SPANNING SUBGRAPHS IN GRAPHS AND HYPERGRAPHS

by

IMDADULLAH KHAN

A Dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Computer Science

written under the direction of

Endre Szemerédi

and approved by

____

____

____

____

New Brunswick, New Jersey

May, 2011

ABSTRACT OF THE DISSERTATION

Spanning Subgraphs in Graphs and Hypergraphs

by

Imdadullah Khan

Dissertation Director: Endre Szemerédi

This thesis consists of three new fundamental results on the existence of spanning subgraphs in graphs and hypergraphs.

Cycle Factors in Graphs

A classical conjecture of El-Zahar states that if H is a graph consisting of r vertex disjoint cycles of length $n_1, n_2, \ldots, n_r$, and G is a graph on $n = n_1 + n_2 + \ldots + n_r$ vertices with minimum degree at least $\sum_{i=1}^{r} \lceil n_i/2 \rceil$, then G contains H as a subgraph. A proof of this conjecture for graphs with $n \geq n_0$ was announced by S. Abbasi (1998) using the Regularity Lemma-Blow-up Lemma method. We give a new, "de-regularized" proof of the conjecture for large graphs that avoids the use of the Regularity Lemma, and thus the resulting n_0 is much smaller.

Perfect Matching in three-uniform hypergraphs

A perfect matching in a three-uniform hypergraph on $n = 3k$ vertices is a subset of $\frac{n}{3}$ disjoint edges. We prove that if H is a three-uniform hypergraph on $n = 3k$ vertices such that every vertex belongs to at least $\binom{n-1}{2} - \binom{2n/3}{2} + 1$ edges, then H contains a

perfect matching. We give a construction to show that our result is best possible.

Perfect Matching in four-uniform hypergraphs

A perfect matching in a four-uniform hypergraph is a subset of $\lfloor \frac{n}{4} \rfloor$ disjoint edges. We prove that if H is a sufficiently large four-uniform hypergraph on $n = 4k$ vertices such that every vertex belongs to more than $\binom{n-1}{3} - \binom{3n/4}{3}$ edges, then H contains a perfect matching. Our bound is tight and settles a conjecture of Hán, Person and Schacht (2009).

Acknowledgements

I would like to thank my family for their never-ending love and support all along. I am indebted to my parents and siblings for their selfless love they all have given me. Their understanding and encouragement have allowed me to concentrate on my education. My wife Sumera has been very encouraging and supportive throughout and made my life easier through this Ph.D. Words cannot express how grateful I am for her support, patience and all the caring.

I am thankful to my advisor Endre Szemerédi for teaching me so many things. Endre inspired me to my research interest, supported me at every step during my Ph.D and guided me from the initial to the final stage of this work. Without his help this work would not have been possible. He has generous and very crucial contributions to this thesis.

I would like to thank Michael Grigoriadis, who took keen interest in my work and shared his enthusiasm with me. Many stimulating discussions with him have had a huge impact on the quality of this work. Michael has always helped me in difficult time.

William Steiger has been an invaluable teacher and has inspired me to work on many interesting problems. I wish to extend my gratitude to Bill for his continual interest in my research progress and well being.

I am also thankful to Bruce Reed for serving on my committee. His comments and feedbacks have improved my papers significantly.

I am heartily thankful to Gabor Sarkozy for many valuable discussions on the proof of El-Zahar conjecture. He helped me understand many hard concepts and guided me through the writeup of one paper.

I thank all my friends and fellow graduate students for all the contributions, big and small, direct and indirect.

Lastly I offer my regards to all those who supported me in any aspect during this work. I would especially thank Carol, Joanne, Komal, Maryann, Anetha and Regina for their kindness.

Dedication

Dedicated to my parents and to Sumera and Suhaima.

Table of Contents

List of Figures

Chapter 1

Introduction

This thesis focuses on structural problems in graph theory and combines algorithmic, combinatorial and probabilistic methods. This work is part of a vivid area of discrete mathematics, which, with its methods and results, has contributed to theoretical computer science and computer science in general. In this thesis we not only solve long standing open problems but also develop powerful techniques that have further potential applications. The unifying theme of the problems we tackle, is the following question: *Does a certain density condition in a graph guarantee that a certain subgraph must exist?*

We provide generalizations of two classical problems in graph theory: The celebrated theorem of Dirac on the existence of Hamiltonian cycle in graphs, and the well-studied and widely applicable perfect matching problem extended to hypergraphs.

