The number of triple systems without even cycles

Dhruv Mubayi* Lujia Wang†

March 19, 2018

Abstract

For $k \geqslant 4$, a loose k-cycle C_k is a hypergraph with distinct edges $e_1, e_2, \ldots, e_k$ such that consecutive edges (modulo k) intersect in exactly one vertex and all other pairs of edges are disjoint. Our main result is that for every even integer $k \geqslant 4$, there exists $c > 0$ such that the number of triple systems with vertex set $[n]$ containing no C_k is at most 2^{cn^2}. An easy construction shows that the exponent is sharp in order of magnitude.

Our proof method is different than that used for most recent results of a similar flavor about enumerating discrete structures, since it does not use hypergraph containers. One novel ingredient is the use of some (new) quantitative estimates for an asymmetric version of the bipartite canonical Ramsey theorem.

1 Introduction

An important theme in combinatorics is the enumeration of discrete structures that have certain properties. Within extremal combinatorics, one of the first influential results of this type is the Erdős-Kleitman-Rothschild theorem [26], which implies that the number of triangle-free graphs with vertex set $[n]$ is $2^{n^2/4+o(n^2)}$. This has resulted in a great deal of work on problems about counting the number of graphs with other forbidden subgraphs such as odd cycles [42], complete bipartite graphs [15, 16], octahedron graph [8] and other graphs with certain chromatic properties [6, 7, 27, 33, 50]; as well as similar question for other discrete structures [10, 11, 18, 19, 37, 48, 49, 51, 53]. In extremal graph theory, these results show that such problems are closely related to the corresponding extremal problems. More precisely, the *Turán* problem asks for the maximum number of edges in a (hyper)graph that does not contain a specific subgraph. For a given r-uniform hypergraph (henceforth r-graph) F, let $\mathrm{ex}_r(n, F)$ be the maximum number of edges among all r-graphs G on n vertices that contain no copy of F as a (not necessarily induced) subgraph. Henceforth we will call G an F-free r-graph. Write $\mathrm{Forb}_r(n, F)$ for the set of F-free r-graphs with vertex set $[n]$. Since each subgraph of an F-free r-graph is also F-free, we trivially obtain

*Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607, Email:mubayi@uic.edu. Research partially supported by NSF grant DMS-1300138.

†Department of Mathematics, Statistics, and Computer Science, University of Illinois, Chicago, IL 60607, Email:lwang203@uic.edu. Research partially supported by NSF grant DMS-1300138.

$|\mathrm{Forb}_r(n,F)| \geqslant 2^{\mathrm{ex}_r(n,F)}$ by taking subgraphs of an F-free r-graph on $[n]$ with the maximum number of edges. On the other hand for fixed r and F,

$$|\mathrm{Forb}_r(n,F)| \leqslant \sum_{i \leqslant \mathrm{ex}_r(n,F)} \binom{\binom{n}{r}}{i} = 2^{O(\mathrm{ex}_r(n,F)\log n)},$$

so the issue at hand is the factor $\log n$ in the exponent. The work of Erdős-Kleitman-Rothschild [26] and Erdős-Frankl-Rödl [27] for graphs, and Nagle-Rödl-Schacht [47] for hypergraphs (see also [46] for the case $r = 3$) improves the upper bound above to obtain

$$|\mathrm{Forb}_r(n,F)| = 2^{\mathrm{ex}_r(n,F)+o(n^r)}.$$

Although much work has been done to improve the exponent above (see [1, 6, 7, 8, 33, 36, 50] for graphs and [10, 11, 22, 49, 13, 52] for hypergraphs), this is a somewhat satisfactory state of affairs when $\mathrm{ex}_r(n,F) = \Omega(n^r)$ or F is not r-partite.

In the case of r-partite r-graphs, the corresponding questions appear to be more challenging since the tools used to address the case $\mathrm{ex}_r(n,F) = \Omega(n^r)$ like the regularity lemma are not applicable. The major open problem here when $r = 2$ is to prove that

$$|\mathrm{Forb}_r(n,F)| = 2^{O(\mathrm{ex}_r(n,F))}.$$

The two cases that have received the most attention are for $r = 2$ (graphs) and $F = C_{2l}$ or $F = K_{s,t}$. Classical results of Bondy-Simonovits [17] and Kovári-Sós-Turán [38] yield $\mathrm{ex}_2(n, C_{2l}) = O(n^{1+1/l})$ and $\mathrm{ex}_2(n, K_{s,t}) = O(n^{2-1/s})$ for $2 \leqslant s \leqslant t$, respectively. Although it is widely believed that these upper bounds give the correct order of magnitude, this is not known in all cases. Hence the enumerative results sought in these two cases were

$$|\mathrm{Forb}_2(n, C_{2l})| = 2^{O(n^{1+1/l})} \qquad \text{and} \qquad |\mathrm{Forb}_2(n, K_{s,t})| = 2^{O(n^{2-1/s})}.$$

In 1982, Kleitman and Winston [34] proved that $|\mathrm{Forb}_2(n, C_4)| = 2^{O(n^{3/2})}$ which initiated a 30-year effort on searching for generalizations of the result to complete bipartite graphs and even cycles. Kleitman and Wilson [35] proved similar results for C_6 and C_8 in 1996 by reducing to the C_4 case. Finally, Morris and Saxton [44] recently proved that $|\mathrm{Forb}_2(n, C_{2l})| = 2^{O(n^{1+1/l})}$ and Balogh and Samotij [15, 16] proved that $|\mathrm{Forb}_2(n, K_{s,t})| = 2^{O(n^{2-1/s})}$ for $2 \leqslant s \leqslant t$. Both these results used the hypergraph container method (developed independently by Saxton and Thomason [52], and by Balogh-Morris-Samotij [13]) which has proved to be a very powerful technique in extremal combinatorics. For example, [13] and [52] reproved $|\mathrm{Forb}_r(n,F)| = 2^{\mathrm{ex}_r(n,F)+o(n^r)}$ using containers.

There are very few results in this area when $r > 2$ and $\mathrm{ex}_r(n,F) = o(n^r)$. The only cases solved so far are when F consists of just two edges that intersect in at least t vertices [9], or when F consists of three edges such that the union of the first two is equal to the third [12] (see also [4, 5, 23, 24] for some related results). These are natural points to begin these investigations since the corresponding extremal problems have been studied deeply.

Recently, Kostochka, the first author and Verstraëte [39, 40, 41], and independently, Füredi and Jiang [30] (see also [31]) determined the Turán number for several other families of

r-graphs including paths, cycles, trees, and expansions of graphs. These hypergraph extremal problems have proved to be quite difficult, and include some longstanding conjectures. Guided and motivated by these recent developments on the extremal number of hypergraphs, we consider the corresponding enumeration problems focusing on the case of cycles.

Definition 1 *For each integer $k \geqslant 3$, a k-cycle C_k is a hypergraph with distinct edges $e_1, \ldots, e_k$ and distinct vertices $v_1, \ldots, v_k$ such that $e_i \cap e_{i+1} = \{v_i\}$ for all $1 \leqslant i \leqslant k-1$, $e_1 \cap e_k = \{v_k\}$ and $e_i \cap e_j = \emptyset$ for all other pairs i, j.*

Sometimes we refer to C_k as a *loose* or *linear* cycle. To simplify notation, we will omit the parameter r when the cycle C_k is a subgraph of an r-graph.

Since $\mathrm{ex}_r(n, C_k) = O(n^{r-1})$, we obtain the upper bound

$$|\mathrm{Forb}_r(n, C_k)| = 2^{O(n^{r-1}\log n)}$$

when r and k are fixed and $n \to \infty$. Our main result is the following theorem, which improves this upper bound.

Theorem 2 (Main Result) *For every even integer $k \geqslant 4$, there exists $c = c(k)$, such that*

$$|\mathrm{Forb}_3(n, C_k)| < 2^{c\,n^2}.$$

Since trivially $\mathrm{ex}_r(n, C_k) = \Omega(n^{r-1})$ for all $r \geqslant 3$, we obtain $|\mathrm{Forb}_3(n, C_k)| = 2^{\Theta(n^2)}$ when k is even. We conjecture that a similar result holds for $r > 3$ and cycles of odd length.

Conjecture 3 [1] *For fixed $r \geqslant 3$ and $k \geqslant 3$ we have $|\mathrm{Forb}_r(n, C_k)| = 2^{\Theta(n^{r-1})}$.*

Almost all recent developments in this area have relied on the method of hypergraph containers that we mentioned above. What is perhaps surprising about the current work is that the proofs do not use hypergraph containers. Instead, our methods employ old and new tools in extremal (hyper)graph theory. The old tools include the extremal numbers for cycles modulo h and results about decomposing complete r-graphs into r-partite ones, and the new tools include the analysis of the shadow for extremal hypergraph problems and quantitative estimates for the bipartite canonical Ramsey problem.

1.1 Definitions and notations

Throughout this paper, we let $[n]$ denote the set $\{1, 2, \ldots, n\}$. Write $\binom{X}{r} = \{S \subset X : |S| = r\}$ and $\binom{X}{\leqslant r} = \{S \subset X : |S| \leqslant r\}$. For $X \subset [n]$, an *r-uniform hypergraph* or *r-graph* H on vertex set X is a collection of r-element subsets of X, i.e. $H \subset \binom{X}{r}$. The vertex set X is denoted by $V(H)$. The r-sets contained in H are *edges*. The *size* of H is $|H|$.

