Journal of Combinatorial Theory, Series A 116 (2009) 637–649

Contents lists available at ScienceDirect

Journal of Combinatorial Theory, Series A

www.elsevier.com/locate/jcta

The maximum size of hypergraphs without generalized 4-cycles

Oleg Pikhurko [a,1], Jacques Verstraëte [b]

[a] Department of Mathematical Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

[b] Department of Mathematics, University of California, San Diego, La Jolla, CA 92093, USA

ARTICLE INFO

Article history:
Received 30 December 2007
Available online 14 November 2008

Keywords:
Generalized 4-cycle
Turán function

ABSTRACT

Let $f_r(n)$ be the maximum number of edges in an r-uniform hypergraph on n vertices that does not contain four distinct edges A, B, C, D with $A \cup B = C \cup D$ and $A \cap B = C \cap D = \emptyset$. This problem was stated by Erdős [P. Erdős, Problems and results in combinatorial analysis, Congr. Numer. 19 (1977) 3–12]. It can be viewed as a generalization of the Turán problem for the 4-cycle to hypergraphs.

Let $\phi_r = \limsup_{n\to\infty} f_r(n)/\binom{n}{r-1}$. Füredi [Z. Füredi, Hypergraphs in which all disjoint pairs have distinct unions, Combinatorica 4 (1984) 161–168] observed that $\phi_r \geqslant 1$ and conjectured that this is equality for every $r \geqslant 3$. The best known upper bound $\phi_r \leqslant 3$ was proved by Mubayi and Verstraëte [D. Mubayi, J. Verstraëte, A hypergraph extension of the bipartite Turán problem, J. Combin. Theory Ser. A 106 (2004) 237–253]. Here we improve this bound. Namely, we show that $\phi_r \leqslant \min(7/4, 1 + 2/\sqrt{r})$ for every $r \geqslant 3$, and $\phi_3 \leqslant 13/9$. In particular, it follows that $\phi_r \to 1$ as $r \to \infty$.

1. Introduction

Erdős [5] stated the following problem. Determine $f_r(n)$, the maximum number of edges in an r-graph on n vertices that does not contain four distinct edges A, B, C, D with $A \cup B = C \cup D$ and $A \cap B = C \cap D = \emptyset$.

For $r = 2$, this reduces to the known and well-studied problem of extremal graph theory to determine the Turán function for the 4-cycle. It was raised by Erdős [3] in 1938 and various results were

E-mail address: jacques@ucsd.edu (J. Verstraëte).

URL: http://www.math.cmu.edu/~pikhurko (O. Pikhurko).

[1] Partially supported by the National Science Foundation, Grant DMS-0457512.

0097-3165/$ – see front matter
doi:10.1016/j.jcta.2008.09.002

obtained in [1,2,4,6–8,10,17]. It is known that $f_2(n) = (\frac{1}{2} + o(1))n^{3/2}$ as $n \to \infty$, see Brown [1] and Erdős, Rényi, and Sós [7].

Thus the computation of $f_r(n)$ can be viewed as a generalization of this problem to hypergraphs. Therefore, we denote the family of forbidden r-graphs as $\mathcal{C}_4^r$ and call each member of $\mathcal{C}_4^r$ a *generalized* 4-*cycle*. When $r = 2$ or 3, there is only one forbidden subgraph up to isomorphism.

Bollobás and Erdős (unpublished, see [5, p. 11]) proved that there are constants $c_1, c_2 > 0$ such that $c_1 n^2 \leqslant f_3(n) \leqslant c_2 n^2$. As Erdős and Frankl pointed out in 1975 (see [9, p. 162]), one can show that $f_r(n) = O(n^{r-1/2})$ for all $r \geqslant 2$. Füredi [9] proved that

$$\binom{n-1}{r-1} + \left\lfloor \frac{n-1}{r} \right\rfloor \leqslant f_r(n) < \frac{7}{2}\binom{n}{r-1}. \tag{1}$$

The lower bound in (1) arises from the family of all r-element subsets of $[n] := \{1, \ldots, n\}$ containing the element n together with an arbitrary family of $\lfloor \frac{n-1}{r} \rfloor$ pairwise disjoint r-element subsets of $[n-1]$. Füredi [9, Example 1.3] also observed that if we replace every 5-set in a Steiner $S_1(n, 5, 2)$-system by all its 3-element subsets, then the resulting $\mathcal{C}_4^3$-free triple system has $\binom{n}{2}$ triples. A Steiner $S_1(n, 5, 2)$-system exists if and only if $n \equiv 1$ or 5 (mod 20), see Hanani [11,12]. Thus this construction improves the lower bound in (1) to $f_3(n) \geqslant \binom{n}{2}$ for such n. Füredi [9, Conjecture 1.4] conjectured that, for $n \geqslant n_0(r)$, $f_r(n) \leqslant \binom{n}{r-1}$ if $r \geqslant 3$ and $f_r(n) = \binom{n-1}{r-1} + \lfloor \frac{n-1}{r} \rfloor$ if $r \geqslant 4$.

Also, Mubayi [14, Conjecture 6.2] conjectured that the $f_r(n)$-problem is *stable* for $r \geqslant 4$, that is, for any $r \geqslant 4$ and $\delta > 0$ there are $\varepsilon > 0$ and n_0 such that any $\mathcal{C}_4^r$-free r-graph with n vertices and at least $(1-\varepsilon)\binom{n}{r-1}$ edges contains a vertex v belonging to at least $(1-\delta)\binom{n}{r-1}$ edges.

By (1), it makes sense to define

$$\phi_r = \limsup_{n\to\infty} \frac{f_r(n)}{\binom{n}{r-1}}, \quad r \geqslant 3. \tag{2}$$

Füredi [9, Proposition 6.1] showed that $f_3(n)/\binom{n}{2}$ converges as $n \to \infty$ but the existence of the limit for $r \geqslant 4$ is still an open question. Recently, Mubayi and Verstraëte [15] showed that $f_r(n) \leqslant 3\binom{n}{r-1} + O(n^{r-2})$ if $r \geqslant 3$ is fixed, thus improving the upper bound in (1). Hence, $1 \leqslant \phi_r \leqslant 3$ for any $r \geqslant 3$.

Here we prove the following results.

Theorem 1. *For every $r \geqslant 3$ we have $\phi_r \leqslant \min(1 + 2/\sqrt{r}, 7/4)$. In particular, $\lim_{r\to\infty} \phi_r = 1$.*

Unfortunately, the optimal upper bound on ϕ_3 given by our proof of Theorem 1 is a cumbersome expression, involving roots of cubic polynomials. Therefore, we decided to state the weaker bound $1 + 2/\sqrt{r}$ as well as to give some simple constant (i.e. 7/4) that is an upper bound on ϕ_r for every r. (Note that $1 + 2/\sqrt{r} < 7/4$ for $r \geqslant 8$.) We refer the reader to the remarks at the end of Section 6 for a discussion of the best upper bounds on ϕ_r given by our proof. These bounds are largest when $r = 3$ and 4.

Given this, the case $r = 3$ seems the most interesting one, especially that it is somewhat exceptional (if Füredi's conjecture is true). Also, the proofs in [9,15] proceeded by reducing the general case to the 3-partite version of the problem for 3-graphs. Therefore, we worked harder on the case $r = 3$ and were able to improve the bound $\phi_3 \leqslant 1.739\ldots$ of Theorem 1 (obtained by optimizing the constants) to $\phi_3 \leqslant 13/9 = 1.444\ldots$ in the following, somewhat more precise form.

Theorem 2. $f_3(n) \leqslant \frac{13}{9}\binom{n}{2}$ *for every $n \geqslant 1$.*

The key ingredient in our being able to improve the previous results comes from a strengthening of an auxiliary result of Füredi [9, Lemma 3.1] on the minimum number of edges that meet every 4-cycle in a graph. The exact statement of the new lemma and a short discussion can be found in Section 3.

Let us mention a few other related results. Mubayi [13] proved that if, in addition to $\mathcal{C}_4^3$, we also forbid the complete 3-partite 3-graph with parts of sizes 1, 2 and 4, then the maximum number of triples on n vertices is indeed at most $\binom{n}{2}$. Mubayi and Verstraëte [16] showed that the maximum size of an r-graph on n vertices without any *minimal* 4-*cycle* (a certain r-graph family including all $\mathcal{C}_4^r$-cycles) is $\binom{n-1}{r-1} + O(n^{r-2})$. Mubayi and Verstraëte [15] stated some generalizations of the $f_r(n)$-problem and presented various bounds.

Here we concentrate on the original function $f_r(n)$. Our paper is organized as follows. Section 2 lists the notation used in this paper. Some auxiliary results for graphs are presented in Section 3 and for hypergraphs in Sections 4–5. Theorem 1 is proved in Section 6 and Theorem 2 in Section 7.

2. Notation

We use the following notation in this paper. We denote $[n] = \{1, \dots, n\}$. Also, $\binom{X}{k} = \{Y \subseteq X\colon |Y| = k\}$ is the set of all k-subsets of a set X. For brevity, we use abbreviations like $ab = \{a, b\}$ and $abc = \{a, b, c\}$ (and even $a = \{a\}$ in the cases when the meaning is clear).