1.1 Cycle Factors in Graphs

We only consider simple and undirected graphs $G = (V, E)$. We denote the degree of a vertex $v \in V$ by $deg(v)$, the minimum degree of a vertex in V, by $\delta(G)$ and a cycle of length i is denoted by C_i. A cycle containing all vertices of the G is called a Hamiltonian cycle. In this case G is called a Hamiltonian graph. A collection of vertex disjoint cycles in G that covers all vertices of G is called a cycle factor in G.

A fundamental theorem in graph theory is the Dirac Theorem [23], which gives a sufficient condition for the Hamiltonicity of a graph as follows:

Theorem 1 (Dirac 1952). *If G is a simple graph on $n \geq 3$ vertices and $\delta(G) \geq n/2$, then G is Hamiltonian.* □

A long-standing and fundamental conjecture of M.H. El-Zahar generalizes this result on Hamiltonicity of graphs.

Conjecture (El-Zahar (1984))**.** *Let H be a graph consisting of r vertex disjoint cycles of length $n_1, n_2, \ldots, n_r$ and G be a graph on $n = n_1 + n_2 + \ldots + n_r$ vertices, with minimum degree at least $\sum_{i=1}^{r} \lceil n_i/2 \rceil$. Then all cycles in H can be packed into G.*

Simple constructions show that the above minimum degree condition is tight. This beautiful conjecture has generated a lot of attention. Many well known results are only among its special cases. El-Zahar [9] verified his conjecture only for $r = 2$. The famous Corrádi-Hajnal theorem [8] is the special case when every $n_i = 3$. Wang [20] resolved the special case for arbitrary n_1 but $n_i = 3$ for all $i \geq 2$. The case $n_i = 4$ is an old conjecture of Erdös and Faudree [10]. More recently, special cases when H is a mixture of only triangles and quadrilaterals (i.e. $n_i = 3$ or 4), were proved by Yan [22] and Wang [21].

On the other hand, several stronger conditions have been shown to suffice for the existence of such a cycle factor. Sauer and Spencer [4] showed that the much stronger condition $\delta(G) \geq 3n/4$ implies the conclusion of El-Zahar's conjecture, which was later strengthened to, $\delta(G) \geq 2n/3$, by Aigner and Brandt [2] and Alon and Fischer [3]. In general, these results are still far from the original conjecture. In his unpublished work, Abbasi [1] announced a proof of the conjecture using the Szemerédi Regularity Lemma. However, as it is common to all results based on the regularity lemma, the result holds only for astronomically large graphs (involving a tower function). Another interesting direction was taken by Johansson [13, 14]; it considers the path and bipartite variants of the El-Zahar's conjectures.

In this thesis, we prove the El-Zahar's conjecture for graphs of order $n \geq n_0$. Our proof does not invoke the regularity lemma, so that the resulting n_0 is much smaller than the n_0 that would generally be obtained by using the regularity lemma. Rather, our proof relies on elementary algorithmic and graph theoretic methods. Our main result of this section can be summarized, in the following more convenient form.

Theorem. *There exists an n_0 such that the following holds. Let H be a graph consisting* *of r vertex disjoint cycles of length $n_1, n_2, \ldots, n_r$, where the number of odd cycles is denoted by k. If G is a graph on $n = n_1 + n_2 + \ldots + n_r \geq n_0$ vertices such that*

$$\delta(G) \geq \frac{n+k}{2},$$

then G contains H as a subgraph.

We prove this theorem, considering two cases based on the structure of the graph G, *the extremal* and *the non-extremal case*. For the former, we use a simple matching algorithm to embed the r cycles of H into G. For the latter, our main tool is the *optimal cover*, consisting of balanced complete tripartite graphs, balanced complete bipartite graphs (both of size $c \log n$) and the remaining almost independent set. We embed all cycles of H into the optimal cover by first eliminating the vertices in the bipartite graphs and the independent set. We then embed the remaining cycles into the complete tripartite graphs using a straightforward greedy algorithm.

