 245, 345\})$. For reasons we shall elaborate on in Section 4.1, the extremal theory of 3-graphs with independent neighbourhoods is very amenable to investigations via the semi-definite method.

The first result we should mention is due to Füredi, Pikhurko and Simonovits [22], who established the Turán density of $F_{3,2}$ (and in fact determined its Turán number $\mathrm{ex}(n, F_{3,2})$ exactly).

Theorem 3.1 (Füredi, Pikhurko and Simonovits).

$$\pi(F_{3,2}) = 4/9. \qquad \square$$

The next four results, however, are new.

Theorem 3.2.

$$\pi(K_4^-, C_5, F_{3,2}) = 12/49.$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

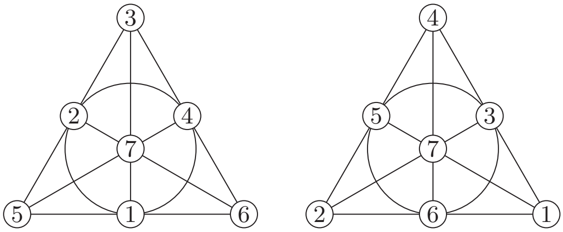

Figure 3. Füredi's double Fano construction.

Proof. The upper bound is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound, which was independently obtained by Füredi [29], comes from taking a balanced blow-up of the 6-regular 3-graph on 7 vertices:

$$H_7 = ([7], \{124, 137, 156, 235, 267, 346, 457, 653, 647, 621, 542, 517, 431, 327\}).$$

The 3-graph H_7 can be obtained as the union of two edge-disjoint copies of the Fano plane on the same vertex set,

$$F_1 = ([7], \{124, 137, 156, 235, 267, 346, 457\}),$$
$$F_2 = ([7], \{653, 647, 621, 542, 517, 431, 327\}),$$

as depicted in Figure 3. This elegant perspective is due to Füredi [29].

Another way to think about H_7 is by considering its link graphs: for every $i \in [7]$, the link graph of i in H_7 is a 6-cycle, which is triangle-free (in fact bipartite). This instantly shows that a blow-up of H_7 is K_4^--free. To see that H_7 and its blow-ups are $F_{3,2}$-free, it is enough to observe that for every $i \neq j$ in [7], the codegree of i and j in H_7 is exactly 2, which is not enough to support a 3-edge, so that their joint neighbourhood remains edge-free in the blow-up. Finally, to see that such a blow-up is C_5-free, note that H_7 is itself C_5-free, so that 5 vertices in distinct parts of the blow-up cannot span a C_5, while, on the other hand, a copy of C_5 in the blow-up cannot involve two vertices in the same part (since any two vertices of C_5 appear together in a 3-edge). □

The next result is similar to an earlier theorem of Frankl and Füredi [20], which we shall discuss in the next section, where we also show how our results differ.

Theorem 3.3.

$$\pi(K_4^-, F_{3,2}) = 5/18.$$

Proof. The upper bound is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound, due to Frankl and Füredi, is obtained by taking a balanced blow-up of the following 5-regular 3-graph on 6 vertices:

$$H_6 = ([6], \{123, 234, 345, 145, 125, 136, 356, 256, 246, 146\}).$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

There are two easy ways to visualize H_6. On the one hand, it is the unique 3-graph on 6 vertices such that for every $i \in [6]$ the link graph of i is a 5-cycle. Alternatively, we may think of it as the unique 3-graph on 6 vertices with all its 5-vertex subgraphs isomorphic to C_5. The first description makes it clear that blow-ups of H_6 are K_4^--free, since the link graphs of H_6 contain no triangles. A blow-up of C_5 clearly has independent neighbourhoods, and a copy of $F_{3,2}$ involves vertices in at most 5 different parts of a blow-up, so the second description establishes that blow-ups of H_6 are C_5-free as well. □

We now turn our attention to problems where we forbid certain 2-*graphs* from appearing in the links of a 3-graph. From this perspective, the previous theorem established that the edge density of a 3-graph with independent neighbourhoods and triangle-free links is at most $5/18 + o(1)$. In Theorems 3.4 and 3.5 we consider the problem of forbidding complete 2-graphs on 4 and 5 vertices respectively from appearing in the links instead of triangles.

Theorem 3.4.

$$\pi(J_4, F_{3,2}) = 3/8.$$

Proof. The upper bound is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound is obtained by taking a balanced blow-up H of K_4. For each vertex x in the resulting 3-graph, the link graph is the disjoint union of an independent set of vertices and a complete 3-partite graph; such a graph clearly cannot contain a complete graph on 4 vertices, establishing that H is J_4-free. To see that H is $F_{3,2}$-free as well, it is enough to note that a copy of $F_{3,2}$ cannot involve two vertices lying in the same part of H, and that H has only 4 parts whereas $F_{3,2}$ has 5 vertices. □

Theorem 3.5.

$$\pi(J_5, F_{3,2}) = 3/8.$$

Remark 3.6. Note this implies Theorem 3.4.

Proof. The upper bound is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound is obtained, as in Theorem 3.4, by taking a balanced blow-up H of K_4. Since J_4 is a subgraph of J_5, and H is J_4-free, H must be J_5-free as well. □

We should make two remarks here. First of all, the flag algebra calculation involved in the proof of Theorem 3.4 is 'easy' in comparison with the calculations involved in the proofs of Theorems 3.2 and 3.3. This, and the pleasing structure of our lower-bound construction, suggest that the underlying Turán density problem should be amenable to

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

more direct combinatorial arguments. Secondly, we might have expected that the extremal configuration for the $(J_5, F_{3,2})$ problem be a balanced blow-up of K_5, yielding link graphs consisting of complete 4-partite graphs together with an independent set. However, K_5 is not $F_{3,2}$-free, and as Theorem 3.5 shows, we do not gain anything from forbidding J_5 rather than J_4.