[1] Recently, Conjecture 3 was proved by Balogh, Narayanan and Skokan [14] using the container method and Ferber, McKinley and Samotij [32] proved a more general result.

Given $S \subset V(H)$, the *neighborhood* $N_H(S)$ of S is the set of all $T \subset V(H) \setminus S$ such that $S \cup T \in H$. The *codegree* of S is $d_H(S) = |N_H(S)|$. When the underlying hypergraph is clear from context, we may omit the subscripts in these definitions and write $N(S)$ and $d(S)$ for simplicity. The *sub-edges* of H are the $(r-1)$-subsets of $[n]$ with positive codegree in H. The set of all sub-edges of H is called the *shadow* of H, and is denoted ∂H.

An *r-partite r-graph* H is an r-graph with vertex set $\bigsqcup_{i=1}^{r} V_i$ (the V_is are pairwise disjoint), and every $e \in H$ satisfies $|e \cap V_i| = 1$ for all $i \in [r]$. When all such edges e are present, H is called a *complete r-partite r-graph.* When $|V_i| = s$ for all $i \in [r]$, a complete r-partite r-graph H is said to be *balanced*, and denoted $K_{s:r}$.

For each integer $k \geqslant 1$, a *(loose, or linear) path of length k* denoted by P_k, is a collection of k edges $e_1, e_2, \ldots, e_k$ such that $|e_i \cap e_j| = 1$ if $i = j+1$, and $e_i \cap e_j = \emptyset$ otherwise.

We will often omit floors and ceilings in our calculations for ease of notation and all logs will have base 2.

2 Proof of the main result

We begin with a sketch of the proof of Theorem 2. The first step is to partition each of the hypergraphs that we must count into a bounded number of sub-hypergraphs, each of which can be encoded by an edge-colored graph (by choosing one pair from each hyperedge, and coloring that with the third vertex, in such a way that each pair is chosen at most once). This follows easily from a (straightforward) lemma of Kostochka, Mubayi and Verstraëte [39] which states that a 3-graph with no loose cycle of a given length contains a pair of vertices of bounded codegree.

The main task is therefore to bound the number of 3-graphs with no loose cycle of length $2l$ that can be encoded by an edge-colored graph G. Our strategy is to first partition the edges of G into complete bipartite graphs $K_{s_1,s_1}, \ldots, K_{s_m,s_m}$, with each $s_i \leqslant \log n$, such that $\sum_{i=1}^{m} s_i = O(n^2/\log n)$; this can be done greedily, using the Kővári–Sós–Turán Theorem [38]; there are at most $2^{o(n^2)}$ choices for the sequence $(m, s_1, \ldots, s_m)$.

The problem is therefore reduced to counting the number of edge-colorings of a complete bipartite graph $K_{s,s}$ such that the associated 3-graph contains no loose cycle of length $2l$. Theorem 5 proves that there are at most $2^{O(s^2)} n^{O(s)}$ such colorings. This suffices to prove our main theorem, since we obtain the following upper bound for the number of 3-graphs with no loose cycle of length $2l$:

$$\begin{aligned}\left(\sum_{m,s_1,\ldots,s_m} \prod_{i=1}^{m} 2^{s_i^2} n^{s_i} \right)^{O(1)} &= \left(2^{o(n^2)} \cdot 2^{\sum_i s_i^2 + \sum_i s_i \log n} \right)^{O(1)} \\ &= \left(2^{o(n^2)} 2^{n^2 + O((n^2/\log n)\cdot \log n)} \right)^{O(1)} = 2^{O(n^2)}.\end{aligned}$$

2.1 Main technical statement

Given a graph G with $V(G) \subset [n]$, a *coloring function* is a function $\chi : G \to [n]$ such that $\chi(e) = z_e \in [n] \setminus e$ for every $e \in G$. We call z_e the *color* of e. The vector of colors $N_G = (z_e)_{e \in G}$ is called an *edge-coloring* of G. The pair (G, N_G) is an *edge-colored* graph. A *color class* is the set of all edges that receive the same color.

Given G, each edge-coloring N_G defines a 3-graph $H(N_G) = \{e \cup \{z_e\} : e \in G\}$, called the *extension of G by N_G*. When there is only one coloring that has been defined, we also use the notation $G^* = H(N_G)$ for the extension. Observe that any subgraph $G' \subset G$ also admits an extension by N_G, namely, $G'^* = \{e \cup \{z_e\} : e \in G'\} \subset G^*$. If $G' \subset G$ and $\chi|_{G'}$ is one-to-one, then G' is called *rainbow colored*. If a rainbow colored G' further satisfies that $z_e \notin V(G')$ for all $e \in G'$, then G' is said to be *strongly rainbow colored*. Note that a strongly rainbow colored graph $C_k \subset G'$ gives rise to 3-graph C_k in $G'^* \subset G^*$.

Definition 4 *For $k \geqslant 3, s \geqslant 1$, let $f(n,k,s)$ be the number of edge-colored complete bipartite graphs $G = K_{s,s}$ with $V(G) \subset [n]$, whose extension G^* is C_k-free.*

The function $f(n,k,s)$ allows us to encode 3-graphs, and our main technical theorem gives an upper bound for this function.

Theorem 5 *Fix an even integer $k \geqslant 4$. Then*

$$f(n,k,s) = 2^{O(s \log n + s^2)}$$

Note that the trivial upper bound is $f(n,k,s) \leqslant n^{2s+s^2} \sim 2^{s^2 \log n}$ (first choose $2s$ vertices, then color each of its s^2 edges using an arbitrary vertex from $[n]$). The proof of Theorem 5 will be given in Sections 3–6.

2.2 A non-bipartite version of Theorem 5

Chung-Erdős-Spencer [20] and Bublitz [3] proved that the complete graph K_n can be decomposed into balanced complete bipartite graphs such that the sum of the sizes of the vertex sets in these bipartite graphs is at most $O(n^2/\log n)$. See also [57, 45] for some generalizations and algorithmic consequences. We state the result without proof as follows.

Theorem 6 *Let $n \geqslant 2$. Then, each n-vertex graph can be decomposed into complete bipartite graphs $K_{s_i,s_i}, i = 1, \ldots, m$, with $s_i \leqslant \log n$ and $\sum_{i=1}^{m} s_i = O(n^2/\log n)$.*

Theorem 5 is about the number of ways to edge-color complete bipartite graphs with parts of size s and vertex set in $[n]$. Next, we use Theorems 5 and 6 to prove a related statement where we do not require the bipartite condition and the restriction to s vertices.

Definition 7 *For $k \geqslant 4$ and even, let $g(n,k)$ be the number of edge-colored graphs G with $V(G) \subset [n]$ such that the extension G^* is C_k-free.*

Lemma 8 *For fixed $k \geqslant 4$ and even,*

$$g(n,k) = 2^{O(n^2)}.$$

Proof. Given graph G, by applying Theorem 6, we may decompose G into balanced complete bipartite graphs $K_{s_1,s_1}, \dots, K_{s_m,s_m}$, with $s_i \leqslant \log n$ and $\sum_{i=1}^m s_i = O(n^2/\log n)$. Then we trivially deduce the following two facts.

- From the second inequality, we have $m = O(n^2/\log n)$.
- Using the fact that these copies of $K_{s_i:s_i}$ are edge disjoint, we have

$$\sum_{i=1}^m s_i^2 \leqslant \binom{n}{2} < n^2.$$

Therefore, to construct an edge-colored G, we need to first choose a sequence of positive integers $(m, s_1, \dots, s_m)$ such that $m \leqslant c_1 n^2/\log n$, with some fixed $c_1 > 0$ and $s_i \leqslant \log n$ for all i. More formally, let

$$S_n = \{(m, s_1, s_2, \dots, s_m) : m \leqslant c_1 n^2/\log n, 1 \leqslant s_i \leqslant \log n, 1 \leqslant i \leqslant m\}.$$

Then

$$|S_n| \leqslant \frac{c_1 n^2}{\log n} (\log n)^{\frac{c_1 n^2}{\log n}} = 2^{\log\left(\frac{c_1 n^2}{\log n}\right) + \frac{c_1 n^2 \log(\log n)}{\log n}} \leqslant 2^{o(n^2)}. \tag{1}$$

Next, we sequentially construct an edge-colored K_{s_i,s_i} for each $i \in [m]$. Since G^* is C_k-free, $K^*_{s_i,s_i}$ is C_k-free. Writing v for a vector $(m, s_1, \dots, s_m)$ and applying Theorem 5 yields

$$g(n,k) \leqslant \sum_{v \in S_n} \prod_{i=1}^m f(n,k,s_i) \leqslant \sum_{v \in S_n} \prod_{i=1}^m 2^{O(s_i \log n + s_i^2)} \leqslant \sum_{v \in S_n} 2^{O(\sum_{i=1}^m s_i \log n + s_i^2)}.$$

By Theorem 6 and (1), this is at most

$$\sum_{v \in S_n} 2^{O((n^2/\log n) \cdot \log n + n^2)} = 2^{O(n^2)},$$

and the proof is complete. □

2.3 Proof of Theorem 2

A crucial statement that we use in our proof is that any r-graph such that every sub-edge has high codegree contains rich structures, including cycles. This was explicitly proved in [39] and we cite their following result.

Lemma 9 (Lemma 3.2 in [39]) *For $r, k \geqslant 3$, if all sub-edges of an r-graph H have codegree greater than rk, then $C_k \subset H$.*

Now we have all the ingredients to complete the proof of our main result.