1.1.1 Perfect matchings in uniform hypergraphs

Many practical problems are formulated as graph matching problems. For example, a graph for which vertices correspond to employees in an organization and edges represent the willingness of two employees to work together. A perfect matching in this graph assigns employees into teams of size two, who are willing to work together. The König-Hall and Tutte theorems give full characterization of bipartite and general graphs, respectively, where a perfect matching can be achieved. The situation drastically changes when the teams have to be of size three or more. The appropriate formulation is now a hypergraph matching problem, where hyperedges represent the willingness of three or more employees to work together. In this thesis we consider, k-uniform hypergraphs, where all hyperedges are of the same size k. In this setting, a graph is a 2-uniform hypergraph, but for $k \geq 3$, the corresponding decision problem becomes **NP-Complete** (from the special case of SET PACKING). Thus we must look for sufficient conditions in k-uniform hypergraphs that guarantee a perfect matching. The unifying theme of a huge body of recent work is to search for a minimum degree condition which would guarantee the existence of a perfect matching.

For graphs (2-uniform hypergraphs) Dirac's theorem (1952) implicitly gives a tight sufficient condition for a perfect matching. For hypergraphs, various notions of *degree* (and *minimum degree*) have been proposed. Given a k-uniform hypergraph $H = (V, E)$ on n vertices, the degree of a d-tuple of vertices $v_1, \ldots, v_d \in \binom{V}{d}$, $1 \leq d \leq k-1$ (denoted by $deg_d(v_1, \ldots, v_d)$) is the number of edges in E containing all of $v_1, \ldots, v_d$. The minimum d-degree, $\delta_d(H)$, is the smallest degree over all d-tuples in $\binom{V}{d}$. The general theme of results is that: *If $\delta_d(H) \geq m_d(k, n)$, then H has a perfect matching.* The function $m_d(k, n)$ is called the threshold for matching. Recent results, mostly due to Han, Kühn, Osthus, Rödl, Ruciński, Schacht, and Szemerédi, primarily address the case $d > 1$ (see the recent survey by Rödl and Ruciński [35]). The case $d = 1$ is quite harder and is studied in Hán, Person and Schacht [26] where the authors posed the following conjecture for $k = 3$ and $d = 1$.

Conjecture (Hán et.al. [26] (2009))**.** *If H is a three-uniform hypergraph on n vertices such that*

$$\delta_1(H) \geq \binom{n-1}{2} - \binom{2n/3}{2},$$

then H has a perfect matching.

A well known construction of a three-uniform hypergraph shows that this conjecture is tight. Hán et.al. [26] proved only an approximate version of their conjecture. In particular, they showed that: *If $\delta_1(H) \geq \left(\frac{5}{9} + \eta\right)\binom{n}{2}$ then H has a perfect matching.*

In Chapter 3 we completely resolve this conjecture. Most recently, in independent work, Kühn, Osthus and Treglown [28] announced the same result. However, the novel idea in this thesis is that, for the first time an old result of Erdös [25] is used to find complete tripartite hypergraphs rather than just hyperedges. The techniques are general and have many other applications. In fact, we used the main ingredient of this section to prove the harder conjecture for the case $k = 4$ and $d = 1$ and to show existence of a Hamiltonian cycle in three-uniform hypergraphs with large vertex degree.

For $k = 4$ and $d = 1$, the Hán et.al conjecture works out to:

Conjecture. *If H is a four-uniform hypergraph on n vertices such that*

$$\delta_1(H) \geq \binom{n-1}{3} - \binom{3n/4}{3},$$

then H has a perfect matching.

The only work toward this conjecture is the recent approximate result of Markström and Ruciński [37], proving: *If $\delta_1(H) \geq \left(\frac{42}{64} + o(1)\right)\binom{n-1}{3}$, then H has a perfect matching.* This result is still quite far from the above conjectured threshold value of minimum degree.