It seems natural to ask about the the behaviour of $\pi(J_t, F_{3,2})$ for large t.

Theorem 3.7 (Pikhurko [32]).

$$\lim_{t\to\infty} \pi(J_t, F_{3,2}) = \frac{4}{9}.$$

Proof. Certainly

$$\pi(J_t, F_{3,2}) \leqslant \pi(F_{3,2}) = \frac{4}{9}$$

for all $t \geqslant 1$. The idea for the lower-bound construction, due to Pikhurko [32], is to use the extremal construction for $F_{3,2}$ and to modify it slightly in order to make it J_t-free as well.

Take a bipartition of $[n]$ into two parts A and B, with $|A| \approx n/3$. Further divide B into $t-1$ equal parts $B_1, B_2, \ldots, B_{t-1}$. Then take as our 3-edges all triples meeting A and two distinct subparts of B to obtain a graph G. We claim that G is $(J_t, F_{3,2})$-free.

Indeed, vertices in our construction have $(t-1)$-partite link graphs (since there is no edge meeting one of $A, B_1, \ldots, B_{t-1}$ in exactly two vertices). Thus G is J_t-free. Also, given b, b' in B, their joint neighbourhood is the edge-free set A; given $a \in A$ and $b \in B$, their joint neighbourhood is a subset of the edge-free set B; and given $a, a' \in A$, their joint neighbourhood is empty. Thus G is $F_{3,2}$-free as claimed.

Now the edge density of G is

$$\begin{aligned} d(G) &= |A|\binom{t-1}{2}\left(\frac{n-|A|}{t-1} + O(1)\right)^2 \Big/ \binom{n}{3} \\ &= \frac{4}{9}\frac{(t-2)}{(t-1)} + O(n^{-1}). \end{aligned}$$

Thus, for any fixed t we have

$$\pi(J_t, F_{3,2}) \geqslant \frac{4}{9}\frac{(t-2)}{(t-1)},$$

which tends to $\pi(F_{3,2}) = 4/9$ as $t \to \infty$. □

In essence, as $t \to \infty$ the construction of Pikhurko given above 'converges' to the extremal configuration for $F_{3,2}$, namely the bipartite graph with $V = A \sqcup B$, $|A| \approx |B|/2$ and all 3-edges of type ABB. In the limit as $t \to \infty$, this is the best we can do. For t fixed, this is another matter. As we proved in Theorems 3.3, 3.4 and 3.5, this construction is not optimal for $t = 3, 4, 5$. In fact, as

$$\frac{4}{9}\frac{(t-2)}{(t-1)} \leqslant \frac{3}{8}$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

for all $t < 8$, we know that Pikhurko's construction cannot be optimal for $t < 8$.

Question 3.8. Is it the case that for all $t : 4 \leqslant t < 8$

$$\pi(J_t, F_{3,2}) = \frac{3}{8}?$$

Is it the case that for all $t \geqslant 8$

$$\pi(J_t, F_{3,2}) = \frac{4}{9}\frac{(t-2)}{(t-1)}?$$

Let us finally note that until very recently all previous known results in extremal 3-graph theory had one of five extremal configurations: the blow-up of a 3-edge [9, 19], H_6 [20], the 'one-way' complete bipartite 3-graph [22] (an unbalanced blow-up of the degenerate 3-graph $([2], \{112\})$), Turán's construction [37] (where the proof also relied on the semi-definite method) and the complete bipartite 3-graph [5, 13, 31, 27]. We can now add two more extremal configurations to this list: the balanced blow-up of H_7 and the balanced blow-up of K_4.

Since the first version of this paper was written, Baber and Talbot [7] have added seven more examples using exhaustive computer search and the semi-definite method. Also Pikhurko [35] has showed that for *every* blow-up or iterated blow-up construction with optimized weights, there exists some finite family of 3-graphs for which the construction is extremal. (His proof relies on a kind of compactness argument, however, and so does not yield explicit families.)

We now come to some Turán problems for which we have been unable to find tight bounds using Flagmatic. Erdős and Sós conjectured that the maximal density of a 3-graph in which all vertices have a bipartite link graph is 1/4.

Conjecture 3.9 (Erdős and Sós: see [20]).

$$\pi(\text{odd cycle in link graph}) = 1/4.$$

If the conjecture is true, then this is an extremely unstable problem. Two different constructions were given by Frankl and Füredi [20].

Construction 3.10 (Frankl and Füredi). Distribute n vertices uniformly along the circumference of a circle. Then define a 3-graph on n vertices by putting a 3-edge xyz in the graph if the centre of the circle lies in the interior of the triangle determined by x, y and z, to obtain a K_4^--free 3-graph.

Construction 3.11 (Frankl and Füredi). Consider a random tournament T on n vertices. Then define a 3-graph on n vertices by putting a 3-edge xyz in the graph if xyz is an oriented triangle in T.

To these constructions, we can add five more.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Construction 3.12. Take a balanced, iterated blow-up of the 3-graph consisting of a single 3-edge, $G = ([3], \{123\})$.

Construction 3.13. Take a balanced iterated blow-up of C_5.

Construction 3.14. Take a balanced iterated blow-up of H_7.

Construction 3.15. Take a balanced iterated blow-up of

$$([7], \{123, 124, 125, 136, 137, 146, 247, 256, 257, 347, 356, 357, 456, 467\}).$$

Construction 3.16. Take a balanced iterated blow-up of

$$([7], \{123, 124, 125, 136, 146, 157, 237, 247, 256, 345, 356, 367, 457, 467\}).$$

The last three constructions are all iterated blow-ups of some 6-regular 3-graph on 7 vertices. The best way to think about them is perhaps in terms of their link graphs: the link graphs in H_7 consist of 6-cycles, whereas the links in Constructions 3.15 and 3.16 are isomorphic to (respectively) a 4-cycle with two pendant edges attached to a pair of adjacent vertices, and a 4-cycle with a path of length 2 attached to one of the vertices. In all three cases, the links are bipartite, and so the links in an iterated blow-up are bipartite as well.