An *r-graph* (or an *r-uniform set system*) on a set X is $\mathcal{G} \subseteq \binom{X}{r}$, a collection of r-subsets of X. We identify r-graphs with their edge sets so that, for example, $|\mathcal{G}|$ denotes the number of edges of $\mathcal{G}$. The *vertex set* of $\mathcal{G}$ is $V(\mathcal{G}) = \bigcup_{E \in \mathcal{G}} E$. When $r = 2$, we use the term *graph*.

Some special r-graphs are as follows.

- $\mathcal{C}_4(ab, cd)$ is the graph $\{ac, ad, bc, bd\}$.
- $\mathcal{C}_4^3(ab, cd, ef)$ is the 3-graph $\{abc, abd, cef, def\}$.
- $\mathcal{C}_4^r$ denotes the family of r-graphs with four distinct edges A, B, C, D such that $A \cup B = C \cup D$ and $A \cap B = C \cap D = \emptyset$.
- $\mathcal{P}(T, P)$ is the 3-graph consisting of all triples E with $|T \cap E| = 2$ and $|P \cap E| = 1$, where T and P and are two disjoint sets of vertices.
- $\mathcal{P}_i$ is a copy of $\mathcal{P}(T, P)$ with $|T| = 3$ and $|P| = i$.
- $\mathcal{K}_5^-$ is the family of all 3-graphs with 5 vertices and at least 8 edges such that if two different triples are missing then these triples intersect in precisely one vertex. (Thus, up to isomorphism, $\mathcal{K}_5^-$ has three different 3-graphs.)

Let $\mathcal{G} \subseteq \binom{X}{2}$ be a graph and $ab \in \binom{X}{2}$. The *μ-multiplicity* of ab in $\mathcal{G}$ is $\mu_{\mathcal{G}}(ab) = |\{x \in X\colon ax, bx \in \mathcal{G}\}|$, the number of 2-paths connecting a to b. The pair ab is a *diagonal* of $\mathcal{G}$ if $\mu_{\mathcal{G}}(ab) \geqslant 2$ (equivalently, if a and b are diametrally opposite points on some 4-cycle in $\mathcal{G}$). Let

$$\mathcal{D}(\mathcal{G}) = \left\{xy \in \binom{X}{2}\colon\ \mu_{\mathcal{G}}(xy) \geqslant 2\right\}$$

be the set of all diagonals of $\mathcal{G}$. The pair ab is a *half-diagonal* if $\mu_{\mathcal{G}}(ab) = 1$. Let

$$\mathcal{E}(\mathcal{G}) = \left\{xy \in \binom{X}{2}\colon\ \mu_{\mathcal{G}}(xy) = 1\right\}.$$

Let $\mathcal{G} \subseteq \binom{X}{r}$ be an r-graph. For $A \subseteq X$, its *link* $(r - |A|)$*-graph* is $\mathcal{G}_A = \{B \subseteq X \setminus A\colon\ A \cup B \in \mathcal{G}\}$. If $|A| = r - 1$, then we view $\mathcal{G}_A$ as a set of vertices rather than as a set of single-element sets. Also, $\mathcal{G}[B] = \{E \in \mathcal{G}\colon\ E \subseteq B\}$ denotes the subgraph induced by a set $B \subseteq X$.

Let $\mathcal{G} \subseteq \binom{X}{3}$ be a 3-graph. Let

$$\mathcal{D}_i(\mathcal{G}) = \left\{ab \in \binom{X}{2}\colon\ \left|\left\{x \in X \setminus ab\colon\ ab \in \mathcal{D}(\mathcal{G}_x)\right\}\right| = i\right\}$$

consist of those pairs ab that are diagonals in exactly i link graphs of $\mathcal{G}$. Define $\mathcal{D}(\mathcal{G}) = \bigcup_{i \geqslant 1} \mathcal{D}_i(\mathcal{G})$ and $\mathcal{E}(\mathcal{G}) = (\bigcup_{x \in X} \mathcal{E}(\mathcal{G}_x)) \setminus \mathcal{D}(\mathcal{G})$. We call the elements of $\mathcal{D}(\mathcal{G})$ (respectively $\mathcal{E}(\mathcal{G})$) *diagonals* (respectively *half-diagonals*) of the 3-graph $\mathcal{G}$.

3. Removing diagonals in graphs

Füredi [9, Lemma 3.1] proved that any bipartite graph $\mathcal{G}$ can be made $\mathcal{C}_4$-free by removing at most $|\mathcal{D}(\mathcal{G})|$ edges. This lemma, interesting on its own, turned out to be very useful in proving upper bounds on $f_r(n)$, see [9,15]. Here we strengthen Füredi's lemma in two directions simultaneously. Firstly, we remove the assumption that $\mathcal{G}$ is bipartite. Secondly, we show that every diagonal of the original graph $\mathcal{G}$ is neither a diagonal nor a half-diagonal in the obtained graph $\mathcal{G}'$. (In general, $\mathcal{G}'$ need not consist of isolated edges only so it may have some other pairs as half-diagonals.)

Lemma 3. *For any graph $\mathcal{G}$ there is an edge set $\mathcal{R} \subseteq \mathcal{G}$ such that $|\mathcal{R}| \leqslant |\mathcal{D}(\mathcal{G})|$ and*

$$\mathcal{D}(\mathcal{G}) \cap \big(\mathcal{D}(\mathcal{G}') \cup \mathcal{E}(\mathcal{G}')\big) = \emptyset, \tag{3}$$

where $\mathcal{G}' = \mathcal{G} \setminus \mathcal{R}$, that is, $\mu_{\mathcal{G}'}(ab) = 0$ for every diagonal ab of $\mathcal{G}$. (In particular, $\mathcal{G}'$ is C_4-free.)

Proof. Let us prove the following claim first.

Claim 1. *For any graph $\mathcal{H}$ that contains at least one 4-cycle, we can remove a non-empty set $\mathcal{R} \subseteq \mathcal{H}$ of edges so that the obtained graph $\mathcal{H}' = \mathcal{H} \setminus \mathcal{R}$ satisfies*

$$\big|\mathcal{D}(\mathcal{H}) \setminus \big(\mathcal{D}(\mathcal{H}') \cup \mathcal{E}(\mathcal{H}')\big)\big| \geqslant |\mathcal{R}|, \tag{4}$$

that is, we have $\mu_{\mathcal{H}'}(ab) = 0$ for at least $|\mathcal{R}|$ diagonals $ab \in \mathcal{D}(\mathcal{H})$.

Proof. Choose any uv which is an edge of at least one C_4-subgraph of $\mathcal{H}$. For $x \in V(\mathcal{H})$, let $D(x) = \{y \in V(\mathcal{H}):\ xy \in \mathcal{D}(\mathcal{H})\}$. Define

$$X = D(u) \cap \mathcal{H}_v = \big\{x \in V(\mathcal{H}):\ ux \in \mathcal{D}(\mathcal{H}),\ vx \in \mathcal{H}\big\},$$
$$Y = D(v) \cap \mathcal{H}_u = \big\{y \in V(\mathcal{H}):\ vy \in \mathcal{D}(\mathcal{H}),\ uy \in \mathcal{H}\big\}.$$

Note that both X and Y are non-empty, because there is a 4-cycle containing the edge uv. Also, $u, v \notin X \cup Y$.

Remove all edges between v and X and between u and Y. It is enough to show that in the obtained graph $\mathcal{H}'$ all pairs vy with $y \in Y$ and ux with $x \in X$ have μ-multiplicity 0. Suppose on the contrary that, for example, for some $w \in V(\mathcal{H}')$ and $y \in Y$ we have $vw, wy \in \mathcal{H}'$. Then $w \neq u$ because uy has been deleted but $wy \in \mathcal{H}'$. Thus $C_4(uw, vy) \subseteq \mathcal{H}$. By definition, $w \in X$. But all edges between v and X have been deleted and cannot belong to $\mathcal{H}'$, a contradiction that finishes the proof of Claim 1. □

Let us return to the proof of the lemma. Starting with $\mathcal{H} = \mathcal{G}$, we iteratively apply Claim 1 and keep removing edges from $\mathcal{H}$ until no C_4-subgraph remains. Suppose we have removed k edges in total. Let $\mathcal{H}'$ be the final graph. Then the number of $ab \in \mathcal{D}(\mathcal{G})$ with $\mu_{\mathcal{H}'}(ab) = 0$ is at least k. Since $\mathcal{H}'$ is C_4-free, all other diagonals ab of $\mathcal{G}$ satisfy $\mu_{\mathcal{H}'}(ab) = 1$. For each $ab \in \mathcal{D}(\mathcal{G}) \cap \mathcal{E}(\mathcal{H}')$ pick an edge $E_{ab} \in \mathcal{H}'$ whose removal would bring the μ-multiplicity of ab to 0. Note that there are at most $|\mathcal{D}(\mathcal{G})| - k$ such pairs ab. Let $\mathcal{G}'$ be obtained from $\mathcal{H}'$ by removing all such edges E_{ab}. Clearly, the edge set $\mathcal{R} = \mathcal{G} \setminus \mathcal{G}'$ satisfies all the conclusions of the lemma. □

We will also need the following simple observation.