In Chapter 4 we settle this conjecture exactly. While the ideas in this section emanate from Chapter 3, the construction of an almost perfect matching (which together with the 'absorbing step' builds a perfect matching) is substantially more difficult. This is so because, unlike the three-uniform case where the pairs in the neighborhood of a vertex span a graph, now the graph induced by the neighbors of a vertex is itself a three-uniform hypergraph. Here, we demonstrate that the new method of using complete k-partite hypergraphs is extremely powerful. In addition, our techniques for building an almost perfect matching provide a systematic way of attacking the problem for general k-uniform hypergraphs. We intend to pursue this in the near future.

We deal with the problem again in two cases, the extremal and non-extremal case. For the extremal case, simple greedy algorithms, together with probabilistic methods, are employed to find a perfect matching. Our procedure for building a perfect matching in the non-extremal case, can be summarized as follows:

- (*Absorbing small sets:*) Find a relatively small "absorbing matching", $\mathcal{M}_0$, that has the property, that any 'not so large' set of vertices can be absorbed into $\mathcal{M}_0$, meaning that there is a matching containing such a set and the vertices of $\mathcal{M}_0$.

- (*Finding an almost perfect matching:*) Find an almost perfect matching in H, that leaves out at most εn vertices. This matching is disjoint from $\mathcal{M}_0$.

- (*Extending to perfect matching*) The remaining at most εn vertices are "absorbed" into $\mathcal{M}_0$ to get a perfect matching.

Chapter 2

Cycle Factors in Graphs

2.1 Introduction

The vertex-set and the edge-set of the graph G is denoted by $V(G)$ and $E(G)$. K_n is the complete graph on n vertices, $K_{r+1}(t)$ is the complete $(r+1)$-partite graph where each class contains t vertices and $K_2(t) = K(t,t)$ is the complete bipartite graph between two vertex classes of size t. C_l (P_l) denotes the cycle (path) on l vertices. We denote by (A, B, E) a bipartite graph $G = (V, E)$, where $V = A + B$, and $E \subset A \times B$. For a graph G and a subset U of its vertices, $G|_U$ is the restriction of G to U. The set of neighbors of $v \in V$ is $N(v)$. Hence the size of $N(v)$ is $|N(v)| = deg(v) = deg_G(v)$, the degree of v. The minimum degree is denoted by $\delta(G)$ and the maximum degree by $\Delta(G)$ in a graph G. When A, B are subsets of $V(G)$, we denote by $e(A, B)$ the number of edges of G with one endpoint in A and the other in B. In particular, we write $deg(v, U) = e(\{v\}, U)$ for the number of edges from v to U. A graph G_n on n vertices is γ-*dense* if it has at least $\gamma\binom{n}{2}$ edges. A bipartite graph $G(k, l)$ is γ-*dense* if it contains at least γkl edges. If a graph is not γ-dense, then it is γ-*sparse*. A graph G is γ-*connected* if for every partition $V(G) = A \cup B$ with $A, B \neq \emptyset$ the bipartite graph between A and B is γ-dense. Throughout the thesis log denotes the base 2 logarithm.

A classical conjecture of El-Zahar states the following.

Conjecture 1 (El-Zahar conjecture)**.** *Let H be a graph consisting of r vertex disjoint cycles of length $n_1, n_2, \ldots, n_r$ satisfying $n_1 + n_2 + \ldots + n_r = n$, and G be a graph on n vertices with minimum degree at least $\sum_{i=1}^{r} \lceil n_i/2 \rceil$, then G contains H as a subgraph.*

Note that the graph $K_{k-1} + K(\lceil \frac{n-k+1}{2} \rceil, \lceil \frac{n-k+1}{2} \rceil)$ has minimum degree $(n + k - 1)/2$ but contains no k vertex disjoint odd length cycles. Thus, the conjecture is best possible. This beautiful conjecture has generated a lot of attention. El-Zahar proved the conjecture for $r = 2$ in [9]. The case that each $n_i = 3$ (i.e. we only have triangles) follows from a result of Corrádi and Hajnal [8]. Wang [20] verified the conjecture for arbitrary n_1 and $n_i = 3, i \geq 2$. The case that each $n_i = 4$ (i.e. we only have 4-cycles) is an old conjecture of Erdős and Faudree [10] (see [5], [14], [16], [18] and [21] for results related to this special case). For the case of triangles and quadrilaterals see [22]. In general it was proved in [2] and in [3] that $\delta(G) \geq 2n/3$ implies the desired conclusion; note that this is a special case of the Bollobás-Eldridge conjecture (see [6]). In [13] Johansson has shown that an El-Zahar type condition implies path factors.