In fact, more generally, if G is a 3-graph with bipartite links, then any iterated blow-up of G also has bipartite links. We can thus construct arbitrarily many non-isomorphic configurations of 3-graphs with bipartite links and 3-edge density $1/4 + o(1)$ by taking any of the above constructions, blowing it up, and then inside each of the parts, we are free to place a copy of any of the other constructions.

Given this instability, the bipartite links conjecture of Erdős and Sós appears very hard. We believe, however, that the independent neighbourhoods version of the problem should be stable with Construction 3.10 being the essentially unique extremal configuration.

Conjecture 3.17.

$$\pi(\text{odd cycle in link graph}, F_{3,2}) = 1/4,$$

with the stable extremal configuration being given by Construction 3.10.

In fact, more generally, we believe that extremal problems for 3-graphs with independent neighbourhoods should be stable.

Conjecture 3.18. Turán problems for 3-graphs with independent neighbourhoods are stable.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

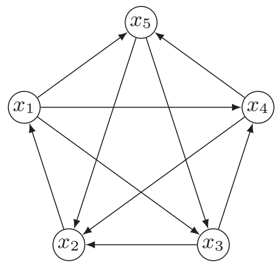

Figure 4. An orientation of $K_5^{(2)}$.

As we shall see in Section 3.3, however, the extremal theory of 3-graphs with independent neighbourhoods still has non-principality: there exist 3-graphs H_1 and H_2 such that

$$\pi(H_1, H_2, F_{3,2}) < \min(\pi(H_1, F_{3,2}),\ \pi(H_2, F_{3,2})).$$

Thus even in this restricted setting we cannot hope for an analogue of the Erdős–Stone theorem from extremal graph theory.

Before we close this section, let us note the bounds we can obtain using Flagmatic for the problems in Conjectures 3.9 and 3.17.

Theorem 3.19.

$$1/4 \leqslant \pi(\text{odd cycle in link graph}, F_{3,2}) < 0.255889,$$
$$1/4 \leqslant \pi(\text{odd cycle in link graph}) < 0.258295.$$

Proof. The upper bounds are from two flag algebra calculations using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bounds are from Construction 3.10. □

Let us finally outline a proof of our claim that Constructions 3.10–3.16 are distinct. (That they have asymptotic density 1/4 and bipartite links is left as an exercise for the reader.)

Constructions 3.12–3.16 can be distinguished by considering their link graphs; they are moreover highly structured, so that with high probability, the random Construction 3.11 cannot be edited into them without changing at least a constant proportion of the 3-edges. (Indeed the probability of, say, $n/3$ vertices having identical neighbourhoods (up to $o(n^3)$ edges) in the rest of the 3-graph is exceedingly small.)

Clearly iterated blow-up constructions are not $F_{3,2}$-free. It is easy to see that Construction 3.11 is not $F_{3,2}$-free either: given 5 vertices x_1, x_2, x_3, x_4, x_5, the orientation

$$\overrightarrow{x_2x_1}, \overrightarrow{x_1x_3}, \overrightarrow{x_3x_2}, \overrightarrow{x_1x_4}, \overrightarrow{x_4x_2}, \overrightarrow{x_1x_5}, \overrightarrow{x_5x_2}, \overrightarrow{x_3x_4}, \overrightarrow{x_4x_5}, \overrightarrow{x_5x_3}$$

(see Figure 4) occurs with probability at least 2^{-10} in a random tournament, so that we expect $F_{3,2}$ to occur as a subgraph in Construction 3.11 with strictly positive density.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

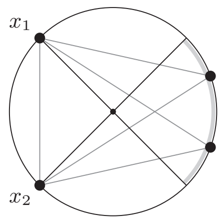

Figure 5. The circle construction has independent neighbourhoods.

Now on the other hand, Construction 3.10 *is* $F_{3,2}$-free. Indeed, consider any two vertices x_1 and x_2 on the circumference of a circle, and let us show that their common neighbourhood is an independent set. If x_1 and x_2 lie on the same diameter, their codegree must be zero, as x_1 and x_2 cannot be vertices of a triangle that has the centre of the circle in its interior. Thus we may assume, without loss of generality, that x_1 and x_2 do not lie on the same diameter. Then the diameters through x_1 and x_2 separate the circumference of the circle into four arcs (see Figure 5). By construction, the common neighbourhood of x_1 and x_2 consists of all the vertices lying on the interior of the arc that contains neither x_1 nor x_2. But by construction this is an independent set of vertices. Thus Construction 3.10 is distinct from all our other constructions.

3.2. Forbidding induced subgraphs

Turán's conjecture is arguably the most famous open problem in extremal combinatorics.

Conjecture 3.20 (Turán).

$$\pi(K_4) = 5/9.$$

Turán's original construction for the lower bound that motivates his conjecture is obtained by taking a balanced tripartition $A \sqcup B \sqcup C$ of the vertex set, and putting in all 3-edges of type AAB, BBC, CCA and ABC. (In our language, this is a blow-up of the degenerate 3-graph $([3], \{123, 112, 223, 331\})$.) Many other other constructions for the problem have since been found. Indeed there are exponentially many non-isomorphic 3-graph configurations on n vertices attaining the bound given by Turán's construction while not containing any copy of K_4: see Brown [12], Kostochka [28], Fon-der-Flaas [18] and Frohmader [21]. If Turán's conjecture is true, the Turán density problem for K_4 is therefore very unstable and thus (for reasons we shall develop in Section 4) unlikely to be resolved by a direct application of the semi-definite method.