Lemma 4. *Let $\mathcal{G}$ be a C_4-free graph on n vertices with $g = |\mathcal{G}|$ edges. Then $|\mathcal{E}(\mathcal{G})| \geqslant 2g - n$.*

Proof. Let $g_1 \geqslant \cdots \geqslant g_n$ be the degrees of $\mathcal{G}$. Since $C_4 \nsubseteq \mathcal{G}$, we have

$$\big|\mathcal{E}(\mathcal{G})\big| = \sum_{i=1}^{n} \binom{g_i}{2}. \tag{5}$$

Note that $\binom{k}{2} \geqslant k - 1$ for every non-negative integer k. This, (5), and the identity $\sum_{i=1}^{n}(g_i - 1) = 2g - n$ imply the lemma. □

4. Multiple diagonals in $\mathcal{C}_4^3$-free 3-graphs

Here we present some auxiliary lemmas about *multiple diagonals* (i.e. pairs in $\bigcup_{i\geqslant 2}\mathcal{D}_i(\mathcal{G})$) for $\mathcal{C}_4^3$-free 3-graphs. All results in this section are obtained by a straightforward case analysis. The reader may wish to skip the proofs and refer only to the statements of the results.

The following lemma implies that $\mathcal{D}_i(\mathcal{G})=\emptyset$ for any $\mathcal{C}_4^3$-free 3-graph $\mathcal{G}$ except possibly for $i\in\{0,1,3\}$. Recall that $\mathcal{P}(X,Y)$ denotes the 3-graph with edges $\{xyz\colon\ xz\in\binom{X}{2},\ y\in Y\}$.

Lemma 5. *For an arbitrary $\mathcal{C}_4^3$-free 3-graph $\mathcal{G}$, each pair ab of vertices is a diagonal for either 0, 1, or 3 link graphs. In the last case, there is a 3-set X such that $\mathcal{P}(X,ab)\subseteq\mathcal{G}$ and, moreover, for every $x\in X$, $\mathcal{C}_4(ab,X\setminus x)$ is the unique cycle of the link graph $\mathcal{G}_x$ that has ab as a diagonal.*

Proof. Suppose that ab is a diagonal for a least two link graphs, say, $ab\in\mathcal{D}(\mathcal{G}_x)\cap\mathcal{D}(\mathcal{G}_{x'})$ for some distinct $x,x'\in V(\mathcal{G})$. Let the corresponding 4-cycles be $\mathcal{C}_4(ab,cd)\subseteq\mathcal{G}_x$ and $\mathcal{C}_4(ab,c'd')\subseteq\mathcal{G}_{x'}$.

Let us first derive a contradiction by assuming that $x'\notin cd$. At least one of two distinct vertices c' and d' is not equal to x, so assume without loss of generality that $d'\neq x$. Likewise, by the symmetry of cd, we can assume that $c\neq d'$. But then $\mathcal{C}_4^3(xc,ab,x'd')\subseteq\mathcal{G}$, a contradiction.

This shows that $x'\in cd$. Thus ab can be a diagonal for at most three link graphs of $\mathcal{G}$. Without loss of generality assume that $x'=c$ (thus $ab\in\mathcal{D}(\mathcal{G}_c)$). It follows that $\mathcal{C}_4(ab,cd)$ is the unique cycle of $\mathcal{G}_x$ having ab for a diagonal: if $av,bv\in\mathcal{G}_x$ with $v\notin cd$, then $\mathcal{C}_4(ab,dv)$ is another 4-cycle with the diagonal ab that omits the vertex $c=x'$, a contradiction to the arguments from the previous paragraph. Clearly, the roles of c and x can be interchanged. Thus the unique 4-cycle in $\mathcal{G}_c$ that has ab for a diagonal has to use the vertex x. If $\mathcal{C}_4(ab,ex)\subseteq\mathcal{G}_c$ with $e\neq d$, then $\mathcal{C}_4^3(ce,ab,dx)\subseteq\mathcal{G}$, a contradiction. Otherwise, the unique cycle is $\mathcal{C}_4(ab,dx)\subseteq\mathcal{G}_c$. It follows that $\mathcal{P}(cdx,ab)\subseteq\mathcal{G}$ and $ab\in\mathcal{D}_3(\mathcal{G})$. By symmetry, $\mathcal{C}_4(cx,ab)$ is the unique 4-cycle of $\mathcal{G}_d$ that has ab as a diagonal. The lemma is proved. □

Lemma 6. *Let $\mathcal{G}$ be a $\mathcal{C}_4^3$-free 3-graph and let $\mathcal{P}(xyz,ab)\subseteq\mathcal{G}$. If $auv,buv\in\mathcal{G}$, then $uv\subseteq xyz$.*

Proof. If $uv\not\subseteq xyz$, then by symmetry we can assume that $xy\cap uv=\emptyset$. But then $\mathcal{C}_4^3(uv,ab,xy)\subseteq\mathcal{G}$, a contradiction. □

Lemma 7. *Let $\mathcal{G}$ be a $\mathcal{C}_4^3$-free 3-graph and let $\mathcal{P}(T,P),\mathcal{P}(T',P')\subseteq\mathcal{G}$ be such that $T\neq T'$, $T=t_1t_2t_2$ and $T'=t_1't_2't_3'$ are 3-element sets while $P\supseteq p_1p_2$ and $P'\supseteq p_1'p_2'$ have at least 2 elements each. If*

$$\left(\binom{T}{2}\cup\binom{P}{2}\right)\cap\left(\binom{T'}{2}\cup\binom{P'}{2}\right)\neq\emptyset,$$

then $|P|=|P'|=2$, $T\cup P=T'\cup P'$, and the 5-set $T\cup P$ spans a $\mathcal{K}_5^-$-subgraph in $\mathcal{G}$.

Proof. We have $p_1p_2\neq p_1'p_2'$ for otherwise the pair p_1p_2 is a diagonal for at least four different link graphs $\mathcal{G}_x$ (namely, for $x\in T\cup T'$), contradicting Lemma 5. Up to a symmetry, there are the following two cases to consider.

Case 1. $T'=\{t_1,t_2,t_3'\}$.

Since $T\neq T'$, we have $t_3\neq t_3'$. Let us show that $p_1'\in p_1p_2t_3$. Indeed, otherwise at least one of $p_1\neq p_2$ is distinct from t_3', say $p_1\neq t_3'$. But then $\mathcal{C}_4^3(p_1t_3,t_1t_2,p_1't_3')\subseteq\mathcal{G}$, a contradiction. By symmetry, $p_2'\in p_1p_2t_3$ and $p_1,p_2\in p_1'p_2't_3'$. Since $p_1p_2\neq p_1'p_2'$, we have $t_3'\in p_1p_2$ and $t_3\in p_1'p_2'$.

If there is a vertex $p_3\in P\setminus p_1p_2$, then the above arguments with p_1 replaced by p_3 imply that $T\cup p_1p_2p_3\subseteq T'\cup p_1'p_2'$, a contradiction. It follows that $|P|=2$ and, by symmetry, $|P'|=2$. This in turn implies that $T\cup P=T'\cup P'$. It routinely follows that at most one triple of $T\cup P$ (namely $P\cup P'$) can be missing from $\mathcal{G}$, settling Case 1.

Case 2. Let $p'_1 = t_1$ and $p'_2 = t_2$.

Suppose first that $T' \nsubseteq T \cup p_1p_2$. By symmetry, we can assume that $t'_1 \notin T \cup p_1p_2$. If $p_1 \notin T'$, then one of $t'_2 \neq t'_3$ is not equal to t_3, say $t'_2 \neq t_3$, and we have $\mathcal{C}_4^3(p_1t_3, t_1t_2, t'_1t'_2) \subseteq \mathcal{G}$, a contradiction. Hence $p_1 \in T'$. Likewise, $p_2 \in T'$. Since $t'_1 \notin p_1p_2$, we have $t'_2t'_3 = p_1p_2$. But then $\mathcal{C}_4^3(t'_1t'_2, t_1t_2, t_3t'_3) \subseteq \mathcal{G}$, a contradiction.

Thus $T' \subseteq T \cup p_1p_2$, that is, $T' = p_1p_2t_3$. Since $p_1, p_2 \in P$ were arbitrary, we conclude that $|P| = 2$. We cannot have another vertex $p'_3 \in P' \setminus p'_1p'_2$ for otherwise $\mathcal{C}_4^3(t_1t_2, p_1p_2, p'_3t_3) \subseteq \mathcal{G}$, a contradiction. Thus $|P'| = 2$ and $T \cup P = T' \cup P'$. It is routine to see that most two triples can be missing from $\mathcal{G}[T \cup P]$, namely T and T' with $|T \cap T'| = 1$.