Finally Abbasi announced a proof of Conjecture 1 for graphs with $n \geq n_0$ in [1]. The proof used the Regularity Lemma-Blow-up Lemma method ([19], [15]) and thus the resulting n_0 was quite large (a tower function).

The main purpose of this chapter is to give a new, "de-regularized" proof of the El-Zahar conjecture for large graphs that avoids the use of the Regularity Lemma and thus we obtain a much smaller n_0. We prove the theorem in the following more convenient form.

Theorem 2. *There exists an n_0 such that the following holds. Let H be a graph consisting of r vertex disjoint cycles of length $n_1, n_2, \ldots, n_r$ satisfying $n_1+n_2+\ldots+n_r = n \geq n_0$, where the number of odd cycles is denoted by k. If G is a graph on n vertices with*

$$\delta(G) \geq \frac{n+k}{2}, \tag{2.1}$$

then G contains H as a subgraph.

2.1.1 Outline of the proof

We will follow a similar outline as with the Regularity method in [1], where the main tool was a so-called *optimal cover* in the reduced graph consisting of triangles, edges and an independent set. However, here a regular pair will be replaced with a complete balanced bipartite graph $K(t,t)$, and a regular triangle will be replaced with a complete balanced tripartite graph $K(t,t,t)$, where $t \geq c \log n$ for some constant c (thus the sizes of the color classes are somewhat smaller, logarithmic instead of linear, but this is still good enough for our purposes).

We will build the optimal cover of this type consisting of complete tripartite and bipartite graphs and an almost independent set in G using the tools developed in section 2.2. We will show that either we can find an optimal cover or we are in one of the following extremal cases.

Extremal Case 1 (EC1) with parameter α: *There exists an $A \subset V(G)$ such that*

- $|A| \geq \frac{n-k}{2} - \alpha n$, *and*
- $d(A) < \alpha$.

Extremal Case 2 (EC2) with parameter α: *There exists an $A \subset V(G)$ such that for $B = V(G) \setminus A$ we have*

- $\frac{n}{2} \geq |A| \geq \frac{n}{2} - \alpha n$, *and*
- $d(A, B) < \alpha$.

These extremal cases will be handled in Section 2.4.1 and 2.4.2. In the non-extremal case in Section 2.3 we will eliminate first the vertices in the independent set and in the complete bipartite graphs by embedding cycles (or cycle parts) into them. In this process we will use parts of the complete tripartite graphs as well, but we maintain the balance of the color classes inside one such complete tripartite graph. Then we finish the embedding of the cycles by applying Lemma 11 inside the complete tripartite graphs.

2.2 Tools

2.2.1 Complete bipartite and tripartite subgraphs

In [1] Abbasi used the Regularity Lemma [19], however, here we use a more elementary approach using only the Kővári-Sós-Turán bound [17]. In what follows we develop some tools that we repeatedly apply to build the cover in the non-extremal case.

We make the following two easy observations.

Fact 3. *If $G(A, B)$ is a bipartite graph with $d(A, B) \geq x$, then there must be at least $\eta|B|$ vertices in B for which the degree in A is at least $(x - \eta)|A|$.*

Indeed, otherwise the total number of edges would be less than

$$\eta|A||B| + (x - \eta)|A||B| = x|A||B|,$$

a contradiction to the fact that $d(A, B) \geq x$.