Proof of Theorem 2. Starting with any 3-graph H on $[n]$ with $C_k \not\subset H$, we claim that there exists a sub-edge with codegree at most $3k$. Indeed, otherwise all sub-edges of H will have codegree more than $3k$, and then by Lemma 9 we obtain a $C_k \subset H$. Let e' be the sub-edge of H with $0 < d_H(e') \leqslant 3k$ such that it has smallest lexicographic order among all such sub-edges. Delete all edges of H containing e' from H (i.e. delete $\{e \in H : e' \subset e\}$). Repeat this process of "searching and deleting" in the remaining 3-graph until there are no such sub-edges. We claim that the remaining 3-graph must have no edges at all. Indeed, otherwise we get a nonempty subgraph all of whose sub-edges have codegree greater than $3k$, and again by Lemma 9, we obtain a $C_k \subset H$.

Given any C_k-free 3-graph H on $[n]$, the algorithm above sequentially decomposes H into a collection of sets of at most $3k$ edges who share a sub-edge (a pair of two vertices) in common. We regard the collection of these pairs as a graph G. Moreover, for each edge $e \in G$, let N_e be the set of vertices $v \in V(H)$ such that $e \cup \{v\}$ is an edge of H at the time e was chosen. So $N_e \in \binom{[n]\setminus e}{\leqslant 3k}$, for all $e \in G$. Thus, we get a map

$$\phi : \mathrm{Forb}_3(n, C_k) \longrightarrow \left\{ (G, N_G) : G \subset \binom{[n]}{2}, N_G = \left(N_e \in \binom{[n] \setminus e}{\leqslant 3k} : e \in G \right) \right\}.$$

We observe that ϕ is injective. Indeed,

$$\phi^{-1}((G, N_G)) = H(N_G) = \{e \cup \{z_e\} : e \in G, z_e \in N_e\},$$

therefore $|\mathrm{Forb}_3(n, C_k)| = |\phi(\mathrm{Forb}_3(n, C_k))|$. Let $P = \phi(\mathrm{Forb}_3(n, C_k))$ which is the set of all pairs (G, N_G) such that $H(N_G)$ is C_k-free. Next we describe our strategy for upper bounding $|P|$.

For each pair $(G, N_G) \in P$ and $e \in G$, we pick exactly one $z_e^1 \in N_e$. Thus we get a pair (G_1, N_{G_1}), where $G_1 = G$, and $N_{G_1} = (z_e^1 : e \in G_1)$. Then, delete z_e^1 from each N_e, let $G_2 = \{e \in G_1 : N_e \setminus \{z_e^1\} \neq \emptyset\}$ and pick $z_e^2 \in N_e \setminus \{z_e^1\}$ to get the pair (G_2, N_{G_2}). For $2 \leqslant i < 3k$, we repeat this process for G_i to obtain G_{i+1}. Since each N_{G_i} contains only singletons, the pair (G_i, N_{G_i}) can be regarded as an edge-colored graph. Note that we may get some empty G_is. This gives us a map

$$\psi : P \longrightarrow \left\{ (G_1, \ldots, G_{3k}) : G_i \subset \binom{[n]}{2} \text{ is edge-colored for all } i \in [3k] \right\}.$$

Moreover, it is almost trivial to observe that ψ is injective, since if $y \neq y'$, then either the underlying graphs of y and y' differ, or the graphs are the same but the color sets differ. In both cases one can easily see that $\psi(y) \neq \psi(y')$. Again, we let $Q = \psi(P)$.

Note that $k \geqslant 4$ and even, by Lemma 8, we have

$$|\mathrm{Forb}_3(n, C_k)| = |P| = |Q| \leqslant \prod_{i=1}^{3k} g(n, k) = \prod_{i=1}^{3k} 2^{O(n^2)} = 2^{O(n^2)}.$$

$\square$

3 Proof of Theorem 5

The rest of the paper is devoted to the proof of Theorem 5. For simplicity of presentation, we write $k = 2l$ where $l \geqslant 2$. We first state our two main lemmas about edge-coloring bipartite graphs then give a proof of Theorem 5.

Lemma 10 *Let $l \geqslant 2, s, t \geqslant 1$, $G = K_{s,t}$ be an edge-colored complete bipartite graph with $V(G) \subset [n]$ and $Z = \{z_e : e \in G\} \subset [n]$ be the set of all colors. If G contains no strongly rainbow colored C_{2l}, i.e. the 3-uniform extension G^* of G is C_{2l}-free, then $|Z| < 2l(s+t)$.*

Lemma 11 *Let $l \geqslant 2, s, t \geqslant 1$, $D = D_l = (4l)2^{(4l)^7}$, $Z \subset [n]$ with $|Z| < 2l(s+t)$, $G = K_{s,t}$ be a complete bipartite graph with vertex set $V(G) \subset [n]$. Then the number of ways to edge-color G with colors from Z such that the extension G^* contains no C_{2l}, is at most $D^{(s+t)^2}$.*

The proofs of these lemmas require several new ideas which will be presented in the rest of the paper. Here we quickly show that they imply Theorem 5.

Proof of Theorem 5. Recall that $l \geqslant 2$, and that $f(n, 2l, s)$ is the number of edge-colored copies of $K_{s,s}$ whose vertex set lies in $[n]$ and whose (3-uniform) extension is C_{2l}-free. To obtain such a copy of $K_{s,s}$, we first choose from $[n]$ its $2s$ vertices, then its at most $4ls$ colors by Lemma 10 and finally we color this $K_{s,s}$ by Lemma 11. This yields

$$f(n, 2l, s) \leqslant n^{2s+4ls} D^{(2s)^2} \leqslant 2^{5ls \log n + 4s^2 \log D} = 2^{O(s \log n + s^2)},$$

where the second inequality holds since $l \geqslant 2$. □

4 Proof of Lemma 10

In this section we prove Lemma 10. Our main tool is an extremal result about cycles modulo h in a graph. This problem has a long history, beginning with a Conjecture of Burr and Erdős that was solved by Bollobás [2] in 1976, see also [55, 25, 43, 54]. In particular, we need the following lemma (see Diwan [25]) whose idea is based on considering the longest path in G and the neighbors of the two endpoints of the path.

Lemma 12 *If G is an n-vertex graph with at least $(h+1)n$ edges, then G contains a cycle of length 2 modulo h.*

Recall that a rainbow colored cycle C_k is a copy of C_k with vertex set $V(C_k)$ in $[n]$ whose edges receives all distinct colors (where colors are vertices in $[n]$); whereas a strongly colored cycle C_k is rainbow colored and the set of all its colors is disjoint from its vertex set $V(C_k)$.

Lemma 13 *Let integers $l \geqslant 2$, $s, t \geqslant 1$, $G = K_{s,t}$ with $V(G) \subset [n]$ be edge-colored. If G contains a strongly rainbow colored cycle of length 2 $\pmod{2l-2}$, then G contains a strongly rainbow colored C_{2l}.*

Proof. Let us assume that C is the shortest strongly rainbow colored cycle of length 2 modulo $2l-2$ in G. Then C has at least $2l$ edges. We claim that C is a C_{2l}. Suppose not, let e be a chord of C (such a chord exists as G is complete bipartite), such that C is cut up into two paths P_1 and P_2 by the two endpoints of e, and $|P_1| = 2l-1$. Let Z_1, Z_2 be the set of their colors respectively. If the color $z_e \notin Z_1 \cup V(P_1) \setminus e$, then $P_1 \cup e$ is a strongly rainbow colored cycle of length $2l$, a contradiction. Therefore $z_e \in Z_1 \cup V(P_1) \setminus e$, but then $z_e \notin Z_2 \cup V(P_2) \setminus e$, yielding a shorter strongly rainbow colored cycle $P_2 \cup e$ of length 2 modulo $2l-2$, a contradiction. □

We now have all the necessary ingredients to prove Lemma 10.

Proof of Lemma 10. Suppose that $|Z| \geqslant 2l(s+t)$. Then $|Z \setminus V(G)| \geqslant (2l-1)(s+t)$. For each color v in $Z \setminus V(G)$, pick an edge e of G with color v. We obtain a strongly rainbow colored subgraph G' of G with at least $(2l-1)(s+t)$ edges. Lemma 12 guarantees the existence of a rainbow colored cycle of length 2 modulo $2l-2$ in G'. By construction, this cycle is strongly rainbow. Lemma 13 then implies that there is a strongly rainbow colored C_{2l} in G. □

5 Proof of Lemma 11

Our proof of Lemma 11 is inspired by the methods developed in [39]. The main idea is to use the bipartite canonical Ramsey theorem. In order to use this approach we need to develop some new quantitative estimates for an asymmetric version of the bipartite canonical Ramsey theorem.

5.1 Canonical Ramsey theory

In this section we state and prove the main result in Ramsey theory that we will use to prove Lemma 11. We are interested in counting the number of edge-colorings of a bipartite graph, such that the (3-uniform) extension contains no copy of C_{2l}. The canonical Ramsey theorem allows us to find nice colored structures that are easier to work with. However, the quantitative aspects are important for our application and consequently we need to prove various bounds for bipartite canonical Ramsey numbers. We begin with some definitions.

Let G be a bipartite graph on vertex set with bipartition $X \sqcup Y$. For any subsets $X' \subset X$, $Y' \subset Y$, let $E_G(X', Y') = G[X' \sqcup Y'] = \{xy \in G : x \in X', y \in Y'\}$, and $e_G(X', Y') = |E_G(X', Y')|$. If X' contains a single vertex x, then $E_G(\{x\}, Y')$ will be simply written as $E_G(x, Y')$. The subscript G may be omitted if it is obvious from context.