The lemma is proved. $\square$

Let us call any element of $\binom{T}{2} \cup \binom{P}{2}$, where $|T| = 3$ and $|P| \geqslant 2$, a *private pair* of the 3-graph $\mathcal{P}(T, P)$.

Lemma 8. *Let $\mathcal{G} \in \mathcal{K}_5^-$. Then for every two distinct points $x, y \in V(\mathcal{G})$ there are two triples in $\mathcal{G}$ whose symmetric difference is xy.*

Proof. Assume $V(\mathcal{G}) = abcde$ with possible missing triples being abc and cde. Up to a symmetry, there are three different cases to consider. If xy is respectively ab, ac, and ad, then the required triples are $uvx, uvy \in \mathcal{G}$, where uv is respectively de, bd, and be. $\square$

Lemma 9. *Let $\mathcal{G}$ be a $\mathcal{C}_4^3$-free* 3*-graph and $ab \in \mathcal{E}(\mathcal{G})$. Then there is exactly one pair uv with $auv, buv \in \mathcal{G}$.*

Proof. Let u be such that $ab \in \mathcal{E}(\mathcal{G}_u)$. This means that there is a vertex v such that $av, bv \in \mathcal{G}_u$, which implies the existence of the desired pair uv. On the other hand, if there was another such pair $u'v'$, then we would obtain a contradiction: if $uv \cap u'v' = \emptyset$, then $\mathcal{C}_4^3(uv, ab, u'v') \subseteq \mathcal{G}$, otherwise $ab \in \mathcal{D}(\mathcal{G})$. $\square$

5. Removing diagonals in 3-graphs

Recall that, for a 3-graph $\mathcal{G}$ and an integer $i \geqslant 0$, $\mathcal{D}_i(\mathcal{G})$ consists of all pairs ab such that ab is a diagonal in exactly i link graphs $\mathcal{G}_x$. Also,

$$\mathcal{D}(\mathcal{G}) = \bigcup_{i \geqslant 1} \mathcal{D}_i(\mathcal{G}) = \bigcup_{x \in V(\mathcal{G})} \mathcal{D}(\mathcal{G}_x),$$

is the set of all diagonals, ignoring their multiplicity.

Here we prove a version of Lemma 3 for 3-graphs. We will show that one can destroy all diagonals of a $\mathcal{C}_4^3$-free 3-graph $\mathcal{G}$ by removing at most $|\mathcal{D}(\mathcal{G})|$ edges. Although we cannot prevent some diagonals of $\mathcal{G}$ becoming half-diagonals of the final 3-graph $\mathcal{G}'$, we nonetheless get some control over their distribution. We do need the assumption that $\mathcal{C}_4^3 \nsubseteq \mathcal{G}$: for example, one has to remove $\Omega(n^3)$ edges in order to destroy all diagonals in the complete 3-graph $\binom{[n]}{3}$. As we already know by Lemma 5, this assumption implies that $\mathcal{D}_i(\mathcal{G})$ is empty except possibly for $i = 0, 1$, or 3. It is not surprising that the main idea behind the proof of Lemma 10 below is to apply Lemma 3 to each link graph $\mathcal{G}_x$.

Lemma 10. *Let $\mathcal{G}$ be a $\mathcal{C}_4^3$-free* 3*-graph with vertex set V. Let $\preccurlyeq$ be an arbitrary linear ordering of V. Then there is a subgraph $\mathcal{G}' \subseteq \mathcal{G}$ such that all the following properties hold.*

1. $|\mathcal{G}| - |\mathcal{G}'| \leqslant |\mathcal{D}(\mathcal{G})|$.
2. *$\mathcal{D}(\mathcal{G}') = \emptyset$, that is, all link graphs of $\mathcal{G}'$ are C_4-free.*
3. *If $ab \in \mathcal{D}_1(\mathcal{G})$, then there do not exist $u, v \in V \setminus ab$ with $auv, buv \in \mathcal{G}'$. (This property implies that $\mathcal{E}(\mathcal{G}') \cap \mathcal{D}_1(\mathcal{G}) = \emptyset$.)*

4. *If $ab \in \mathcal{D}_3(\mathcal{G})$, say*

$$ab \in \mathcal{D}(\mathcal{G}_x) \cap \mathcal{D}(\mathcal{G}_y) \cap \mathcal{D}(\mathcal{G}_z), \quad \text{with } x \prec y \prec z, \tag{6}$$

and $auv, buv \in \mathcal{G}'$, then necessarily $uv = yz$.

Proof. We take the vertices of $\mathcal{G}$ one by one in the $\preccurlyeq$-ordering. For each $x \in V$, we construct a set $\mathcal{R}_x$ of edges so that $\mathcal{G}' = \mathcal{G} \setminus (\bigcup_{x\in V} \mathcal{R}_x)$ satisfies all the properties. In order to establish Property 1, we also define an injection $\ell : \bigcup_{x\in V} \mathcal{R}_x \to \mathcal{D}(\mathcal{G})$. Since we will need to refer to the original graph $\mathcal{G}$ later, it remains unchanged throughout the proof. For $ab \in \mathcal{D}(\mathcal{G})$, let

$$f(ab) = \min_{\preccurlyeq}\{y \in V:\ ab \in \mathcal{D}(\mathcal{G}_y)\}, \tag{7}$$

be the $\preccurlyeq$-smallest vertex $y \in V$ with $ab \in \mathcal{D}(\mathcal{G}_y)$.

Suppose we are about to start working on the next vertex $x \in V$. Define $V_{\prec x} = \{y \in V:\ y \prec x\}$ and

$$\mathcal{H}^x = \mathcal{G} \setminus \left(\bigcup_{y\in V_{\prec x}} \mathcal{R}_y\right). \tag{8}$$

One can view $\mathcal{H}^x$ as the 'current' 3-graph at the moment when we have just deleted the sets $\mathcal{R}_y$ for all y preceding x.

Apply Lemma 3 to $\mathcal{H}^x_x$, the link graph of the vertex x in the 3-graph $\mathcal{H}^x$. The lemma returns an edge set $\mathcal{R} \subseteq \mathcal{H}^x_x$. Let $\mathcal{R}'_x = \{abx:\ ab \in \mathcal{R}\} \subseteq \mathcal{H}^x$ and $\mathcal{H}' = \mathcal{H}^x \setminus \mathcal{R}'_x$. Extend the function ℓ to $\mathcal{R}'_x$ by mapping $\mathcal{R}'_x$ injectively into $\mathcal{D}(\mathcal{H}^x_x) \subseteq \mathcal{D}(\mathcal{G}_x)$. (Note that $|\mathcal{R}'_x| \leqslant |\mathcal{D}(\mathcal{H}^x_x)|$ by Lemma 3.) We will argue later that no new value of ℓ coincides with a previous value.

Let $\mathcal{R}''_x = \emptyset$. Take one by one pairs $ab \in \mathcal{D}(\mathcal{G}_x) \setminus \mathcal{D}(\mathcal{H}^x_x)$. If $x = f(ab)$, where $f(ab)$ is defined by (7), and there is a vertex w with $aw, bw \in \mathcal{H}'_x$ (of course, if w exists, it is unique because $\mathcal{H}'_x$ is $\mathcal{C}_4$-free), and $awx \notin \mathcal{R}''_x$, then add awx to $\mathcal{R}''_x$ and define $\ell(awx) = ab$. (To avoid any ambiguity, we may agree that, for example, $a \prec b$.) Otherwise, we do nothing with $\mathcal{R}''_x$ for this pair ab. Having processed all pairs $ab \in \mathcal{D}(\mathcal{G}_x) \setminus \mathcal{D}(\mathcal{H}^x_x)$ this way, let $\mathcal{R}_x = \mathcal{R}'_x \cup \mathcal{R}''_x$. Note that $\mathcal{R}_x \subseteq \mathcal{H}^x$ is disjoint from $\bigcup_{y\in V_{\prec x}} \mathcal{R}_y$ by (8). The function ℓ has already been defined on the elements of this set; we have

$$\ell(\mathcal{R}_x) \subseteq \mathcal{D}(\mathcal{G}_x). \tag{9}$$

This finishes the definition of $\mathcal{R}_x$.

Being done with x, take the next vertex of V with respect the $\preccurlyeq$-ordering. If x is the last vertex, then we have defined $\mathcal{R}_y$ for every $y \in V$ and we let $\mathcal{G}' = \mathcal{G} \setminus (\bigcup_{y\in V} \mathcal{R}_y)$.

We have achieved that, for every $ab \in \mathcal{D}(\mathcal{G}_x)$ with $x = f(ab)$, the link graph $(\mathcal{H}^x \setminus \mathcal{R}_x)_x$ contains no 2-path connecting a to b. Indeed, this follows from Lemma 3 if $ab \in \mathcal{D}(\mathcal{H}^x_x)$ and from the definition of $\mathcal{R}''_x$ otherwise. Thus, for every $ab \in \mathcal{D}(\mathcal{G})$ we have

$$\mu_{\mathcal{G}'_x}(ab) = \mu_{(\mathcal{H}^x\setminus\mathcal{R}_x)_x}(ab) = 0, \quad \text{where } x = f(ab). \tag{10}$$

Let us show all the claims with respect to the final 3-graph $\mathcal{G}'$.