Razborov observed, however, that Turán's construction is the only one in which no 4-set of vertices spans exactly one 3-edge. Adding this restriction, he was able to give a proof of a weakening of Turán's conjecture using the semi-definite method. Formally, let us call G_1 the unique (up to isomorphism) 3-graph on 4 vertices with exactly one 3-edge. Then the following holds.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Theorem 3.21 (Razborov [37]).

$$\pi(K_4,\ \text{induced } G_1) = 5/9. \qquad \square$$

Thus in this case Razborov was able to circumvent the instability of the K_4 problem to obtain his result.

Related work on Turán's conjecture along these lines can be found in [38, 7]. Mention should also be made of the work of Pikhurko [34], who used Razborov's flag algebra computation and combinatorial arguments to determine the Turán number $\mathrm{ex}(n, \{K_4,\ \text{induced } G_1\})$ exactly.

Proceeding similarly to Razborov, we considered the following conjecture, which is also attributed to Turán.

Conjecture 3.22 (Turán).

$$\pi(K_5) = 3/4.$$

As in Conjecture 3.20, more than one extremal configuration attaining the conjectured bound is known. One K_5-free 3-graph with density $3/4 + o(1)$ is obtained by taking a complete balanced bipartite 3-graph. Another example, due to Keevash and Mubayi [26], is obtained by taking a balanced blow-up of K_4 and, writing $A \sqcup B \sqcup C \sqcup D$ for the corresponding 4-partition of the vertex sets, adding all 3-edges of type AAC, AAD, BBD, BBA, CCA, CCB, DDB and DDC. In our notation, this is the complement of the blow-up of the degenerate 3-graph

$$([4], \{111, 222, 333, 444, 112, 223, 334, 441\}).$$

This is easily seen to be distinct from the first example. Many more configurations exist: Sidorenko exhibited infinite families of non-isomorphic K_5-free constructions with asymptotic density $3/4$ (see Constructions 4–7 in [40]). Thus, if Conjecture 3.22 is true, then the Turán density problem for K_5 is very unstable and, just as in Conjecture 3.20, we are unlikely to arrive at tight bounds for $\pi(K_5)$ by using only the semi-definite method.

We are, however, able to obtain an analogue of Razborov's result: observe that in a complete bipartite graph, a 5-set of vertices cannot span exactly 8 edges. On the other hand, consider for example the construction of Keevash and Mubayi: taking one vertex from part A and two vertices from each of parts B and C yields a 5-set spanning exactly 8 edges.

Let us therefore write $\mathcal{G}$ for the family of 3-graphs on 5 vertices with exactly 8 edges. (There are only two such 3-graphs up to isomorphism; considering K_5 as a graph on the vertex set [5], these are $K_5 \setminus \{123, 145\}$ and $K_5 \setminus \{123, 124\}$ respectively.) Then the following holds.

Theorem 3.23.

$$\pi(K_5,\ \text{induced copy of a member of } \mathcal{G}) = 3/4.$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Proof. The upper bound is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound is from consideration of a complete balanced bipartite 3-graph. □

Just as for Theorem 3.21, it would be nice to have a more direct, combinatorial proof of Theorem 3.23; the proof above does not give much insight into the problem.

The strategy of introducing extra restrictions that we know must be satisfied by our desired extremal configuration in order to obtain a better bound is not new. An earlier result of a similar flavour (but proved without resorting to the semi-definite method) is the following theorem of Frankl and Füredi [20].

Theorem 3.24 (Frankl and Füredi [20]).

$$\pi(K_4^-, \text{ induced } G_1) = 5/18.$$ □

In fact Frankl and Füredi showed rather more: they determined the Turán number for this problem and showed that the unique extremal graph is a balanced blow-up of H_6. Even more, they proved that all K_4^--free 3-graphs with no induced copy of G_1 are either (possibly unbalanced) blow-ups of H_6 or are of the form given by Construction 3.10 in the previous subsection.

Observe that the density version of Frankl and Füredi's result which we stated above is very similar to Theorem 3.3. Indeed, the two results share the same lower-bound construction. Let us observe that forbidding a 3-graph from containing a copy of K_4^- or $F_{3,2}$ is strictly weaker than forbidding a 3-graph from containing a copy of K_4^- or an induced copy of G_1 (which is equivalent to requiring that all 4-sets span exactly 0 or 2 edges). Theorem 3.3 is thus a nominally stronger result than Theorem 3.24.

Lemma 3.25. *Suppose G is a 3-graph in which 4-sets span exactly 0 or 2 edges. Then G is $(K_4^-, F_{3,2})$-free. The converse is false.*

Proof. Let G be a 3-graph in which 4-sets span exactly 0 or 2 edges. Then G is trivially K_4^--free. Suppose it contained $F_{3,2}$ as a subgraph. By relabelling vertices, we have that G contains 5 vertices a, b, c, d, e such that abc, ade, bde, cde are all edges of G. Now the 4-set $\{a, b, d, e\}$ already spans 2 edges, so it cannot span any more. Thus neither of abd, abe lies in $E(G)$. Similarly, none of acd, ace and bcd, bce can lie in $E(G)$. Now consider the 4-set $abcd$. This spans exactly one edge, the other three having been forbidden; but this contradicts the fact that G is a 3-graph in which 4-sets span exactly 0 or 2 edges.