Definition 14 *Let G be an edge-colored bipartite graph with $V(G) = X \sqcup Y$.*

- *G is* monochromatic *if all edges in $E(X, Y)$ are colored by the same color.*
- *G is* weakly X-canonical *if $E(x, Y)$ is monochromatic for each $x \in X$.*

- *G is X-*canonical *if it is* weakly X-canonical *and for all distinct $x, x' \in X$ the colors used on $E(x, Y)$ and $E(x', Y)$ are all different.*

In all these cases, the color z_x of the edges in $E(x, Y)$ is called a canonical color.

Lemma 15 *Let $G = K_{a,b}$ be an edge-colored complete bipartite graph with bipartition $A \sqcup B$, with $|A| = a, |B| = b$. If G is weakly A-canonical, then there exists a subset $A' \subset A$ with $|A'| = \sqrt{a}$ such that $G[A' \sqcup B] = K_{\sqrt{a},b}$ is A'-canonical or monochromatic.*

Proof. Take a maximal subset A' of A such that the coloring on $E(A', B)$ is A'-canonical. If $|A'| \geqslant \sqrt{a}$, then we are done. So, we may assume that $|A'| < \sqrt{a}$. By maximality of A', there are less then $\sqrt{a}$ canonical colors. By the pigeonhole principle, there are at least $|A|/|A'| \geqslant a/\sqrt{a} = \sqrt{a}$ vertices of A sharing the same canonical color, which gives a monochromatic $K_{\sqrt{a},b}$. □

Our next lemma guarantees that in an "almost" rainbow colored complete bipartite graph, there exists a rainbow complete bipartite graph.

Lemma 16 *For any integer $c \geqslant 2$, and $p > c^4$, if $G = K_{p,p}$ is an edge-colored complete bipartite graph, in which each color class is a matching, then G contains a rainbow colored $K_{c,c}$.*

Proof. Let $A \sqcup B$ be the vertex set of G. Pick two c-sets X, Y from A and B respectively at random with uniform probability. For any pair of monochromatic edges e, e', the probability that they both appear in the induced subgraph $E(X, Y)$ is

$$\left(\frac{\binom{p-2}{c-2}}{\binom{p}{c}}\right)^2 = \left(\frac{c(c-1)}{p(p-1)}\right)^2.$$

On the other hand, the total number of pairs of monochromatic edges is at most $p^3/2$, since every color class is a matching. Therefore the union bound shows that, when $p > c^4$, the probability that there exists a monochromatic pair of edges in $E(X, Y)$ is at most

$$\frac{p^3}{2}\left(\frac{c(c-1)}{p(p-1)}\right)^2 < \frac{pc^4}{2(p-1)^2} < 1.$$

Consequently, there exists a choice of X and Y such that the $E(X, Y)$ contains no pair of monochromatic edges. Such an $E(X, Y)$ is a rainbow colored $K_{c,c}$. □

Now we are ready to prove the main result of this section which is a quantitative version of a result from [40]. Note that the edge-coloring in this result uses an arbitrary set of colors. Since the conclusion is about "rainbow" instead of "strongly rainbow", it is not essential to have the set of colors disjoint from the vertex set of the graph.

Theorem 17 (Asymmetric bipartite canonical Ramsey theorem) *For any integer $l \geqslant 2$, there exists real numbers $\epsilon = \epsilon(l) > 0, s_0 = s_0(l) = 2^{(4l)^7}$, such that if $G = K_{s,t}$ is an edge-colored complete bipartite graph on vertex set $X \sqcup Y$ with $|X| = s, |Y| = t$ with $s > s_0$ and $s/\log s < t \leqslant s$, then one of the following holds:*

- *G contains a rainbow colored $K_{4l,4l}$,*
- *G contains a $K_{q,2l}$ on vertex set $Q \sqcup R$, with $|Q| = q, |R| = 2l$ that is Q-canonical, where $q = s^\epsilon$,*
- *G contains a monochromatic $K_{q,2l}$ on vertex set $Q \sqcup R$, with $|Q| = q, |R| = 2l$, where $q = s^\epsilon$.*

Note that in the last two cases, it could be $Q \subset X, R \subset Y$ or the other way around.

Proof. We will show that $\epsilon = 1/18l$. First, fix a subset Y' of Y with $|Y'| = t^{1/4l}$ and let

$$W = \left\{x \in X : \text{there exists a } Y'' \in \binom{Y'}{2l} \text{ such that } E_G(x, Y'') \text{ is monochromatic}\right\}.$$

If $|W| > s/2l$, then the number of $Y'' \in \binom{Y'}{2l}$ such that $E_G(x, Y'')$ is monochromatic for some x (with repetition) is greater than $s/2l$. On the other hand, $|\binom{Y'}{2l}| < |Y'|^{2l} = \sqrt{t}$. By the pigeonhole principle, there exists a $Y'' \in \binom{Y'}{2l}$ such that at least

$$\frac{s}{2l\sqrt{t}} \geqslant \frac{s}{2l\sqrt{s}} \geqslant s^{1/3}$$

vertices x have the property that $E_G(x, Y'')$ is monochromatic. Let Q_1 be a set of $s^{1/3}$ such x. Then we obtain a weakly Q_1-canonical $K_{s^{1/3},2l}$ on $Q_1 \sqcup Y''$ which, by Lemma 15, contains a canonical or monochromatic $K_{s^{1/6},2l}$. Since $\epsilon < 1/6$, this contains a $K_{s^\epsilon,2l}$ as desired.

We may now assume that $|W| \leqslant s/2l$. By definition of W and the pigeonhole principle, $E_G(x, Y')$ contains at least $|Y'|/2l$ (distinct) colors for every $x \in X \setminus W$. Hence, for each $x \in X \setminus W$ we can take $|Y'|/2l$ distinctly colored edges from $E(x, Y')$ to obtain a subgraph G' of G on $(X \setminus W) \sqcup Y'$ with $|X \setminus W||Y'|/2l$ edges.

Pick a subset $X' \subset X \setminus W$ with $|X'| = s^{1/16l^2}$ and $e_{G'}(X', Y') \geqslant |X'||Y'|/2l$. This is possible by an easy averaging argument. Let

$$Z = \left\{y \in Y' : \text{there exists an } X'' \in \binom{X'}{2l} \text{ such that } E_{G'}(X'', y) \text{ is monochromatic}\right\}.$$

If $|Z| > |Y'|/20l$, then the number of $X'' \in \binom{X'}{2l}$ such that $E_{G'}(X'', y)$ is monochromatic for some y (with repetition) is greater than $|Y'|/20l$. On the other hand, $|\binom{X'}{2l}| < |X'|^{2l} = s^{1/8l}$. By the pigeonhole principle, there exists a $X'' \in \binom{X'}{2l}$ such that at least

$$\frac{|Y'|}{20ls^{1/8l}} = \frac{t^{1/4l}}{20ls^{1/8l}} \geqslant \frac{s^{1/4l}}{(\log s)^{1/4l}20ls^{1/8l}} \geqslant s^{1/9l} = s^{2\epsilon}$$

vertices y have the property that $E_{G'}(X'', y)$ is monochromatic. Let Q_2 be a set of $s^{2\epsilon}$ such y. We find a weakly Q_2-canonical $K_{2l,s^{2\epsilon}}$ on $X'' \sqcup Q_2$. Again, by Lemma 15, a copy of $K_{2l,s^{\epsilon}}$ that is monochromatic or canonical is obtained.

Finally, we may assume that $|Z| \leqslant |Y'|/20l$. Then

$$\begin{aligned} e_{G'}(X', Y' \setminus Z) &\geqslant e_{G'}(X', Y') - |X'||Z| \geqslant \frac{1}{2l}|X'||Y'| - \frac{1}{20l}|X'||Y'| = \frac{9}{20l}|X'||Y'| \\ &\geqslant \frac{9}{20l}|X'||Y' \setminus Z|. \end{aligned}$$

Since each vertex $y \in Y' \setminus Z$ has the property that $E_{G'}(X', y)$ sees each color at most $2l-1$ times, for each $y \in Y' \setminus Z$ we may remove all edges from $E_{G'}(X', y)$ with duplicated colors (keep one for each color). We end up getting a bipartite graph G'' on $X' \sqcup (Y' \setminus Z)$ with at least $9|X'||Y' \setminus Z|/40l^2$ edges. By the Kővári–Sós–Turán theorem [38], there is a $c = c(l) > 0$ such that G'' contains a copy K of $K_{p,p}$ where $p > c \log s$. More precisely, writing $|X'| = m = s^{1/16l^2}$ and $|Y' \setminus Z| = n \geqslant (1 - 1/20l)t^{1/4l}$, the graph G'' contains a copy of $K_{p,p}$ if $|G''| \geqslant 2nm^{1-1/p} > (p-1)^{1/p}nm^{1-1/p} + (p-1)m$, which is an upper bound for the bipartite Turán number for $K_{p,p}$. Since we have proved $|G''| > 9mn/40l^2$, a short calculation shows that we can let $p = (4l)^{-3}\log s$ and therefore $c(l) = 1/(4l)^3$.