Property 1. It suffices to show that $\ell : \bigcup_{x\in V} \mathcal{R}_x \to \mathcal{D}(\mathcal{G})$ is an injection.

Let $ab \in \mathcal{D}(\mathcal{G})$ be arbitrary and let $x = f(ab)$. Let us show that if $ab \in \mathcal{D}(\mathcal{G})$ belongs to $\ell(\mathcal{R}_u)$ for some $u \in V$, then $u = x$. This is clearly true for $ab \in \mathcal{D}_1(\mathcal{G})$ by (9), so suppose that $ab \in \mathcal{D}_3(\mathcal{G})$. By Lemma 5, there is a pair $yz \in \binom{V}{2}$ with $\mathcal{P}(xyz, ab) \subseteq \mathcal{G}$. Since $x = f(ab)$, we have $x \prec y$ and $x \prec z$. By (9), we have $ab \in \mathcal{D}(\mathcal{G}_u)$. Hence $u \in xyz$. Suppose on the contrary that $u \neq x$, say $u = y$. Since $x = f(ab)$ and $\mathcal{H}^y \subseteq \mathcal{H}^x \setminus \mathcal{R}_x$, (10) implies that at least one of the triples axy and bxy is missing from $\mathcal{H}^y$. But $\mathcal{C}_4(ab, xz)$ is the unique 4-cycle of $\mathcal{G}_y$ having ab for a diagonal by the second part of Lemma 5. Hence $ab \notin \mathcal{D}(\mathcal{H}^y)$. Thus $ab \notin \ell(\mathcal{R}'_y)$. Also, $ab \notin \ell(\mathcal{R}''_y)$ because $y \neq f(ab)$. So $ab \notin \ell(\mathcal{R}_y)$, a contradiction that proves that $u = x$ as claimed.

Hence, it is enough to show that ℓ is injective on $\mathcal{R}_x = \mathcal{R}'_x \cup \mathcal{R}''_x$ for every $x \in V$. Note that $\ell(\mathcal{R}'_x) \subseteq \mathcal{D}(\mathcal{H}^x_x)$ is disjoint from $\ell(\mathcal{R}''_x) \subseteq \mathcal{D}(\mathcal{G}_x) \setminus \mathcal{D}(\mathcal{H}^x_x)$. The injectivity of ℓ on each of $\mathcal{R}'_x$ and $\mathcal{R}''_x$ is obvious from the definition. Thus ℓ is an injection, as required.

Property 2 clearly holds because Lemma 3 was applied to the x-link graph of the current hypergraph for each $x \in V$.

Property 3. Suppose on the contrary that ab and uv contradict it. Let $ab \in \mathcal{D}(\mathcal{G}_x)$. Choose a 4-cycle demonstrating this fact, say $\mathcal{C}_4(ab, cd) \subseteq \mathcal{G}_x$ for some vertices c and d. We cannot have $x \in uv$ because $x = f(ab)$ and there is no 2-path connecting a to b in $\mathcal{G}'_x$ by (10). If $uv = cd$, then $\mathcal{P}(cdx, ab) \subseteq \mathcal{G}$, which shows that $ab \in \mathcal{D}(\mathcal{G}_x) \cap \mathcal{D}(\mathcal{G}_c) \cap \mathcal{D}(\mathcal{G}_d)$, a contradiction to our assumption $ab \in \mathcal{D}_1(\mathcal{G})$. Otherwise (if $uv \neq cd$), we can assume by symmetry that $c \notin uv$, but then $\mathcal{C}_4^3(uv, ab, xc) \subseteq \mathcal{G}$, a contradiction proving Property 3.

Property 4. Suppose that vertices a, b, x, y, z and a pair uv satisfy all assumptions of Property 4. Lemma 5 implies that $\mathcal{P}(xyz, ab) \subseteq \mathcal{G}$ and Lemma 6 implies that $uv \subseteq xyz$. Since $x = f(ab)$, the link graph $\mathcal{G}'_x$ cannot contain a 2-path connecting a and b by (10). Thus we have $uv = yz$, as required.

The lemma is completely proved. □

Remark. In fact, Property 2 follows from Properties 3–4 (and the $\mathcal{C}_4^3$-freeness of $\mathcal{G}$) but it is convenient to have it explicitly stated.

The following lemma is needed in the proof of Theorem 2 but not in that of Theorem 1.

Lemma 11. *Let $\mathcal{G}$ be an arbitrary $\mathcal{C}_4^3$-free 3-graph on an n-set V. Then we can find a set of edges $\mathcal{R} \subseteq \mathcal{G}$ such that $\mathcal{G}' = \mathcal{G} \setminus \mathcal{R}$ does not contain a $\mathcal{K}_5^-$-subgraph and*

$$\left|\mathcal{D}(\mathcal{G}') \cup \mathcal{E}(\mathcal{G}')\right| \leqslant \binom{n}{2} - |\mathcal{R}|. \tag{11}$$

Proof. Let $\mathcal{A} \subseteq \binom{V}{5}$ be the family of the vertex sets of all $\mathcal{K}_5^-$-subgraphs of $\mathcal{G}$.

Let us show that

$$A, B \in \mathcal{A},\ \ A \neq B \quad \Longrightarrow \quad |A \cap B| \leqslant 1. \tag{12}$$

Let $A, B \in \mathcal{A}$. Since $\mathcal{G}$ is $\mathcal{C}_4^3$-free, Lemma 8 implies that $|A \cap B| = 2$ is impossible. Suppose next that $A \cap B = xyz$. Let $A \setminus xyz = uv$. Every two of the triples uvx, uvy, uvz share two vertices so at most one can be missing from $\mathcal{G}$ because $\mathcal{G}[A] \in \mathcal{K}_5^-$. Assume that $uvx, uvy \in \mathcal{G}$. Lemma 8, when applied to xy and B, produces $u', v' \in B$ with $u'v'x, u'v'y \in \mathcal{G}$. This gives $\mathcal{C}_4^3(uv, xy, u'v') \subseteq \mathcal{G}$, a contradiction. Finally, let us derive a contradiction by assuming that $|A \cap B| = 4$. Since this can make our task only harder, assume that two edges are missing from $\mathcal{G}[A]$ (respectively $\mathcal{G}[B]$). Let $a \in A$ (respectively $b \in B$) be the vertex shared by these two edges. The 3-graph $\mathcal{G}[A]$ contains two $\mathcal{P}_2$-subgraphs whose private pairs form two triangles sharing the vertex a. (Recall that $\mathcal{P}_2$ is a copy of $\mathcal{P}(xyz, uv)$ and its private pairs are uv, xy, xz, and yz.) A quadruple $X = A \cap B$ of vertices contains either 2 disjoint private pairs (if $a \notin X$) or a triangle with a pendant edge (if $a \in X$). The analogous claims hold for B. Since $A \neq B$, the 3-graphs $\mathcal{G}[A]$ and $\mathcal{G}[B]$ cannot share a private pair by Lemma 7. It follows that $a, b \notin X$ and the missing edges are auv, axy, bux, bvy, where $X = uvxy$. But then $\mathcal{C}_4^3(uy, ab, vx) \subseteq \mathcal{G}$. This contradiction proves (12).

Let $\mathcal{R} = \mathcal{M} \cup \mathcal{S}$, where $\mathcal{M} = \bigcup_{A \in \mathcal{A}} \mathcal{G}[A]$ consists of what we call *main triples* and

$$\mathcal{S} = \left\{E \in \mathcal{G}\colon\ \exists A \in \mathcal{A},\ |E \cap A| = 2\right\},$$

consists of all *secondary triples*. Let $\mathcal{G}' = \mathcal{G} \setminus \mathcal{R}$. Let us define the logic predicate $\mathcal{L}(x, y, uv, A)$ which is true if and only if $A \in \mathcal{A}$, $u, v, y \in A$, $x \in V \setminus A$, and $uvx, uvy \in \mathcal{G}$.

In order to prove the lemma it is enough to specify a set $\mathcal{R}' \subseteq \binom{V}{2}$ of pairs such that $|\mathcal{R}'| \geqslant |\mathcal{R}|$ and

$$\mathcal{R}' \cap \left(\mathcal{D}(\mathcal{G}') \cup \mathcal{E}(\mathcal{G}')\right) = \emptyset. \tag{13}$$

Let $\mathcal{M}'=\bigcup_{A\in\mathcal{A}}\binom{A}{2}$ consist of what we call *main pairs*. Let us call a pair $xy\in\binom{V}{2}$ a *secondary pair* if there are u, v, and A satisfying $\mathcal{L}(x,y,uv,A)$ or $\mathcal{L}(y,x,uv,A)$. Let $\mathcal{S}'\subseteq\binom{V}{2}$ consist of all secondary pairs. Let us show that $\mathcal{R}'=\mathcal{M}'\cup\mathcal{S}'$ is the required set.