To see that the converse is false, consider a 3-graph on 4 vertices with 1 edge. This is obviously $(K_4^-, F_{3,2})$-free but violates the condition that 4-sets span exactly 0 or 2 edges. The same is true of any of its blow-ups. □

Finally, let us stress just how different forbidding induced subgraphs is to forbidding subgraphs. We have shown that $\pi(K_4^-, F_{3,2}) = 5/18$. In marked contrast is the following.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Theorem 3.26.

$$\pi(\text{induced } K_4^-,\ F_{3,2}) = 3/8.$$

Proof. The upper bound is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound is from consideration of a balanced blow-up of K_4. □

Note that Theorems 3.26 and 3.4 are implied by Theorem 3.5 and the observation that a blow-up of K_4 is $F_{3,2}$-free, J_4-free and contains no induced K_4^-. Indeed, suppose an $F_{3,2}$-free 3-graph G contains a copy of J_5. This consists of a 5-set S together with a vertex $x \notin S$ and all $|S^{(2)}|$ possible 3-edges containing x and two vertices from S. Since G is $F_{3,2}$-free, it must also be K_5-free, and hence at least one 3-edge $e = \{abc\}$ from $S^{(3)}$ is missing in G. The 4-set $\{xabc\}$ then spans an induced copy of K_4^- in G.

3.3. Non-principal pairs

By Theorem 3.3, $\pi(K_4^-, F_{3,2}) = 5/18$. On the other hand, Frankl and Füredi gave a lower bound of $2/7$ for $\pi(K_4^-)$ by considering a balanced iterated blow-up of H_6 [20], while Füredi, Pikurkho and Simonovits [22] showed $\pi(F_{3,2}) = 4/9$. Gathering all this together we have

$$\pi(K_4^-, F_{3,2}) = \frac{5}{18} < \min(\pi(K_4^-), \pi(F_{3,2})).$$

Thus $(K_4^-, F_{3,2})$ is an example of a *non-principal* pair of 3-graphs – that is to say, a pair F, F' with $\pi(F, F') < \min(\pi(F), \pi(F'))$. Non-principality for 3-graphs was conjectured by Mubayi and Rödl [31] and first exhibited by Balogh [8]. Mubayi and Pikhurko [30] then built on Balogh's ideas to give the first example of a non-principal pair of 3-graphs, and Razborov [37] used the semi-definite method to show that (K_4^-, C_5) is also a non-principal pair. We can exhibit yet another non-principal pair of 3-graphs.

Theorem 3.27.

$$\pi(K_4, J_4) < 0.479371 < 1/2 \leqslant \pi(J_4).$$

Proof. The upper bound on $\pi(K_4, J_4)$ is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound for $\pi(J_4)$, due to Bollobás, Leader and Malvenuto [10], is a balanced iterated blow-up of the complement of the Fano plane. □

Given that $\pi(K_4) \geqslant 5/9$, it follows that (K_4, J_4) is a fourth non-principal pair of 3-graphs. (It is in fact very similar to the example given by Mubayi and Pikhurko [30], who showed that (K_4, J_5) is a non-principal pair.) Note that we can show $\pi(K_4, J_4) \geqslant 2/5$ by considering an iterated blow-up of

$$([6], \{123, 124, 125, 134, 135, 146, 156, 236, 245, 246, 256, 345, 346, 356\}),$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

but 2/5 is quite far from the upper bound.

Question 3.28. What is $\pi(K_4, J_4)$?

Finally, let us remark that the extremal theory of 3-graphs with independent neighbourhoods also exhibits non-principality: by Theorems 3.2 and 3.3,

$$\pi(K_4^-, C_5, F_{3,2}) = 12/49 < \pi(K_4^-, F_{3,2}) = 5/18 < \pi(C_5, F_{3,2}) = 4/9,$$

where in the last line we have used the fact that $\pi(C_5, F_{3,2}) = \pi(F_{3,2})$ (which holds since the extremal configuration for $F_{3,2}$ is C_5-free.) Thus even in the case of 3-graphs with independent neighbourhoods we cannot hope for some analogue of the Erdős–Stone theorem from extremal graph theory.

Non-principality is in general hard to prove by hand. It can, however, be a useful tool when attacking Turán density problems. A common strategy when studying $\pi(\mathcal{F})$ for some family $\mathcal{F}$ is to try showing that $\pi(\mathcal{F}, G)$ is less than the conjectured valued of $\pi(\mathcal{F})$ for some nice, dense 3-graph G, and then use the presence of a (large) number of copies of G in a putative $\mathcal{F}$-extremal example to bound the edge density. (See for example [13] for a nice example of this technique.) Usually, provided that $\pi(\mathcal{F}, G) < \pi(\mathcal{F}) \leqslant \pi(G)$ is actually true, that we have a (conjectured) extremal $\mathcal{F}$-free construction, and that G and the graphs in $\mathcal{F}$ are not too large, Flagmatic can be expected to show non-principality holds.

4. Some additional remarks

4.1. The complexity barrier

We have already remarked in Section 2.2 that the semi-definite method cannot at present hope to give exact Turán density results for 3-graphs on 7 or more vertices. In this subsection, we shall consider some problems for small 3-graphs that we believe are still intractable, at least using the flag algebra method.

In contrast to the situation for graphs, we do not expect stability in general in extremal 3-graph theory. Indeed, we saw in Section 3.2 that if the conjectures of Turán and Sós are true then the Turán problems for K_4 and K_5 are unstable. In fact generally the K_t problem is conjectured to be unstable, non-isomorphic families of constructions having been given by Keevash and Mubayi [26]. We mentioned another example of conjectured instability in Section 3.1 when we considered the Erdős–Sós conjecture on odd cycles in link graphs and added new constructions to the two given by Frankl and Füredi [20].

Whatever the method used, unstable problems tend of course to be more difficult to handle than stable ones, and the semi-definite method is no exception to this trend. The bounds yielded by Flagmatic on the three problems mentioned above are

$$5/9 \leqslant \pi(K_4) < 0.561666,$$
$$3/4 \leqslant \pi(K_5) < 0.769533,$$
$$1/4 \leqslant \pi(\text{odd cycles in link graph}) < 0.258295$$

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

respectively, and we do not believe that these can be made tight even by an increase in computational firepower. A heuristic justification for our pessimism is as follows. The semi-definite method obtains bounds by considering how flags can intersect with each other; this information is then used to give inequalities which must be satisfied by the admissible subgraph densities. In an unstable problem, however, several very different global intersection structures are possible, and what is a correct, sharp subgraph density inequality in one structure may well be false in another. Indeed some admissible subgraphs may be present in one extremal configuration with strictly positive density, but absent in another. As remarked in Section 2.2, a hypothetical tight flag algebra bound would have to be tight on all such subgraphs simultaneously; this seems a rather unlikely situation to hope for. In this sense, unstable problems appear to be beyond the scope of the semi-definite method at present.