Let $K = K_{p,p} \subset G''$ and $V(K) = A \sqcup B$. For each $x \in A$, the edge set $E(x, B)$ is rainbow colored, and for each $y \in B$, the edge set $E(A, y)$ is rainbow colored. Therefore each color class in K is a matching. By Lemma 16 and $s > s_0 = 2^{(4l)^4/c} = 2^{(4l)^7}$, we can find a rainbow colored $K_{4l,4l}$ in K as desired. □

5.2 The induction argument for Lemma 11

We are now ready to prove Lemma 11. Let us recall the statement.

Lemma 11 Let $l \geqslant 2, s, t \geqslant 1$, $D = D_l = (4l)2^{(4l)^7}$, $Z \subset [n]$ with $|Z| < 2l(s+t)$, $G = K_{s,t}$ be a complete bipartite graph with vertex set $V(G) \subset [n]$. Then the number of ways to edge-color G with colors from Z such that the extension G^* contains no C_{2l}, is at most $D^{(s+t)^2}$.

Here is a sketch of the proof. We proceed by induction on $s+t$. For the base cases $s+t = O(1)$ and $t < s/\log s$, we just upper bound the number of colorings using the trivial bound σ^{st}, where $\sigma = |Z|$ is less than $2l(s+t)$ by hypothesis. For the induction step, we apply Theorem 17 to show that any coloring of $G = K_{s,t}$ contains a rainbow $K_{4l,4l}$ or a $K_{q,2l}$ that is either Q-canonical or monochromatic, where $q = |Q| = s^{\epsilon}$. The case of a rainbow $K_{4l,4l}$ is very easy to handle so we focus on the other two cases. So we are counting colorings of G that can be constructed as follows: first pick a q-set Q from X and a $2l$-set R from Y; next color $E(Q, R)$ in a Q-canonical or monochromatic fashion, then color $E(Q, Y \setminus R)$ to obtain a coloring of $E(Q, Y)$; finally color $E(X \setminus Q, Y)$.

In both cases (monochromatic and canonical), the number of ways to pick the q-set is at most s^q. The main step in the proof is to show that the number of ways to color $E(Q, Y)$ is bounded by $2^{O(qs)} = 2^{s^{1+\epsilon}}$, instead of the trivial $\sigma^{qs} = 2^{O(s^{1+\epsilon}\log s)}$. We will show this claim in the last two subsections. The idea is to first color a strongly rainbow path starting in Q since this creates restrictions on the possible colorings of the remaining edges.

Finally, the number of colorings of $E(X \setminus Q, Y)$ is at most $D^{(s+t-q)^2}$ by the induction hypothesis. Altogether, the number of colorings of $E(X,Y)$ is at most

$$s^q \cdot 2^{O(sq)} \cdot D^{(s+t-q)^2} \leqslant D^{(s+t)^2}.$$

Proof of Lemma 11. Let the vertex set of G be $S \sqcup T$ with $|S| = s$ and $|T| = t$. We apply induction on $s+t$. Denote $|Z| := \sigma < 2l(s+t)$. The number of ways to color G is at most σ^{st}. As long as $s+t \leqslant D/2l$, we have

$$\sigma^{st} \leqslant D^{st} \leqslant D^{(s+t)^2}$$

and this concludes the base case(s).

For the induction step, we may henceforth assume $s+t > D/2l$, and the statement holds for all smaller values of $s+t$. Let us also assume without loss of generality that $t \leqslant s$.

Next, we deal with the case $t \leqslant s/\log s$. Since $D = (4l)2^{(4l)^7} > 16l^2$, $s > (s+t)/2 > D/4l > 4l$ and the number of ways to color G is at most

$$\sigma^{st} \leqslant (2l(s+t))^{st} \leqslant 2^{\frac{s^2\log(2l(s+t))}{\log s}} \leqslant 2^{\frac{s^2\log(4ls)}{\log s}} \leqslant 2^{2s^2} \leqslant 2^{(s+t)^2\log D} = D^{(s+t)^2}.$$

Therefore, for the rest of the proof we may assume $s/\log s < t \leqslant s$ and $s > D/4l = 2^{(4l)^7}$. Since s and t differ from each other by only a little, we would like to give the two partite sets (recall that $V(G) = S \sqcup T$) of G pseudonyms, so that we can discuss the appearance of some coloring pattern of some subgraph and its symmetric case at the same time.

Definition 18 *Let $\{X,Y\} = \{S,T\}$. Given $X' \subset X$ and $Y' \subset Y$, let $\#E(X',Y')$ be the number of ways to color the edges in $E(X',Y')$.*

The following lemma provides the essential idea of the induction step.

Claim 19 *Let $q = s^\epsilon < s/\log s < t < s$ and $s > D/4l = 2^{(4l)^7}$. Let $G = K_{s,t}$ on the vertex set $S \sqcup T$, let $\{X,Y\} = \{S,T\}$. Suppose that there is a subset $Q \subset X$ with $|Q| = q$ such that $\#E(Q,Y) \leqslant 2^{70l^2qs}$. Then $\#E(X,Y) < D^{(s+t)^2}/(4s^q)$.*

Proof. Delete Q from X and apply the induction hypothesis (of Lemma 11) to obtain $\#E(X \setminus Q, Y) \leqslant D^{(s+t-q)^2}$. Together with the condition $\#E(Q,Y) \leqslant 2^{70l^2qs}$ we have

$$s^q \cdot \#E(X,Y) \leqslant s^q \cdot \#E(Q,Y) \cdot \#E(X \setminus Q, Y).$$

Take logs, we get

$$\begin{aligned} q\log s + \log \#E(X,Y) &\leqslant q\log s + 70l^2 qs + \log\left(D^{(s+t-q)^2}\right) \\ &= q\log s + 70l^2 qs - 2qs\log D + q(q-2t)\log D + (s+t)^2\log D \\ &\leqslant 71l^2 qs - 2qs\log D + (s+t)^2\log D \\ &< \log\frac{1}{4} + (s+t)^2\log D, \end{aligned}$$

where the second inequality holds since $\log s < s$ and $q - 2t < 0$, while the last inequality holds since we take $D = (4l)2^{(4l)^7}$, so $(71l^2 - 2\log D)qs < -2$. Therefore, we have $\#E(X,Y) < D^{(s+t)^2}/(4s^q)$. □

Since we may assume that $s/\log s < t \leqslant s$, and $s > D/4l = 2^{(4l)^7} = s_0$, the conditions of Theorem 17 hold. Let $N_G = (z_e)_{e\in G}$ be an edge-coloring of G using colors in Z. By Theorem 17, such an edge-colored G will contain a subgraph G' that is either

- a rainbow colored $K_{4l,4l}$, or
- a Q-canonical $K_{q,2l}$, or
- a monochromatic $K_{q,2l}$,

where $|Q| = q = s^\epsilon$ and $\epsilon = 1/18l$.

Claim 20 *G' cannot be a rainbow colored $K_{4l,4l}$.*

Proof of Claim 20. Suppose for a contradiction that $G' = K_{4l,4l}$ is rainbow colored and Z' is the set of colors used on G'. Then $|Z' \setminus V(G')| \geqslant 16l^2 - 8l$. Pick an edge of each color in $Z' \setminus V(G')$ to obtain a strongly rainbow colored subgraph G'' of G' with $|G''| = 16l^2 - 8l \geqslant (2l-1)8l$. By Lemma 12, G'' contains a strongly rainbow colored cycle of length $2 \mod 2l-2$. Lemma 13 now implies the existence of a strongly rainbow colored C_{2l} in G'', which forms a linear C_{2l} in G^*, a contradiction. □

Thus, we are guaranteed that for each edge-coloring of G that we want to count there is a subgraph $G' = K_{q,2l}$ of G that is colored in either a Q-canonical or monochromatic fashion. Let the vertex set of G' be $Q \sqcup R$, where $Q \in \binom{X}{q}$ and $R \in \binom{Y}{2l}$. Define $|X| = a, |Y| = b$, so $\{a,b\} = \{s,t\}$.

There are four combinations according to the choice of (X,Y) and the coloring patterns. If we can show that in each case $\#E(Q,Y) \leqslant 2^{70l^2 qs}$, then we are done. Because in each case, to count the number of colorings, we may first choose a q-set from X then apply Claim 19 to color G. The total number of colorings is at most

$$s^q \cdot \#E(X,Y) \leqslant s^q \cdot \frac{D^{(s+t)^2}}{4s^q} = \frac{D^{(s+t)^2}}{4}.$$

Therefore, our goal of the last two subsections is to show the following. When $Q \subset X$ is fixed, if there exists an $R \subset Y$ such that $E(Q,R)$ is either Q-canonical or monochromatic, we have $\#E(Q,Y) \leqslant 2^{70l^2 qs}$.

5.2.1 The canonical case

Recall that for each $x \in Q$, the edges in $E(x, R)$ all have the same color z_x which is called a *canonical color*. Let $Z_c = \{z_x : x \in Q\}$ be the set of all canonical colors. For each edge xy with $x \in Q, y \in Y \setminus (R \cup Z_c)$, a color $z_{xy} \neq z_x$ is called a *free color*. We will count the number of colorings of $E(Q, Y)$, and then remove Q to apply the induction hypothesis. For each coloring N_G, consider the following partition of $Y \setminus (R \cup Z_c)$ into two parts:

$$
\begin{aligned}
Y_0 &= \{y \in Y \setminus (R \cup Z_c) : E(y, Q) \text{ sees at most } 11l - 1 \text{ distinct free colors}\},\\
Y_1 &= \{y \in Y \setminus (R \cup Z_c) : E(y, Q) \text{ sees at least } 11l \text{ distinct free colors}\}.
\end{aligned}
$$

We claim that the length of strongly rainbow colored paths that lie between Q and Y_1 is bounded.