Let us check (13) first. Suppose on the contrary to (13) that $xy\in\mathcal{R}'$ and $ab\in\binom{V}{2}$ satisfy $abx, aby\in\mathcal{G}'$. If xy lies inside some $A\in\mathcal{A}$ (i.e. it is a main pair), then $a\notin A$ and $b\notin A$ for otherwise, for example, aby belongs to $\mathcal{R}=\mathcal{G}\setminus\mathcal{G}'$, a contradiction. By Lemma 8, there are $a',b'\in A$ with $a'b'x, a'b'y\in\mathcal{G}$. But then $\mathcal{C}_4^3(ab,xy,a'b')\subseteq\mathcal{G}$, a contradiction. So suppose that xy is a secondary pair, which is witnessed by $\mathcal{L}(x,y,uv,A)$. Since $aby\in\mathcal{G}\setminus\mathcal{R}$, we have $a\notin A$ and $b\notin A$. But then $\mathcal{C}_4^3(ab,xy,uv)\subseteq\mathcal{G}$, a contradiction proving (13).

Thus, in order to finish the proof of the lemma, it is enough to show that $|\mathcal{R}|\leqslant|\mathcal{R}'|$. Inequality (12) implies that $|\mathcal{M}'|=10|\mathcal{A}|\geqslant|\mathcal{M}|$.

It remains to consider secondary pairs and triples. Let xy be an arbitrary secondary pair with $\mathcal{L}(x,y,uv,A)$ being true. Observe that xy cannot be a subset of some $B\in\mathcal{A}$. Indeed, otherwise $A\cap B=\{y\}$ by (12) while there are $u'v'\in B$ with $u'v'x, u'v'y\in\mathcal{G}$ by Lemma 8, giving $\mathcal{C}_4^3(u'v',xy,uv)\subseteq\mathcal{G}$, a contradiction. Therefore, we have $\mathcal{M}'\cap\mathcal{S}'=\emptyset$. (It also holds that $\mathcal{M}\cap\mathcal{S}=\emptyset$ but we do not need this fact.)

Claim 1. *Suppose that $\mathcal{L}(x,y,uv,A)$ holds for some x, y, uv, and A. Then there are no $u'v'$ and A' such that $A'\neq A$ and $\mathcal{L}(x,y,u'v',A')$ is true. Moreover, there are at most 2 choices of an unordered pair $u'v'$ satisfying $\mathcal{L}(x,y,u'v',A)$.*

Proof. If the first statement is false, then $A\cap A'=\{y\}$ by (12), and $\mathcal{C}_4^3(uv,xy,u'v')\subseteq\mathcal{G}$, a contradiction.

Suppose on the contrary to the second statement that the link graph $\mathcal{G}_x$ has at least three witness pairs inside the 4-element set $A\setminus y$. Three of these pairs form either a triangle $\{u_1u_2,u_1u_3,u_2u_3\}$, a star $\{u_0u_1,u_0u_2,u_0u_3\}$, or a path $\{u_1u_2,u_2u_3,u_3u_4\}$.

Suppose that we have the triangle. Let z be the unique vertex of $A\setminus u_1u_2u_3y$. Then every two of the triples u_1yz, u_2yz, and u_3yz share two vertices, so at most one of the triples can be missing from $\mathcal{G}$ (because $\mathcal{G}[A]\in\mathcal{K}_5^-$). By symmetry, assume that $u_1yz, u_2yz\in\mathcal{G}$. But then $\mathcal{C}_4^3(u_3x,u_1u_2,yz)\subseteq\mathcal{G}$, a contradiction.

Suppose that we have the 3-star. Like before, at least two of the triples u_1u_2y, u_1u_3y, and u_2u_3y are present in $\mathcal{G}$, say u_1u_2y and u_1u_3y. But then $\mathcal{C}_4^3(u_0x,u_2u_3,u_1y)\subseteq\mathcal{G}$, a contradiction.

Finally, we cannot have the 3-path for otherwise $\mathcal{C}_4^3(u_1u_2,yx,u_3u_4)\subseteq\mathcal{G}$. Claim 1 is proved. □

Let us define the auxiliary bipartite graph $\mathcal{H}$ with parts $\mathcal{S}$ and $\mathcal{S}'$, where for every satisfied predicate $\mathcal{L}(x,y,uv,A)$ we put an edge between xy and uvx. Note that $uvx\in\mathcal{G}$ is necessarily a secondary triple because it intersects $A\in\mathcal{A}$ in exactly two vertices, u and v. Also, we do not have to worry about multiple edges in $\mathcal{H}$ because if $\{xy,uvx\}\in\mathcal{H}$ then there is the unique A with $\mathcal{L}(x,y,uv,A)$ by Claim 1.

Let us show that for every edge $\{xy,uvx\}\in\mathcal{H}$ we have

$$d(xy)\leqslant d(uvx), \tag{14}$$

where d denotes the degree of a vertex in the graph $\mathcal{H}$.

Suppose first that there are no A' and $u'v'$ such that $\mathcal{L}(y,x,u'v',A')$ holds. Claim 1 implies that $d(xy)\leqslant 2$. On the other hand, pick the (unique) set $A\in\mathcal{A}$ with $A\cap uvx=uv$. Then there are at least two choices of $z\in A\setminus uv$ with $\mathcal{L}(x,z,uv,A)$. (Indeed, since $\mathcal{G}[A]\in\mathcal{K}_5^-$, at most one triple uvz with $z\in A\setminus uv$ can be missing from $\mathcal{G}$.) Thus $d(uvx)\geqslant 2$, as required.

It remains to assume that both $\mathcal{L}(x,y,uv,A)$ and $\mathcal{L}(y,x,u'v',A')$ hold for some A, A', and $u'v'$. By Claim 1, A and A' are uniquely determined while we have at most two choices for each of uv and $u'v'$. Hence $d(xy)\leqslant 4$. By considering a possible generalized 4-cycle $\mathcal{C}_4^3(uv,xy,u'v')$, we conclude that $uv\cap u'w'\neq\emptyset$, say $u=u'$. By (12), $A\cap A'=\{u\}$. Thus uvx intersects each of A and A' in exactly two vertices. By the argument of the previous paragraph, there are at least two choices of $z\in A\setminus uv$ satisfying $\mathcal{L}(x,z,uv,A)$ and at least two choices of $z'\in A'\setminus ux$ satisfying $\mathcal{L}(v,z',ux,A')$. Moreover, the four corresponding secondary pairs are pairwise different. (Two of them contain x but not v and two contain v but not x.) Hence, $d(uvx) \geqslant 4$. This proves (14).

Let us assign weights to the edges of $\mathcal{H}$ so that an edge $\{xy, uvx\} \in \mathcal{H}$ gets weight $1/d(xy)$. Then the total edge weight is equal to $|\mathcal{S}'|$. On the other hand, for every $uvx \in \mathcal{S}$, the sum of the weights of all edges incident to uvx is at least 1 by (14). Hence $|\mathcal{S}'| \geqslant |\mathcal{S}|$ and we obtain

$$|\mathcal{R}| \leqslant |\mathcal{M}| + |\mathcal{S}| \leqslant |\mathcal{M}'| + |\mathcal{S}'| = |\mathcal{R}'|,$$

proving the lemma. □

6. Upper bounds on $f_r(n)$ for general r

Proof of Theorem 1. Given $r \geqslant 3$, we choose some real σ (to be specified later) with $0 < \sigma < 1$. Let n be sufficiently large and let $\mathcal{G}$ be an arbitrary $\mathcal{C}_4^r$-free r-graph on an n-set V.

Let $C \subseteq V$ be a random uniformly distributed subset of size $(r-2)t$, where

$$t = \left\lfloor \frac{(1-\sigma)n}{r-2} \right\rfloor. \tag{15}$$

Let $S = V \setminus C$ and $s = |S|$. Take a random partition of C into $(r-2)$-sets $C_1, \dots, C_t$, all partitions being equally likely. Let $T = [t]$. (We assume that $[t] \cap V = \emptyset$.)

Define a 3-graph $\mathcal{H}$ on $S \cup T$ by including those triples abx such that $ab \in \binom{S}{2}$, $x \in T$, and $ab \cup C_x \in \mathcal{G}$. Since a permutation of V does not change the distribution of $(C, C_1, \dots, C_t)$, any two r-subsets of V are equally likely to contribute a triple to $\mathcal{H}$. This common probability is $\binom{s}{2}t/\binom{n}{r}$ because the complete r-graph $\binom{V}{r}$ would contribute exactly $\binom{s}{2}t$ triples. Let us choose $C, C_1, \dots, C_t$ so that $|\mathcal{H}|$ is at least its expected value, that is,

$$|\mathcal{H}| \geqslant |\mathcal{G}| \frac{\binom{s}{2}t}{\binom{n}{r}}. \tag{16}$$

Since every edge of $\mathcal{H}$ intersects T in exactly one vertex, that is,

$$|E \cap T| = 1, \quad \forall E \in \mathcal{H}, \tag{17}$$

it is straightforward to check that $\mathcal{C}_4^3 \not\subseteq \mathcal{H}$ (for otherwise $\mathcal{C}_4^r \subseteq \mathcal{G}$). Thus we can apply Lemma 10 to $\mathcal{H}$, with an arbitrary ordering $\preccurlyeq$, obtaining a subgraph $\mathcal{H}' \subseteq \mathcal{H}$ that satisfies Properties 1–4.