Another hurdle we have to face is that of stable problems with 'complex' extremal configurations. Let us define more precisely what we mean by this. Recall the definition of *blow-up* and *iterated blow-up* introduced in Section 2.1. Currently all known stable extremal configurations for 3-graphs consist of blow-ups of some (possibly degenerate) 3-graphs. Frankl and Füredi, however, gave an iterated blow-up construction for the K_4^- problem which is conjectured to be best possible. Since Frankl and Füredi's paper, Mubayi and Rödl [31] (for the C_5 problem) and Bollobás, Leader and Malvenuto [10] (for the J_4 problem) have both given us instances of the Turán density problem where an iterated blow-up construction is conjectured to be best possible. To these let us now add a fourth.

Conjecture 4.1.

$$\pi(K_4^-, C_5) = 1/4.$$

The lower bound in Conjecture 4.1 is attained for example by a balanced iterated blow-up of the 3-edge $([3], \{123\})$, or by a balanced iterated blow-up of H_7. To give motivation for our conjecture, let us note that we can get the following bounds on $\pi(K_4^-, C_5)$.

Proposition 4.2.

$$1/4 \leqslant \pi(K_4^-, C_5) < 0.251073.$$

Proof. The upper bound is from a flag algebra calculation using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bound is from an iterated blow-up of the 3-edge: this has bipartite links, and is thus K_4^--free. Moreover, 5-sets of vertices are easily seen to span $0, 1, 2, 3$ or 4 edges, which is not sufficient for a copy of C_5 to appear as a subgraph. □

Deferring our discussion of the limits of the semi-definite method for the moment, let us state why one should reasonably expect iterated blow-up constructions to be the best

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

possible for the $K_4^- = J_3$ and the J_4 problem, or indeed for the J_t problem in general. (Why it should crop up in problems involving C_5 seems a little more mysterious.)

Suppose we have a non-degenerate 3-graph H on l vertices which is J_t-free. Then any iterated blow-up of H will be J_t-free. Indeed, let G be an iterated blow-up of H. Let $x \in V(G)$ and let us show that its link graph is $K_t^{(2)}$-free. Consider a t-set of vertices $\{a_1, a_2, \dots, a_t\}$ in G_v. If all of the a_i lie in the same level 1 part of G as v, we can drop down to a lower level of the iterated construction, so we may assume without loss of generality that $v \in A_0$ and $a_1 \in A_1$, where A_0, A_1 are two distinct level 1 parts. As H was non-degenerate, there are no edges of type $A_0A_0A_0$, $A_1A_1A_1$, $A_0A_0A_1$ or $A_0A_1A_1$ in G. Thus, for the purpose of finding a copy of $K_t^{(2)}$ in G_v, we may assume that $v, a_1, a_2, \dots, a_t$ all lie in different level 1 parts $A_0, A_1, A_2, \dots, A_t$ of G. But then the subgraph of G induced by $v, a_1, a_2, \dots, a_t$ is isomorphic to a subgraph of H, which by hypothesis has $K_t^{(2)}$-free link graphs. Thus G has $K_t^{(2)}$-free link graphs and is J_t-free as claimed. It follows from this that for the J_t problem non-iterated blow-up constructions cannot be best possible. (Note that blowing up a 3-graph containing a degenerate edge trivially gives a copy of J_t, so our argument above does indeed cover all possible cases.)

Iterated blow-up constructions are therefore far from pathological, and one should expect them to crop up frequently in extremal 3-graph theory. Their structure is, however, much harder to grasp than that of their blow-up relatives. For example, the blow-up of a 3-graph H (with no degenerate edge of the form vvv) will always be $|V(H)|$-partite. In contrast, for any $N \in \mathbb{N}$ sufficiently large (non-trivial) iterated blow-ups will fail to be N-partite: the level 1 edges force at least two parts; then looking into one of the parts, the level 2 edges force at least one more part; then looking into one of the subparts, the level 3 edges force at least one more part, and so on. This is one reason we would not expect the structure of iterated blow-up configurations to be properly captured by the flag algebra calculus.

Proposition 4.3.

$$2/7 \leqslant \pi(K_4-) \leqslant 0.286889,$$
$$2\sqrt{3} - 3 \leqslant \pi(C_5) \leqslant 0.468287,$$
$$1/2 \leqslant \pi(J_4) \leqslant 0.504081.$$

Proof. The upper bounds are from three flag algebra calculations using Flagmatic (see Section 2.4 for how to obtain a certificate). The lower bounds are from (respectively) a balanced iterated blow-up of H_6 [20], a blow-up of $([2], \{112\})$ with $|A_1| \approx \sqrt{3}|A_2|$ and the construction iterated inside A_2 [31], and a balanced iterated blow-up of the complement of the Fano plane [10]. □

We do not believe that the above three upper bounds can be made tight by the semi-definite method, and similarly we do not expect Conjecture 4.1 to be resolved in this way either.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Let us give here some heuristic justification for our pessimism regarding these bounds, beyond the mere fact that they fail to be tight. Given a non-trivial graph H on t vertices and an integer k, the number of non-isomorphic subgraphs of order k with strictly positive density in large blow-ups of H will grow polynomially in k.

By contrast, the number of non-isomorphic subgraphs of order k found in a non-trivial iterated blow-up of H will typically be superpolynomial.