Claim 21 *If there exists a strongly rainbow colored path $P = P_{2l-2} \subset E(Q, Y_1)$ with both end-vertices $u, v \in Q$, then there exists a C_{2l} in G^*.*

Proof of Claim 21. Clearly, P extends to a linear P_{2l-2} in G^*. We may assume both $z_u, z_v \notin V(P^*)$, where $P^* = \{e \cup \{z_e\} : e \in P\}$ is the extension of P. Otherwise, suppose w.l.o.g. $z_u \in V(P^*)$, let y be the vertex next to u in P, let S_y be of maximum size among sets

$$\{x \in Q : xy \text{ all colored by distinct free colors}\}.$$

Since $y \in Y_1$, $|S_y| \geqslant 11l$. Note that $|V(P^*)| = 4l - 3$ and $|V(P^*) \cap Y_1| \geqslant l - 1$, we have $|S_y \setminus V(P^*)| \geqslant 11l - (4l - 3 - (l - 1)) \geqslant 8l$. Since $|V(P^*)| < 4l$, $E(y, S_y)$ is rainbow, and G' is Q-canonical, there must be at least $4l$ vertices in $S_y \setminus V(P^*)$ whose canonical color is not in $V(P^*)$. Among these $4l$ vertices there is at least one u' with $z_{u'y} \notin V(P^*)$. Replacing u by u', we get a strongly rainbow colored path of length $2l - 2$ with $z_u \notin V(P^*)$.

Now, Since $|R| = 2l$, we can find a vertex $y \in R$ such that $y \notin \{z_e : e \in P\}$. Further, since both $z_u, z_v \notin V(P^*)$ and $z_u \neq z_v$, the set of edges

$$P^* \cup \{uyz_u, vyz_v\}$$

forms a copy of C_{2l} in G^*. $\square$

Thanks to this observation about strongly rainbow paths, we can bound the number of colorings on $E(Q, Y_1)$ as follows.

Claim 22 $\#E(Q, Y_1) \leqslant (2l)^q \cdot (32l^2)^{bq} \cdot (qb)^{2lq} \cdot \sigma^{6lq+8l^2b}$.

Proof of Claim 22. By Claim 21, according to the length of the longest strongly rainbow colored path starting at a vertex, Q can be partitioned into $2l - 3$ parts $\bigsqcup_{i=1}^{2l-3} Q_i$, where

$$
\begin{aligned}
Q_i = \{x \in Q : &\text{ the longest strongly rainbow colored path}\\
&\text{starting at } x \text{ and contained in } E(Q, Y_1) \text{ has length } i\}.
\end{aligned}
$$

For each i, let $q_i = |Q_i|$. We now bound the number of colorings of the edges in $E(Q_i, Y_1)$.

Firstly, for each $x \in Q_i$, choose an i-path $P_x \subset E(Q, Y_1)$ starting at x and color it strongly rainbow. The number of ways to choose and color these paths for all the vertices $x \in Q_i$ is

$$\prod_{x \in Q_i} \#P_x \leqslant \left((qb)^{\lceil (i+1)/2 \rceil} \sigma^i\right)^{q_i} \leqslant (qb\sigma)^{iq_i}.$$

Fix an $x \in Q_i$. Partition Y_1 into 3 parts depending on whether y is on the extension P_x^* of the path starting at x, or the color of xy is on P_x^* or else, i.e. $Y_1 = \bigsqcup_{j=1}^3 Y_{i,x}^{(j)}$, where

$$\begin{aligned}
Y_{i,x}^{(1)} &= Y_1 \cap V(P_x^*), \\
Y_{i,x}^{(2)} &= \{y \in Y_1 \setminus Y_{i,x}^{(1)} : z_{xy} \in V(P_x^*)\}, \\
Y_{i,x}^{(3)} &= Y_1 \setminus (Y_{i,x}^{(1)} \cup Y_{i,x}^{(2)}).
\end{aligned}$$

Depending on the part of Y_1 that a vertex y lies in, we can get different restrictions on the coloring of the edges in $E(y, Q_i)$.

- If $y \in Y_{i,x}^{(1)}$, then z_{xy} has as many as σ choices. Note that $|P_x^*| = 2i + 1$, and $|Y_{i,x}^{(1)}| \leqslant i + \lceil i/2 \rceil \leqslant 2i$. This gives $\#E(x, Y_{i,x}^{(1)}) \leqslant \sigma^{2i}$.
- If $y \in Y_{i,x}^{(2)}$, then $z_{xy} \in V(P_x^*)$, so there are at most $2i + 1$ choices for this color and $\#E(x, Y_{i,x}^{(2)}) \leqslant (2i + 1)^b$.
- Lastly, let $|Y_{i,x}^{(3)}| = b_{i,x}$. If $y \in Y_{i,x}^{(3)}$, then xy extends P_x into a strongly rainbow colored path $P'_x = P_x \cup \{xy\}$ of length $i + 1$, which forces the edges $x'y$ to be colored by $V(P'^*_x)$ for each $x' \in Q_i \setminus V(P'^*_x)$. Otherwise, the path $P'_x \cup \{x'y\}$ is a strongly rainbow colored path of length $i + 2$ starting at a vertex $x' \in Q_i$, contradicting the definition of Q_i. Therefore, $z_{x'y}$ has at most $2i+3$ choices if $x' \in Q_i \setminus V(P'^*_x)$. Putting this together, for each $y \in Y_{i,x}^{(3)}$, we have

  $$\#E(Q_i \setminus V(P'^*_x), y) \leqslant (2i + 3)^{q_i}.$$

  Noticing that $|Q_i \cap V(P'^*_x)| \leqslant i + 1 + \lceil (i + 1)/2 \rceil \leqslant 2i + 1$, we have

  $$\#E(Q_i, y) \leqslant \#E(Q_i \cap V(P'^*_x), y) \cdot \#E(Q_i \setminus V(P'^*_x), y) \leqslant \sigma^{2i+1}(2i + 3)^{q_i}.$$

  Counting over all $y \in Y_{i,x}^{(3)}$, we have

  $$\#E(Q_i, Y_{i,x}^{(3)}) \leqslant \prod_{y \in Y_{i,x}^{(3)}} \#E(Q_i, y) \leqslant \sigma^{(2i+1)b_{i,x}}(2i + 3)^{q_i b_{i,x}}.$$

Hence the number of ways to color $E(x, Y_1) \cup E(Q_i, Y^{(3)}_{i,x})$ is at most

$$2^b \cdot \#E(x, Y^{(1)}_{i,x}) \cdot \#E(x, Y^{(2)}_{i,x}) \cdot \#E(Q_i, Y^{(3)}_{i,x}) \leqslant 2^b \cdot \sigma^{2i} \cdot (2i+1)^b \cdot \sigma^{(2i+1)b_{i,x}}(2i+3)^{q_i b_{i,x}}.$$

The term 2^b arises above since $Y^{(1)}_{i,x}$ has already been fixed by choosing and coloring P_x, so we just need to partition $Y_1 \setminus Y^{(1)}_{i,x}$ to get $Y^{(2)}_{i,x}$ and $Y^{(3)}_{i,x}$.

Now we remove x from Q_i, $Y^{(3)}_{i,x}$ from Y_1 and repeat the above steps until we have the entire $E(Q_i, Y_1)$ colored. Note that $\sum_{x \in Q_i} b_{i,x} \leqslant b$, and that $i \leqslant 2l - 3$ which implies $2i + 3 < 4l$. We obtain

$$\begin{aligned}
\#E(Q_i, Y_1) &\leqslant \prod_{x \in Q_i} \#P_x \cdot \# \left(E(x, Y_1) \cup E(Q_i, Y^{(3)}_{i,x}) \right) \\
&\leqslant (qb\sigma)^{iq_i} \prod_{x \in Q_i} 2^b \cdot \sigma^{2i} \cdot (2i+1)^b \cdot \sigma^{(2i+1)b_{i,x}}(2i+3)^{q_i b_{i,x}} \\
&\leqslant (qb\sigma)^{2lq_i} + \prod_{x \in Q_i} 2^b \cdot \sigma^{4l+4lb_{i,x}} \cdot (4l)^{b+q_i b_{i,x}} \\
&\leqslant (qb\sigma)^{2lq_i} \cdot 2^{bq_i} \cdot \sigma^{4lq_i+4lb} \cdot (4l)^{bq_i+bq_i} \\
&= (32l^2)^{bq_i} \cdot (qb)^{2lq_i} \cdot \sigma^{6lq_i+4lb}.
\end{aligned}$$

Because $\sum_{i=1}^{2l-3} q_i = q$, taking the product over $i \in [2l-3]$, we obtain

$$\begin{aligned}
\#E(Q, Y_1) \leqslant (2l-3)^q \prod_{i=1}^{2l-3} \#E(Q_i, Y_1) &\leqslant (2l-3)^q \prod_{i=1}^{2l-3} (32l^2)^{bq_i} \cdot (qb)^{2lq_i} \cdot \sigma^{6lq_i+4lb} \\
&\leqslant (2l)^q \cdot (32l^2)^{bq} \cdot (qb)^{2lq} \cdot \sigma^{6lq+8l^2b},
\end{aligned}$$

where $(2l-3)^q$ counts the number of partitions of Q into the Q_i. □

Since $G' = E(Q, R)$ is Q-canonical,

$$\#E(Q, R) \leqslant \sigma^q.$$

As $|Z_c| \leqslant q$,

$$\#E(Q, Y \cap Z_c) \leqslant \sigma^{q^2}.$$

By definition of Y_0,

$$\#E(Q, Y_0) \leqslant (\sigma^{11l}(11l+1)^q)^b \leqslant (\sigma^{11l}(12l)^q)^b.$$