By (17), we have $\mathcal{D}(\mathcal{H}) \subseteq \binom{S}{2} \cup \binom{T}{2}$, that is, no xy with $x \in T$ and $y \in S$ can be a diagonal in $\mathcal{H}$. Hence, by Property 1 of Lemma 10 we have

$$h := |\mathcal{H}| - |\mathcal{H}'| \leqslant \left|\mathcal{D}(\mathcal{H}) \cap \binom{S}{2}\right| + \binom{t}{2}. \tag{18}$$

By Lemma 5, $\mathcal{D}_i(\mathcal{G}) = \emptyset$ except possibly for $i \in \{0, 1, 3\}$. Also, (17) implies that any subgraph $\mathcal{P}(xyz, u) \subseteq \mathcal{H}$ satisfies $u \in T$. Thus $\mathcal{D}_3(\mathcal{H}) \cap \binom{S}{2} = \emptyset$. By Property 3 of Lemma 10, we conclude that

$$\mathcal{E}(\mathcal{H}') \cap \binom{S}{2} \subseteq \mathcal{E}(\mathcal{H}) \cap \binom{S}{2} \subseteq \binom{S}{2} \setminus \mathcal{D}(\mathcal{H}).$$

Hence, by (18),

$$\left|\mathcal{E}(\mathcal{H}') \cap \binom{S}{2}\right| \leqslant \binom{s}{2} - \left|\mathcal{D}(\mathcal{H}) \cap \binom{S}{2}\right| \leqslant \binom{s}{2} + \binom{t}{2} - h. \tag{19}$$

Let us derive a contradiction by assuming that $uv \in \mathcal{E}(\mathcal{H}'_x) \cap \mathcal{E}(\mathcal{H}'_y)$ for some $uv \in \binom{S}{2}$ and two different $x, y \in T$. Choose witnesses c and d with $cu, cv \in \mathcal{H}'_x$ and $du, dv \in \mathcal{H}'_y$. Then $c, d \in S$, so they are different from x and y. If $c = d$, this gives $\mathcal{C}_4(uv, xy) \subseteq \mathcal{H}'_c$, a contradiction to Property 2 of Lemma 10. Otherwise, $\mathcal{C}_4^3(cx, uv, dy) \subseteq \mathcal{H}$, a contradiction again. Hence,

$$\mathcal{E}(\mathcal{H}'_x) \cap \mathcal{E}(\mathcal{H}'_y) \cap \binom{S}{2} = \emptyset, \quad \forall xy \in \binom{T}{2}. \tag{20}$$

Table 1

r	3	4	5	6	7	8
Optimal $\sigma(r)$	0.6388...	0.5233...	0.4570...	0.4119...	0.3786...	0.3525...
Upper bound on ϕ_r	1.7397...	1.7442...	1.7159...	1.6826...	1.6506...	1.6214...

We apply Lemma 4 to each link graph $\mathcal{H}'_x$ with $x \in T$. (Recall that $\mathcal{H}'_x$ is $\mathcal{C}_4$-free by Property 2 of Lemma 10.) We conclude by (17) and (20) that

$$2|\mathcal{H}'| - st = \sum_{x \in T} \left(2|\mathcal{H}'_x| - s\right) \leqslant \left|\mathcal{E}(\mathcal{H}') \cap \binom{S}{2}\right|.$$

This, (18), and (19) imply that

$$|\mathcal{H}| = |\mathcal{H}'| + h \leqslant \frac{\binom{s}{2} + \binom{t}{2} - h + st}{2} + h \leqslant \binom{s}{2} + \binom{t}{2} + \frac{st}{2}.$$

This, (16), equality $s = n - (r-2)t$, and (15) imply that, when r is fixed and $n \to \infty$,

$$\begin{aligned} |\mathcal{G}| &\leqslant |\mathcal{H}| \times \frac{\binom{n}{r}}{\binom{s}{2}t} \leqslant \left(\binom{s}{2} + \binom{t}{2} + \frac{st}{2}\right) \times \frac{\binom{n}{r}}{\binom{s}{2}t} \\ &= \left(\frac{\sigma^2(r^2 - 5r + 7) + \sigma(r-4) + 1}{r(r-2)(-\sigma^3 + \sigma^2)} + o(1)\right)\binom{n}{r-1}. \end{aligned} \tag{21}$$

If we set $\sigma = 1/\sqrt{r}$, then we obtain

$$|\mathcal{G}| \leqslant \left(1 + \frac{2}{\sqrt{r}} - \frac{2\sqrt{r} - 3}{(\sqrt{r} - 1)(r-2)\sqrt{r}} + o(1)\right)\binom{n}{r-1},$$

which implies that $\phi_r \leqslant 1 + 2/\sqrt{r}$. The bound $\phi_r \leqslant 7/4$ follows from the assignment $\sigma = 2/r$. □

Remark. The optimal $\sigma(r)$, the one that would minimize the coefficient at $\binom{n}{r-1}$ in (21), can be found by solving a cubic equation and, in fact, satisfies $\sigma(r) = (1 + o(1))/\sqrt{r}$ as $r \to \infty$. Table 1 lists the numerical values of the upper bound on ϕ_r given by the proof of Theorem 1 with the optimal $\sigma(r)$ for some small r.

7. The case $r = 3$

Proof of Theorem 2. Let $\mathcal{G}$ be an arbitrary $\mathcal{C}_4^3$-free 3-graph on an n-set V. First, we apply Lemma 11 to G to obtain an edge set $\mathcal{R} \subseteq \mathcal{G}$. Let $\mathcal{G}' = \mathcal{G} \setminus \mathcal{R}$ be the corresponding $\mathcal{K}_5^-$-free subgraph of $\mathcal{G}$. Let $m = |\mathcal{R}| + |\mathcal{D}(\mathcal{G}')|$. Lemma 11 ensures that

$$\left|\mathcal{E}(\mathcal{G}')\right| \leqslant \binom{n}{2} - m. \tag{22}$$

Next, we fix some ordering $\preccurlyeq$ of V and apply Lemma 10 to $\mathcal{G}'$ with respect to $\preccurlyeq$ to obtain a subgraph $\mathcal{H} \subseteq \mathcal{G}'$. We have $|\mathcal{G} \setminus \mathcal{H}| \leqslant m$.

For $xyz \in \binom{V}{3}$ let $P(xyz) = \{a \in V\colon\ \mathcal{P}(xyz, a) \subseteq \mathcal{G}'\}$ consist of those vertices $a \in V$ such that $axy, axz, ayz \in \mathcal{G}'$. For every pair $yz \in \binom{V}{2}$ we do the following. If there exists a vertex $x \in V$ such that $x \prec y$, $x \prec z$, and $|P(xyz)| \geqslant 2$, then let $x(yz)$ be this vertex x and let

$$S(yz) = \left\{a \in P(xyz)\colon\ ayz \in \mathcal{H}\right\} = P(xyz) \cap \mathcal{H}_{yz}.$$

(Note that if x exists, it is unique by Lemma 7 because $\mathcal{K}_5^- \nsubseteq \mathcal{G}'$.) Otherwise, we set $x(yz) = 0$ and $S(yz) = \emptyset$, where $0 \notin V$ is some fixed element with $0 \prec v$ for every $v \in V$. Also, we let $s(yz) = |S(yz)|$.

By Property 2 of Lemma 10, we have $\mathcal{D}(\mathcal{H})=\emptyset$. Let us show that

$$\mathcal{E}(\mathcal{H})\subseteq\mathcal{E}(\mathcal{G}')\cup\left(\bigcup_{yz\in\binom{V}{2}}\binom{S(yz)}{2}\right). \tag{23}$$

Suppose that $ab\in\mathcal{E}(\mathcal{H})\setminus\mathcal{E}(\mathcal{G}')$, say $auv, buv\in\mathcal{H}$ for some $u,v\in V$. Since $\mathcal{H}\subseteq\mathcal{G}'$, we have $\mathcal{E}(\mathcal{H})\subseteq\mathcal{E}(\mathcal{G}')\cup\mathcal{D}(\mathcal{G}')$. Thus $ab\in\mathcal{D}(\mathcal{G}')$. By Property 3 of Lemma 10, we have $ab\in\mathcal{D}_3(\mathcal{G}')$. By Lemma 5, there is a triple $xyz\in\binom{V}{3}$ with $ab\subseteq P(xyz)$. Without loss of generality assume that $x\prec y\prec z$. By Property 4 of Lemma 10, $uv=yz$. By the definition of $S(yz)$, we have $a,b\in S(yz)$, proving (23).