Fact 4. *If $G(A, B)$ is a 2η-dense bipartite graph, $|A| = c_1 \log n$ and $|B| = c_2 n$ then we can find a complete bipartite subgraph $G'(A', B')$ of G such that $A' \subset A, B' \subset B, |A'| \geq \eta|A|$ and $|B'| \geq \eta c_2 n^{(1-c_1)}$.*

To see this first apply Fact 26 to get a subset of vertices $B_1 \in B$ such that $|B_1| \geq \eta|B|$ and $\forall b \in B_1$ $deg(b, A) \geq \eta|A|$. Now consider the neighborhoods in A, of the vertices in B_1. Since there can be at most $2^{|A|} = n^{c_1}$ such neighborhoods, by averaging there must be a neighborhood that appears for at least $\frac{|B_1|}{n^{c_1}} \geq \frac{\eta c_2 n}{n^{c_1}} = \eta c_2 n^{(1-c_1)}$ vertices of B. This means that we can find the desired complete bipartite graph.

The following lemmas are repeatedly used in section 2.3.

Lemma 5. *For $r \in \{1, 2\}$, let $H(A_1, \ldots, A_r)$ be a complete r-partite graph with $|A_i| = c_1 \log n$, $1 \leq i \leq r$ and B be a set of vertices disjoint from A_i's, with $|B| = c_2 n^{c_3}$, $c_3 > rc_1$. If for $\eta > 0$ we have*

$$d(B, H) \geq \left(\frac{r-1}{r} + r\eta\right)$$

then there is a complete $(r+1)$-partite graph $H'(A'_1, \ldots, A'_r, B')$ such that for $(1 \leq i \leq r)$ we have $A'_i \subset A_i$ and $|A'_i| = r\eta A_i$, and $B' \subset B$ and $|B'| \geq \eta c_2 n^{c_3 - rc_1} \gg |A_i|$.

Proof. First applying Fact 26 we get a subset $B_1 \subset B$ such that $|B_1| \geq \eta|B|$ and $\forall b \in B_1$ $deg(b, H) \geq ((r-1) + r\eta)|A_1|$. In particular every vertex in B_1 has at least $r\eta|A_1|$ neighbors in each color class of H. Now by Fact 27 we get a subset of vertices $B' \subset B_1$ of size at least $\eta c_2 n^{c_3 - rc_1}$ that has the same neighborhood of size at least

$r\eta|X_1|$ in each color class of H, from which we get the required complete $(r+1)$-partite graph. $\square$

Remark 1. *To get a complete $(r+1)$-partite graph in Lemma 5 we only need a set of vertices $B_1 \subset B$ such that $\forall\, b \in B_1$ $deg(b, A_i) \geq \eta|A_i|$ for $1 \leq i \leq r$.*

Remark 2. *Typically we use this lemma to get a balanced complete $(r+1)$-partite graph, which can be achieved by arbitrarily discarding some vertices of B'. When $r = 3$, we will split the complete 4-partite graph guaranteed by Lemma 5 into 4 disjoint balanced complete 3-partite graphs each with a color class of size $\eta|A_1|$.*

The main tool to make the complete tripartite and bipartite graphs in the cover is the following lemma.

Lemma 6 (Theorem 3.1 on page 328 in[6])**.** *There is a constant $0 < c_1 = c_1(\varepsilon, s)$ such that if $0 < \epsilon < 1/s$ and we have a graph G with*

$$e(G) \geq \left(1 - \frac{1}{s} + \epsilon\right) \frac{n^2}{2}$$

then G contains a $K_{s+1}(t)$, where $t = \lfloor c_1 \log n \rfloor$.

For $s = 1$ this is essentially the Kővári-Sós-Turán bound [17] and for general s this was proved by Bollobás, Erdős and Simonovits [7]. Here we will use the result only for $s = 1$ and $s = 2$.

When k is very small (close to 0) in (2.1) then we may not be able to apply lemma 6 for $s = 2$ to get complete tripartite graphs of appropriate sizes in our cover. Therefore we prove the following stability result which shows that in Lemma 6, for $s = 2$, we can slightly lower the necessary density condition when the graph is α-non-extremal. Note that when k is very small, $(k/n < 4\eta)$ then by definition of extremal case 1, G is α-extremal if there exists an $A \subset V(G)$ such that $|A| \geq (1-\alpha)n/2$ and $d(A) < \alpha$.