Therefore to color $E(Q, Y)$, we need to first choose the subsets R and $Z_c \cap Y$ of Y and then take a partition to get Y_0 and Y_1. We color each of $E(Q, R), E(Q, Y \cap Z_c), E(Q, Y_0)$ and $E(Q, Y_1)$. This gives

$$\begin{aligned}
\#E(Q, Y) &\leqslant b^{2l} b^q 2^b \cdot \#E(Q, R) \cdot \#E(Q, Y \cap Z_c) \cdot \#E(Q, Y_0) \cdot \#E(Q, Y_1) \\
&\leqslant b^{2l} b^q 2^b \cdot \sigma^q \cdot \sigma^{q^2} \cdot (\sigma^{11l}(12l)^q)^b \cdot [(2l)^q \cdot (32l^2)^{bq} \cdot (qb)^{2lq} \cdot \sigma^{6lq+8l^2b}] \\
&= b^{2l} 2^b \cdot (2lb)^q \cdot (384l^3)^{qb} \cdot (qb)^{2lq} \cdot \sigma^{q^2+(6l+1)q+(8l^2+11l)b}.
\end{aligned}$$

Recall that $q = s^\epsilon < s/\log s < t \leqslant s$, $\sigma \leqslant 2l(s+t) \leqslant 4ls$, and $s > 2^{(4l)^7} > 4l$. There are two cases according to the choices of a and b, i.e. $(a,b) = (s,t)$ and $(a,b) = (t,s)$. But in either case, we have $b \leqslant s$, hence

$$\begin{aligned}\#E(Q,Y) &\leqslant s^{2l}2^s \cdot (2ls)^q \cdot (384l^3)^{qs} \cdot (qs)^{2lq} \cdot (4ls)^{q^2+(6l+1)q+(8l^2+11l)s} \\ &\leqslant s^{2l}2^s \cdot s^{2q} \cdot (384l^3)^{qs} \cdot s^{4lq} \cdot s^{2q^2+2(6l+1)q+2(8l^2+11l)s}.\end{aligned}$$

Take logs,

$$\begin{aligned}\log \#E(Q,Y) &\leqslant \big(2l + 2q + 4lq + 2q^2 + 2(6l+1)q + 2(8l^2+11l)s\big)\log s + s + qs\log(384l^3) \\ &\leqslant 3q^2\log s + (16l^2 + 22l + 1)s\log s + qs\left(\log 384 + \log(l^3)\right) \\ &\leqslant l^2 s\log s + 27l^2 s\log s + qs\,(9 + 3\log l) \\ &\leqslant 28l^2qs + 3l^2qs < 70l^2qs.\end{aligned}$$

The second last inequality holds since $\log s < s^{1/18l} = s^\epsilon = q$ when $s > 2^{(4l)^7}$.

5.2.2 The monochromatic case

Recall that the vertex set of $G' = K_{q,2l}$ is $Q \sqcup R$, where $Q \in \binom{X}{q}$ and $R \in \binom{Y}{2l}$. The term *canonical color* now refers to the only color z_c that is used to color all edges of G', and $Z_c = \{z_c\}$ still means the set of canonical colors. A *free color* is a color that is not z_c. As before we will count the number of colorings of $E(Q,Y)$, and then remove Q to apply the induction hypothesis.

Let $Y_1 = Y \setminus (R \cup Z_c)$. Similar to Claim 21, we claim that the length of a strongly rainbow colored path between Q and Y_1 is bounded.

Claim 23 *If there exists a strongly rainbow colored path $P = P_{4l-2} \subset E(Q,Y_1)$ with both end-vertices $u,v \in Q$, then there exists a C_{2l} in G^*.*

Proof of Claim 23. We observe that z_c appears in the path or the color of the path at most once, as P is strongly rainbow. Hence, by the pigeonhole principle, there exists a sub-path P' of length $2l-2$ such that $z_c \notin V(P'^*)$ and both end-vertices u,v of P' are in Q.

Now, Since $|R| = 2l$, we can find two vertices $y, y' \in R$ such that $y, y' \notin \{z_e : e \in P'\}$. Thus, the edges

$$P'^* \cup \{uyz_c, vy'z_c\}$$

yield a copy of C_{2l} in G^*. □

Again, we first use this claim to color $E(Q,Y_1)$.

Claim 24 $\#E(Q,Y_1) \leqslant (4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{12lq+32l^2b}$.

Proof of Claim 24. The proof proceeds exactly the same as that of Claim 22, except that Q is partitioned into $4l-3$ parts $\bigsqcup_{i=1}^{4l-3} Q_i$. So in the calculation at the end, we have $i \leqslant 4l-3$ which gives $2i+3 < 8l$ and

$$\begin{aligned}\#E(Q_i, Y_1) &\leqslant \prod_{x\in Q_i} \#P_x \cdot \# \left(E(x, Y_1) \cup E(Q_i, Y_{i,x}^{(3)})\right) \\ &\leqslant (qb\sigma)^{iq_i} \prod_{x\in Q_i} 2^b \cdot \sigma^{2i} \cdot (2i+1)^b \cdot \sigma^{(2i+1)b_{i,x}} (2i+3)^{q_i b_{i,x}} \\ &\leqslant (qb\sigma)^{4lq_i} \prod_{x\in Q_i} 2^b \cdot \sigma^{8l+8lb_{i,x}} \cdot (8l)^{b+q_i b_{i,x}} \\ &\leqslant (qb\sigma)^{4lq_i} \cdot 2^{bq_i} \cdot \sigma^{8lq_i+8lb} \cdot (8l)^{bq_i+bq_i} \\ &\leqslant (128l^2)^{bq_i} \cdot (qb)^{4lq_i} \cdot \sigma^{12lq_i+8lb}.\end{aligned}$$

Again, note that $\sum_{i=1}^{4l-3} q_i = q$. Taking the product over $i \in [4l-3]$, we obtain

$$\begin{aligned}\#E(Q, Y_1) \leqslant (4l-3)^q \prod_{i=1}^{4l-3} \#E(Q_i, Y_1) &\leqslant (4l-3)^q \prod_{i=1}^{4l-3} (128l^2)^{bq_i} \cdot (qb)^{4lq_i} \cdot \sigma^{12lq_i+8lb} \\ &\leqslant (4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{12lq+32l^2b},\end{aligned}$$

where $(4l-3)^q$ counts the number of partitions of Q into the Q_i. $\square$

Similarly, to color $E(Q, Y)$, we need to choose the subsets R and $Y \cap Z_c$, and what remains is Y_1. Consequently,

$$\begin{aligned}\#E(Q, Y) &\leqslant b^{2l} b \cdot \#E(Q, R) \cdot \#E(Q, Y \cap Z_c) \cdot \#E(Q, Y_1) \\ &\leqslant b^{2l} b \cdot \sigma \cdot \sigma^q \cdot [(4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{12lq+32l^2b}] \\ &= b^{2l+1} (4l)^q \cdot (128l^2)^{qb} \cdot (qb)^{4lq} \cdot \sigma^{1+(12l+1)q+32l^2b}.\end{aligned}$$

Recall that $q = s^\epsilon < s/\log s < t \leqslant s$, $\sigma \leqslant 2l(s+t) \leqslant 4ls$ and $s > 2^{(4l)^7} > 4l$. There are two cases according to the choices of a and b, i.e. $(a, b) = (s, t)$ and $(a, b) = (t, s)$. In either case $b \leqslant s$, hence

$$\begin{aligned}\#E(Q, Y) &\leqslant s^{2l+1} (4l)^q \cdot (128l^2)^{qs} \cdot (qs)^{2lq} \cdot (4ls)^{1+(12l+1)q+32l^2s} \\ &\leqslant s^{2l+1} (4l)^q \cdot (128l^2)^{qs} \cdot (s)^{4lq} \cdot s^{2+2(12l+1)q+64l^2s}\end{aligned}$$

Take logs,

$$\begin{aligned}\log \#E(Q, Y) &\leqslant (2l + 1 + 4lq + 2 + 2(12l+1)q + 64l^2 s) \log s + q \log(4l) + qs \log(128l^2) \\ &\leqslant 65l^2 s \log s + qs\,(7 + 2 \log l) \\ &\leqslant 70l^2 qs.\end{aligned}$$

Again, the last inequality holds since $\log s < q$ when $s > 2^{(4l)^7}$. $\square$

Acknowledgment. We are very grateful to Rob Morris for clarifying some technical parts of the proof in [44] at the early stages of this project. After those discussions, we realized that the method of hypergraph containers would not apply easily to prove Theorem 2 and we therefore developed new ideas. We are also grateful to Jozsef Balogh for providing us with some pertinent references, and to Jie Han for pointing out that our proof of Theorem 2 applies for $r > 3$ and odd k. Finally, many thanks to all referees for their helpful suggestions on the presentation of this paper. After their suggestions, the manuscript has undergone an extensive rewriting.

References

[1] N. Alon, J. Balogh, B. Bollobás and R. Morris, The structure of almost all graphs in a hereditary property, *J. Combin. Theory, Ser. B*, 101 (2011), 85–110.