Let us give an example. Consider a blow-up of $([2], \{112\})$. This has $k-1$ distinct, non-isomorphic subgraphs of order k. (There are $k+1$ choices possible for the number of vertices in part 1, but placing $0, 1$ or k vertices in part 1 yields the same subgraph, namely the empty graph on k vertices.) On the other hand, write $f(k)$ for the number of non-isomorphic subgraphs of order k in a blow-up of $([2], \{112\})$ iterated inside part 2. Then, by considering the number of vertices in part 1, it is easy to see that

$$f(k) \geqslant \sum_{i=2}^{k-1} f(k-i). \tag{4.1}$$

Now we know from the subgraph count in the blow-up case that $f(k)$ grows at linear rate at least. The estimate (4.1) then implies that $f(k)$ in fact grows faster than any polynomial.

This superpolynomial growth rate is an objective measure of the fact that iterated blow-ups are significantly more 'complex' as 3-graph configurations than blow-up constructions. Computationally speaking, it is very bad news for an approach based on the flag algebra calculus. As we remarked in Section 2.1, if we obtain the correct upper bound on a Turán density problem, the flag algebra bound must be tight on all subgraphs which appear with strictly positive density in an extremal construction, whereas some slack is expected for the rest of the admissible subgraphs. In this sense iterated blow-up constructions require us to prove far more delicate inequalities than mere blow-up constructions: the far richer subgraph structure of iterated blow-ups leaving us with much less room to spare in our optimization, making our task significantly harder. We therefore expect that most attempts to attack problems admitting iterated blow-ups as extremal constructions with Flagmatic will run into the limits set by the SDP solver and fail to get tight bounds.

4.2. Further questions

The most obvious challenge left open by our discussion above is as follows. Say that a Turán problem is *simple* if the number of subgraphs of order k which can occur with density bounded below by some $\varepsilon > 0$ in an extremal configuration grows polynomially in k, and that a Turán problem is *complex* otherwise.

Question 4.4. Can we obtain an exact Turán density result for a complex problem? More precisely, can we give an explicit example of a finite family of 3-graphs $\mathcal{F}$ for which we can prove that the extremal configurations are complex?

Pikhurko [35] has recently shown that for all iterated blow-up configurations, there exists some finite family $\mathcal{F}$ of 3-graphs for which the said configuration is extremal (with suitably optimized weights placed on the different parts). However, his proof relies on a compactness argument and so does not give explicit families.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Table 3.

Forbidden graphs	Lower bound for π	Upper bound for π	(Conjectured) extremal configuration(s)
K_4^-, C_5, $F_{3,2}$	12/49	12/49	Blow-up of H_7.
K_4^-, C_5	1/4	0.251073	Iterated blow-up of a 3-edge.
$F_{3,2}$, odd cycle in links	1/4	0.255886	Geometric [20]; see Construction 3.10.
odd cycle in links	1/4	0.258295	Many: see Section 3.1.
K_4^-, $F_{3,2}$	5/18	5/18	Blow-up of H_6 [20].
K_4^-	2/7	0.286889	Iterated blow-up of H_6 [20].
J_4, $F_{3,2}$	3/8	3/8	Blow-up of K_4.
J_5, $F_{3,2}$	3/8	3/8	Blow-up of K_4.
$F_{3,2}$, induced K_4^-	3/8	3/8	Blow-up of K_4.
$F_{3,2}$	4/9	4/9 [22]	Bipartition of the vertex set into two parts A and B with $\lvert A\rvert \approx 2\lvert B\rvert$, all edges of type AAB [22].
J_4, K_4	2/5	0.479371	
J_4	1/2	0.504081	Iterated blow-up of the complement of the Fano plane [10].
C_5	$2\sqrt{3}-3$	0.468287	Bipartition of the vertex set into two parts A and B with $\lvert A\rvert \approx \sqrt{3}\lvert B\rvert$, all edges of type AAB, then iterate inside B [31].
K_4, induced G_1	5/9	5/9 [37]	Turán's construction.
K_4	5/9	0.561666 [37]	Many: see [12, 18, 21, 28].
K_5, 5-set spanning 8 edges	3/4	3/4	Complete bipartite graph.
K_5	3/4	0.769533 [5]	Many: see [40].

In Section 3.2 we proved a number of results in the extremal theory of 3-graphs with independent neighbourhoods. As the extremal construction for $F_{3,2}$ is K_4-free, it is easy to see that $\pi(K_t, F_{3,2}) = 4/9$ for all $t \geqslant 4$. Having considered both the J_t (complete graphs in links) and the odd cycle in links problem, the most natural question to ask next is perhaps: What happens if, instead of forbidding all odd cycles, we only forbid odd cycles of a given length in the link graphs? For example, we have the following two questions.

Question 4.5. Is $\pi(F_{3,2},$ odd cycle of length at least 5 in link$) = 1/4$?

Question 4.6. Is $\pi(F_{3,2},$ odd cycle of length at most 5 in link$) = 1/4$?

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

Note that if a vertex in a 3-graph G has a triangle in its link graph, then for any odd length $l \geqslant 3$, sufficiently large blow-ups of G will have link graphs containing odd cycles of length l; were it not for the nature of Construction 3.10, this would suggest that the answer to Question 4.5 is 'yes'. Also, Theorem 3.3 tells us that the answer to Question 4.6 is 'no' if we replace 5 by 3 (since $\pi(K_4^-, F_{3,2}) = 5/18$), making the question more open-ended than suggested by the upper bounds we are able to obtain for the problem using Flagmatic.

4.3. Summary of results and constructions

In Table 3 we set out the constructions and Flagmatic bounds for the Turán density problems discussed in the paper.