Lemma 7. *For every $0 < \varepsilon \ll \alpha$ there exist and integer $n_0 = n_0(\varepsilon, \alpha)$ and a constant $0 < c_2 = c_2(\varepsilon, \alpha)$, such that if G is an α-non-extremal graph on $n \geq n_0$ vertices with $\delta(G) \geq (1/2 - \varepsilon)n$, then G contains a $K_3(t)$, where $t = \lfloor c_2 \log n \rfloor$.*

Proof. By Lemma 6 we find a complete bipartite subgraph $K_2(t_1) = (A_1, A_2)$ with $t_1 = \frac{1}{8}\log n$. Let $B = V(G) \setminus (A_1 \cup A_2)$ and let $C \subset B$ be the set of vertices that have at least εt_1 neighbors in both A_1 and A_2. If $|C| \geq \varepsilon^2 n$ then by the remark following Lemma 5 we can find the required complete 3-partite graph, therefore we assume that $|C| < \varepsilon^2 n$ and consider the remaining vertices of B (for simplicity we still denote it by B).

For $i \in \{1, 2\}$, let $B_i = \{b \in B : deg(b, A_i) < \varepsilon|A_i|\}$. By the above observation ($|C|$ is very small) we have $B = B_1 \cup B_2$. From the minimum degree condition and by the definition of B_i we have that

$$\left(\frac{1}{2} - 2\varepsilon\right) n|A_i| \leq e(A_i, B) \leq \varepsilon|A_i||B_i| + |A_i|(|B| - |B_i|)$$

which gives us $|B_i| \leq (1/2 + 3\varepsilon)n$. Since $B = B_1 \cup B_2$ we have that $|B_1 \cap B_2| < 8\varepsilon n$ and $|B_i| \geq (1/2 - 4\varepsilon)n$.

We group the vertices in B_2 by their neighborhoods in A_1 i.e. each group contains vertices that have the same neighborhood in A_1. There can be at most $2^{t_1} < n^{1/8}$ groups. We may safely ignore the exceptional groups that have at most $\sqrt{n}$ vertices in them since they can contain at most $n^{5/8}$ vertices in them. Similarly those groups may be disregarded that have neighborhood in A_1 of size less than $2|A_1|/3$. Indeed, from the minimum degree condition and the size of B_2 we have that $e(A_1, B_2) \geq (1-8\varepsilon)|A_1||B_2|$, and thus the total number of vertices in such groups is at most $20\varepsilon n$.

Now since G is α-non-extremal we have that $d(B_2) \geq \alpha$, so there must exist either one group with internal density more than $\alpha/4$ or a pair of groups with cross density at least $\alpha/4$. In the prior case we are done because by Lemma 6 (with $s = 1$) we can find a $K_2(c_1 \log n)$ and since the two color classes have the same neighborhood in A we can find a $K_3(c_2 \log n)$; in the latter case, since for any two groups the neighborhoods in A have an intersection of size at least $\frac{1}{3}|A_1|$, we can again find a $K_3(c_2 \log n)$. $\square$

Finally in the extremal case we will use the following simple simple facts on the sizes of a maximum set of vertex disjoint paths in G (see [6]).

Lemma 8. *In a graph G on n vertices, we have*

$$\nu_1(G) \geq \min\left\{\delta(G), \left\lfloor\frac{n}{2}\right\rfloor\right\} \text{ and } \nu_2(G) \geq (\delta(G)-1)\frac{n}{6\Delta(G)}$$

where $\nu_i(G)$ denotes the size of maximum set of vertex disjoint paths of length i in G.

2.2.2 Embedding into complete tripartite graphs

One of the key ideas in the proof is to show the statement for nearly balanced complete tripartite graphs, and then what we are left to do is to reduce the general case to these special graphs. Thus in this section we will assume that we have a complete tripartite graph G on n vertices with color classes V_1, V_2 and V_3. Let $|V_i| = m_i$ and assume $m_1 \leq m_2 \leq m_3$. The content of this section can be found in [1], but for the sake of completeness we present the proofs here.