[2] B. Bollobás, Cycles modulo k, *Bull. London Math. Soc.*, 9 (1977), 97–98.

[3] S. Bublitz, Decomposition of graphs and monotone formula size of homogeneous fuctions, *Acta Informatica*, 23 (1986), 689–696.

[4] J. Balogh, T. Bohman and D. Mubayi, Erdős–Ko–Rado in random hypergraphs, *Combin. Probab. Comput.*, 18 (2009), 629–646.

[5] B. Bollobás, B. Narayanan and A. Raigorodskii, On the stability of the Erdős–Ko–Rado theorem, *J. Combin. Theory, Ser. A*, 137 (2016), 64–78.

[6] J. Balogh, B. Bollobás and M. Simonovits, On the number of graphs without forbidden subgraph, *J. Combin. Theory, Ser. B*, 91 (2004), 1–24.

[7] ———, The typical structure of graphs without given excluded subgraphs, *Random Structures Algorithms*, 34 (2009), no. 3, 305–318.

[8] ———, The fine structure of octahedron-free graphs, *J. Combin. Theory Ser. B*, 101 (2011), no. 2, 67–84.

[9] J. Balogh, S. Das, M. Delcourt, H. Liu, M. Sharifzadeh, Intersecting families of discrete structures are typically trivial, *J. Combin. Theory, Ser. A*, 132 (2015), 224–245.

[10] J. Balogh and D. Mubayi, Almost all triple systems with independent neighborhoods are semi-bipartite, *J. Combin. Theory Ser. A*, 118 (2011), no. 4, 1494–1518.

[11] ———, Almost all triangle-free triple systems are tripartite, *Combinatorica*, 32 (2012), no. 2, 143–169.

[12] J. Balogh and A.Z. Wagner, On the number of union-free families, *Israel J. Math.*, Vol. 219, Issue 1, (2017), 431–448.

[13] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.*, **28** (2015), 669–709.

[14] J. Balogh, B. Narayanan, J. Skokan, The number of hypergraphs without linear cycles, arXiv:1706.01207, preprint, 2017.

[15] J. Balogh and W. Samotij, The number of $K_{m,m}$-free graphs, *Combinatorica*, **31** (2011), 131–150.

[16] ______, The number of $K_{s,t}$-free graphs, *J. London Math. Soc.*, **83** (2011), 368–388.

[17] J. A. Bondy and M. Simonovits, Cycles of even length in graphs, *J. Combin. Theory Ser. B*, 16 (1974), 97–105.

[18] G. Brightwell and S. Goodall, The number of partial orders of fixed width, *Order*, 13 (1996), no. 4, 315–337.

[19] G. Brightwell, D.A. Grable and H. J. Prömel, Forbidden induced partial orders, *Discrete Math.*, 201 (1999), no. 1-3, 53–80.

[20] F.R.K. Chung, P. Erdős and J. Spencer, On the decomposition of graphs into complete bipartite subgraphs *Studies in Pure Math.* (1983), 95–101 Mem. of P. Turán.

[21] C. Collier-Cartaino, N. Graber and T. Jiang, Linear Turán numbers of r-uniform linear cycles and related Ramsey numbers, arXiv:1404.5015v2, preprint, 2014.

[22] D. Conlon and W.T. Gowers, Combinatorial theorems in sparse random sets, *Ann. of Math.*, 184 (2016), 367–454.

[23] S. Das, T. Tran, Removal and Stability for Erdős–Ko–Rado *SIAM J. Discrete Math.*, 30 (2), 1102–1114.

[24] P. Devlin, J. Kahn, On "stability" in the Erdős–Ko–Rado theorem, *SIAM J. Discrete Math.*, 30 (2), 1283–1289.

[25] A. Diwan, Cycles of even lengths modulo k, *J. Graph Theory*, 65 no. 3 (2010), 246–252.

[26] P. Erdős, D.J. Kleitman and B.L. Rothschild, Asymptotic enumeration of K_n-free graphs, in *Colloquio Internazionale sulle Teorie Combinatorie* (Rome, 1973), Vol. II, 19–27. *Atti dei Convegni Lincei* **17**, Accad. Naz. Lincei, Rome, 1976.

[27] P. Erdős, P. Frankl and V. Rödl, The asymptotic number of graphs not containing a fixed subgraph and a problem for hypergraphs having no exponent, *Graphs and Combinatorics*, **2** (1986), 113–121.

[28] P. Erdős, On extremal problems of graphs and generalized graphs, *Israel J. Math.*, 2 (1964), 183–190.

[29] P. Erdős and A.H. Stone, On the structure of linear graphs, *Bull. Amer. Math. Soc.*, 52 (12) (1946), 1087–1091.

[30] Z. Füredi, T. Jiang, Hypergraph Turán numbers of linear cycles, *J. Combin. Theory Ser. A*, 123 (1) (2014) 252–270.

[31] Z. Füredi, T. Jiang and R. Seiver, Exact solution of the hypergraph Turán problem for k-uniform linear paths, *Combinatorica* 34 ,(2014), 299–322.

[32] A. Ferber, G.A. McKinley and W. Samotij, Supersaturated sparse graphs and hypergraphs, *International Mathematics Research Notices* accepted, 2017.

[33] C. Hundack, H.J. Prömel and Angelika Steger, Extremal graph problems for graphs with a color-critical vertex, *Combin. Probab. Comput.*, 2 (1993), no. 4, 465–477.

[34] D. Kleitman and K. Winston, On the number of graphs without 4-cycles, *Discrete Math.*, **41** (1982), 167–172.

[35] D. Kleitman and D. Wilson, On the number of graphs which lack small cycles, manuscript, 1996.

[36] Ph.G. Kolaitis, H.J. Prömel and B.L. Rothschild, K_{l+1}-free graphs: asymptotic structure and a 0-1 law, *Transactions of the Amer. Math. Soc.*, Vol. 303, No. 2 (Oct., 1987), 637–671.

[37] ——, Asymptotic enumeration of partial orders on a finite set, *Trans. Amer. Math. Soc.*, 205 (1975), 205–220.

[38] T. Kővári, V. Sós and P. Turán, On a problem of K. Zarankiewicz, *Colloquium Math.*, 3 (1954), 50–57.

[39] A. Kostochka, D. Mubayi and J. Verstraëte, Turán problems and shadows I: paths and cycles, *J. Combin. Theory, series A*, 129 (2015), 57–79.

[40] ——, Turán problems and shadows II: trees, *J. Combin. Theory Ser. B*, 122 (2017), 457–478.

[41] ——, Turán problems and shadows III: expansions of graphs, *SIAM J. on Discrete Math.*, Vol. 29, No. 2 (2015), 868–876.

[42] E. Lamken and B.L. Rothschild, The numbers of odd-cycle-free graphs, Finite and infinite sets, Vol. I, II (Eger, 1981),*Colloq. Math. Soc. János Bolyai*, Vol. 37, North-Holland, Amsterdam, 1984, 547–553.

[43] C.-H. Liu, J. Ma, Cycle lengths and minimum degree of graphs, *J. Combin. Theory, Ser. B*, 128 (2018), 66–95.

[44] R. Morris and D. Saxton, The number of C_{2l}-free graphs, *Adv. Math.*, 298 (2016), 534–580.

[45] D. Mubayi and G. Turan, Finding bipartite subgraphs efficiently, *Information Processing Letters*, Vol. 110, Issue 5, 1 (2010), 174–177.

[46] B. Nagle and V. Rödl, The asymptotic number of triple systems not containing a fixed one, *Discrete Math.*, 235 (2001), no. 1-3, 271–290, *Combinatorics*, (Prague, 1998).

[47] B. Nagle, V. Rödl and M. Schacht, Extremal hypergraph problems and the regularity method, Topics in Discrete Mathematics, *Algorithms Combin.*, **26** (2006), 247–278.

[48] D. Osthus, D. Kühn, T. Townsend, and Y. Zhao, On the structure of oriented graphs and digraphs with forbidden tournaments or cycles, *J. of Combin. Theory, Ser. B*, 124 (2017), 88–127.

[49] Y. Person and M. Schacht, Almost all hypergraphs without Fano planes are bipartite, *Proceedings of the Twentieth Annual ACM-SIAM Symposium on Discrete Algorithms* (2009), 217–226.

[50] H.J. Prömel and A. Steger, The asymptotic number of graphs not containing a fixed color-critical subgraph, *Combinatorica*, 12 (1992), no. 4, 463–473.

[51] R.W. Robinson, Counting labeled acyclic digraphs, *New directions in the theory of graphs Proc. Third Ann Arbor Conf., Univ. Michigan, Ann Arbor, Mich., 1971*, (1973), 239–273.

[52] D. Saxton and A. Thomason, Hypergraph containers, *Inventiones Math.*, **201** (2015), 1–68.

[53] R.P. Stanley, Acyclic orientations of graphs, *Discrete Math.*, 5 (1973), 171–178.

[54] B. Sudakov and J. Verstraëte, The extremal function for cycles of length l mod k, *Elec. J. of Combin.*, **22** (2015), #P00.

[55] C. Thomassen, Graph decomposition with applications to subdivisions and path systems modulo k, *J. Graph Theory*, 7 (1983), 261–271.

[56] P. Turán, On an extremal problem in graph theory, *Matematikai és Fizikai Lapok* (in Hungarian), 48 (1941), 436–452.

[57] Z. Tuza, Covering of graphs by complete bipartite subgraphs; complexity of 0-1 matrices, *Combinatorica*, **4** (1) (1984), 111–116.