For $yz\in\binom{V}{2}$, recall that $\mathcal{H}_{yz}=\{x\in V\colon\ xyz\in\mathcal{H}\}$; let $h_{yz}=|\mathcal{H}_{yz}|$. Then, since $\mathcal{H}\subseteq\mathcal{G}$ is $\mathcal{C}_4^3$-free and every link graph $\mathcal{H}_x$ is C_4-free by Property 2 of Lemma 10, we conclude by Lemma 9, (22) and (23) that

$$\sum_{yz\in\binom{V}{2}}\binom{h_{yz}}{2}=\left|\mathcal{E}(\mathcal{H})\right|\leqslant\binom{n}{2}-m+\sum_{yz\in\binom{V}{2}}\binom{s(yz)}{2}. \tag{24}$$

Let $\mathcal{S}_i=\{yz\in\binom{V}{2}\colon\ s(yz)=i\}$, $s_i=|\mathcal{S}_i|$, and $\mathcal{S}_{\geqslant 3}=\bigcup_{i\geqslant 3}\mathcal{S}_i$.

Claim 1. *The following inequality holds*:

$$\sum_{yz\in\binom{V}{2}}h_{yz}\leqslant 2\binom{n}{2}-m+\sum_{i\geqslant 2}(i-1)s_i. \tag{25}$$

Proof. Let us maximize $\sigma=\sum_{yz\in\binom{V}{2}}h_{yz}$ over non-negative integers h_{yz}, given that (24) holds, $h_{yz}\geqslant s(yz)$ for every $yz\in\binom{V}{2}$, and

$$\sum_{yz\in\binom{V}{2}}\binom{h_{yz}}{2}\leqslant\binom{n}{2}. \tag{26}$$

(The last inequality holds for otherwise we get either a copy of $\mathcal{C}_4^3$ in $\mathcal{H}$ or a copy of C_4 in a link graph of $\mathcal{H}$, a contradiction.)

Take an optimal integer vector $\mathbf{h}=(h_{yz})_{yz\in\binom{V}{2}}$. Suppose first that $h_{yz}>s(yz)$ for some $yz\in\mathcal{S}_{\geqslant 3}$. By (26), there is a pair ab with $h_{ab}\leqslant 1$. We decrease h_{yz} by 1 and increase h_{ab} by 2. The left-hand side of (24) changes by

$$\binom{h_{ab}+2}{2}-\binom{h_{ab}}{2}+\binom{h_{yz}-1}{2}-\binom{h_{yz}}{2}\leqslant\binom{3}{2}-\binom{1}{2}+\binom{3}{2}-\binom{4}{2}=0.$$

Thus we still have a feasible solution while the sum of the entries of $\mathbf{h}$ strictly increases. This contradicts the optimality of $\mathbf{h}$. Thus we have $h_{yz}=s(yz)$ for all $yz\in\mathcal{S}_{\geqslant 3}$. We have

$$-\binom{n}{2}+\sum_{yz\in\binom{V}{2}}h_{yz}=\sum_{yz\in\binom{V}{2}}(h_{yz}-1)\leqslant\sum_{yz\in\binom{V}{2}\setminus\mathcal{S}_{\geqslant 3}}\binom{h_{yz}}{2}+\sum_{i\geqslant 3}(i-1)s_i.$$

Also, by (24),

$$\sum_{yz\in\binom{V}{2}\setminus\mathcal{S}_{\geqslant 3}}\binom{h_{yz}}{2}=\sum_{yz\in\binom{V}{2}}\binom{h_{yz}}{2}-\sum_{i\geqslant 3}\binom{i}{2}s_i\leqslant\left(\binom{n}{2}-m+\sum_{i\geqslant 2}\binom{i}{2}s_i\right)-\sum_{i\geqslant 3}\binom{i}{2}s_i.$$

Claim 1 follows from the last two inequalities. □

Each yz with $s(yz)\geqslant 2$ comes from some $\mathcal{P}(xyz, P(xyz))\subseteq\mathcal{G}'$ with $|P(xyz)|\geqslant s(yz)$ while every such $\mathcal{P}_i$-subgraph of $\mathcal{G}'$ with $i\geqslant 2$ gives at most one such pair yz because of the relation $x\prec y,z$ in the definition of $S(yz)$. Since $\mathcal{K}_5^- \not\subseteq \mathcal{G}'$, the private pairs of these $\mathcal{P}_i$-subgraphs are distinct from each other by Lemma 7 and all belong to $\mathcal{D}(\mathcal{G}')$. Thus

$$\sum_{i\geqslant 2}\left(3+\binom{i}{2}\right)s_i \leqslant \left|\mathcal{D}(\mathcal{G}')\right| = m - |\mathcal{R}|. \tag{27}$$

It is routine to see that $i-1 \leqslant \frac{1}{3}(3+\binom{i}{2})$ for every $i \geqslant 2$. Thus (25) and (27) imply that

$$\begin{aligned} 3|\mathcal{H}| &= \sum_{yz\in\binom{V}{2}} h_{yz} \leqslant 2\binom{n}{2} - m + \sum_{i\geqslant 2}(i-1)s_i \\ &\leqslant 2\binom{n}{2} - m + \frac{1}{3}\bigl(m-|\mathcal{R}|\bigr) = 2\binom{n}{2} - \frac{2m}{3} - \frac{|\mathcal{R}|}{3}. \end{aligned}$$

We conclude that

$$|\mathcal{G}| \leqslant m + |\mathcal{H}| \leqslant m + \frac{1}{3}\left(2\binom{n}{2} - \frac{2m}{3} - \frac{|\mathcal{R}|}{3}\right) \leqslant \frac{7m}{9} + \frac{2}{3}\binom{n}{2}.$$

Finally, $m = |\mathcal{D}(\mathcal{G}')| + |\mathcal{R}| \leqslant \binom{n}{2}$, giving the required. □

8. Concluding remarks

By analyzing the proof of Theorem 2, it should be possible to derive a contradiction from assuming that $f_3(n) = (13/9 + o(1))\binom{n}{2}$ for an infinite sequence of n. This would imply that there is a constant $c > 0$ such that $\phi_3 \leqslant 13/9 - c$. Unfortunately, a rigorous proof of this would be rather long and messy, especially if one tries to optimize the value of c. Therefore, we decided to settle for the current bound of 13/9, with a reasonably short and clear proof.

Acknowledgments

The authors thank the anonymous referees for helpful remarks.

References

[1] W.G. Brown, On graphs that do not contain a Thomsen graph, Canad. Math. Bull. 9 (1966) 281–289.
[2] C.R.J. Clapham, A. Flockhart, J. Sheehan, Graphs without four-cycles, J. Graph Theory 13 (1989) 29–47.
[3] P. Erdős, On sequences of integers no one of which divides the product of two others and on some related problems, Inst. Math. Mech. Univ. Tomsk 2 (1938) 74–82.
[4] P. Erdős, Some recent results on extremal problems in graph theory. Results, in: Theory of Graphs, Internat. Sympos., Rome, 1966, Gordon and Breach, New York, 1967, pp. 117–123 (in English); pp. 124–130 (in French).
[5] P. Erdős, Problems and results in combinatorial analysis, Congr. Numer. 19 (1977) 3–12.
[6] P. Erdős, A. Rényi, On a problem in the theory of graphs, Publ. Math. Inst. Hungar. Acad. Sci. 7 (1962) 623–641 (in Hungarian).
[7] P. Erdős, A. Rényi, V.T. Sós, On a problem in the theory of graphs, Studia Sci. Math. Hungar. 1 (1966) 215–235.
[8] Z. Füredi, Graphs without quadrilaterals, J. Combin. Theory Ser. A 34 (1983) 187–190.
[9] Z. Füredi, Hypergraphs in which all disjoint pairs have distinct unions, Combinatorica 4 (1984) 161–168.
[10] Z. Füredi, On the number of edges of quadrilateral-free graphs, J. Combin. Theory Ser. B 68 (1996) 1–6.
[11] H. Hanani, The existence and construction of balanced incomplete block designs, Ann. Math. Statist. 32 (1961) 361–386.
[12] H. Hanani, A balanced incomplete block design, Ann. Math. Statist. 36 (1965) 711.
[13] D. Mubayi, Some exact results and new asymptotics for hypergraph Turán numbers, Combin. Probab. Comput. 11 (2002) 299–309.
[14] D. Mubayi, Structure and stability of triangle-free set systems, Trans. Amer. Math. Soc. 359 (2007) 275–291.
[15] D. Mubayi, J. Verstraëte, A hypergraph extension of the bipartite Turán problem, J. Combin. Theory Ser. A 106 (2004) 237–253.
[16] D. Mubayi, J. Verstraëte, Minimal paths and cycles in set systems, European J. Combin. 28 (2007) 1681–1693.
[17] Y. Yuansheng, P. Rowlinson, On extremal graphs without four-cycles, Util. Math. 41 (1992) 204–210.