References

[1] JSON standard. http://tools.ietf.org/html/rfc4627.
[2] The on-line encyclopedia of integer sequences. http://oeis.org.
[3] Alon, N. and Shapira, A. (2003) Testing subgraphs in directed graphs. In *Proc. 35th Annual ACM Symposium on Theory of Computing*, ACM, pp. 700–709.
[4] Alon, N. and Shapira, A. (2005) A characterization of the (natural) graph properties testable with one-sided error. In *46th Annual IEEE Symposium on Foundations of Computer Science, 2005*, IEEE, pp. 429–438.
[5] Baber, R. (2011) Some results in extremal combinatorics. PhD thesis, University College London.
[6] Baber, R. and Talbot, J. (2011) Hypergraphs do jump. *Combin. Probab. Comput.* **20** 161–171.
[7] Baber, R. and Talbot, J. (2012) New Turán densities for 3-graphs. *Electron. J. Combin.* **19** #19.
[8] Balogh, J. (2002) The Turán density of triple systems is not principal. *J. Combin. Theory Ser. A* **100** 176–180.
[9] Bollobás, B. (1974) Three-graphs without two triples whose symmetric difference is contained in a third. *Discrete Math.* **8** 21–24.
[10] Bollobás, B., Leader, I. and Malvenuto, C. (2011) Daisies and other Turán problems. *Combin. Probab. Comput.* **20** 743–747.
[11] Borchers, B. (1999) CSDP: A library for semidefinite programming. *Optim. Methods Software* **11** 613–623.
[12] Brown, W. G. (1983) On an open problem of Paul Turán concerning 3-graphs. In *Studies in Pure Mathematics: To the Memory of Paul Turán*, Birkhäuser, pp. 91–93.
[13] de Caen, D. and Füredi, Z. (2000) The maximum size of 3-uniform hypergraphs not containing a Fano plane. *J. Combin. Theory Ser. B* **78** 274–279.
[14] Erdős, P. and Stone, A. H. (1946) On the structure of linear graphs. *Bull. Amer. Math. Soc* **52** 1087–1091.
[15] Falgas-Ravry, V. and Vaughan, E. R. A note on stability and the semi-definite method. Preprint.
[16] Falgas-Ravry, V. and Vaughan, E. R. (2011) On applications of Razborov's flag algebra calculus to extremal 3-graph theory. arXiv:1110.1623
[17] Falgas-Ravry, V. and Vaughan, E. R. (2012) Turán H-densities for 3-graphs. *Electron. J. Combin.* **19** #40.
[18] Fon-Der-Flaass, D. G. (1988) Method for construction of (3, 4)-graphs. *Math. Notes* **44** 781–783.
[19] Frankl, P. and Füredi, Z. (1983) A new generalization of the Erdős–Ko–Rado theorem. *Combinatorica* **3** 341–349.
[20] Frankl, P. and Füredi, Z. (1984) An exact result for 3-graphs. *Discrete Math.* **50** 323–328.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press

[21] Frohmader, A. (2008) More constructions for Turán's (3, 4)-conjecture. *Electron. J. Combin.* **15** R137.

[22] Füredi, Z., Pikhurko, O. and Simonovits, M. (2005) On triple systems with independent neighbourhoods. *Combin. Probab. Comput.* **14** 795–813.

[23] Goldberg, D. (1991) What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys* (CSUR) **23** 5–48.

[24] Grzesik, A. (2012) On the maximum number of C_5's in a triangle-free graph. *J. Combin. Theory Ser. B* **102** 1061–1066.

[25] Hirst, J. (2011) The inducibility of graphs on four vertices. arXiv:1109.1592

[26] Keevash, P. (2011) Hypergraph Turán problems. In *Surveys in combinatorics 2011*, Cambridge University Press, pp. 83–140.

[27] Keevash, P. and Mubayi, D. (2012) The Turán number of $F_{3,3}$. *Combin. Probab. Comput.* **21** 451–456.

[28] Kostochka, A. V. (1982) A class of constructions for Turán's (3, 4)-problem. *Combinatorica* **2** 187–192.

[29] Mubayi, D. Personal communication.

[30] Mubayi, D. and Pikhurko, O. (2008) Constructions of non-principal families in extremal hypergraph theory. *Discrete Math.* **308** 4430–4434.

[31] Mubayi, D. and Rödl, V. (2002) On the Turán number of triple systems. *J. Combin. Theory Ser. A* **100** 136–152.

[32] Pikhurko, O. Personal communication.

[33] Pikhurko, O. (2010) An analytic approach to stability. *Discrete Math.* **310** 2951–2964.

[34] Pikhurko, O. (2011) The minimum size of 3-graphs without a 4-set spanning no or exactly three edges. *Europ. J. Combin.* **32** 1142–1155.

[35] Pikhurko, O. (2012) On possible Turán densities. arXiv:1204.4423

[36] Razborov, A. A. (2007) Flag algebras. *J. Symbolic Logic* **72** 1239–1282.

[37] Razborov, A. A. (2010) On 3-hypergraphs with forbidden 4-vertex configurations. *SIAM J. Discrete Math.* **24** 946–963.

[38] Razborov, A. A. (2011) On the Fon-der-Flaass interpretation of extremal examples for Turán's (3,4)-problem. *Proc. Steklov Inst. Math.* **274** 247–266.

[39] Rödl, V. and Schacht, M. (2009) Generalizations of the removal lemma. *Combinatorica* **29** 467–501.

[40] Sidorenko, A. (1995) What we know and what we do not know about Turán numbers. *Graphs Combin.* **11** 179–199.

[41] Simonovits, M. (1968) A method for solving extremal problems in graph theory, stability problems. In *Theory of Graphs: Proc. Colloq., Tihany, 1966*, Academic, pp. 279–319.

[42] Vaughan, E. R. (2012) Flagmatic User's Guide, version 1.0. http://maths.qmul.ac.uk/~ev/flagmatic/usersguide.pdf.

https://doi.org/10.1017/S0963548312000508 Published online by Cambridge University Press