Lemma 9. *If n is even and $m_3 \leq 2m_1$, then G contains a perfect matching.*

Proof. First we take a matching of size $m_2 - m_1$ between V_2 and V_3. Hence we may assume that $m_1 = m_2$. Since G has an even number of vertices, m_3 is even. We pick $m_3/2$ edges between V_1 and V_3 and $m_3/2$ edges between V_2 and V_3. Then there are exactly $m_1 - m_3/2$ vertices left in both V_1 and V_2, thus there is a perfect matching between them. □

Lemma 10. *If $m_3 \leq 2m_1 - k$, then G contains all maximum degree two and minimum degree one graphs with n vertices and k odd components (cycles or paths).*

Proof. Let H be a maximum degree two graph with k odd cycles or paths. Let C be any cycle in H whose size $|C|$ is greater than three (the procedure is similar for a path P of length greater than three). We replace the cycle C by a cycle C' and an edge e where $|C'| = |C| - 2$, resulting in a new maximum degree two graph H'. We claim that if H' is a subgraph of G, then H is a subgraph of G as well. Indeed, suppose there is a good embedding $f : V(H') \to V(G)$. Since G is a complete tripartite graph we can "merge" C' and e, i.e. we can easily find a cycle in G (corresponding to C) that spans the vertices in $f(C') \cup f(e)$.

By repeatedly applying the above argument we may assume that H consists of k triangles or paths of length three and a matching. Now choose k triangles or paths of length three in G by choosing one vertex from each V_i. Note that this is possible as

$$2m_1 \geq m_3 + k \geq \frac{m_1 + m_2 + m_3}{3} + k \geq k + k = 2k.$$

The remaining graph satisfies the conditions of Lemma 9 and thus contains a perfect matching. □

Lemma 11. *If $m_3 \leq (1 + 2/7)m_1$, then G contains all maximum degree two and minimum degree one graphs with n vertices that do not contain a C_3 or a P_3 as a component.*

Proof. Let H be a maximum degree two and minimum degree one graph with n vertices that does not contain a C_3 or a P_3 as a component graph and let us denote the number of odd components (cycles or paths of size at least 5) in H by k. Then we have

$$5k \leq m_1 + m_2 + m_3 \leq \left(3 + \frac{4}{7}\right) m_1 = \frac{25}{7} m_1.$$

Therefore,

$$2m_1 - k \geq \left(2 - \frac{5}{7}\right) m_1 = \left(1 + \frac{2}{7}\right) m_1 \geq m_3.$$

Hence, the conditions of Lemma 10 are satisfied and H is a subgraph of G. □

2.3 The non-extremal case

Throughout this section we assume that we have a graph G satisfying (2.1) such that Extremal Cases 1 and 2 do not hold for G. We shall assume that n is sufficiently large and use the following main parameters

$$0 < \eta \ll \alpha \ll 1, \tag{2.2}$$

where $a \ll b$ means that a is sufficiently small compared to b. In order to present the results transparently we do not compute the actual dependencies, although it could be done. We will use the constant

$$c = \min\{c_1(\eta, 2), c_1(\eta, 1), c_2(\eta, \alpha)\}$$

where c_1 is from Lemma 6 and c_2 is from Lemma 7.

Let $\gamma = k/n$, then we have $\delta(G) \geq (1+\gamma)n/2$. For technical reasons we work with a slightly weaker minimum degree condition, we assume that

$$\delta(G) \geq \left(\frac{1+\gamma}{2} - \eta\right) n. \tag{2.3}$$

In the non-extremal case this slightly smaller value of minimum degree is sufficient.

2.3.1 The optimal cover

Before we start the actual embedding we need some preparations in our host graph G. We are going to work with cover $\mathcal{C} = (\mathcal{T}, \mathcal{M}, \mathcal{I})$, where $\mathcal{T}$ is a collection of disjoint balanced complete tripartite graphs, $\mathcal{M}$ is a collection of disjoint balanced complete bipartite graphs and $\mathcal{I}$ is a set of vertices. Using the following iterative procedure we build an *optimal cover*, where

- we cannot *increase significantly* the number of vertices covered by $\mathcal{T}$, (by at least $\eta^3 n$ vertices),
- we cannot *increase significantly* the number of vertices covered by $\mathcal{M}$, (by at least $\eta^3 n$ vertices), without reducing the number of vertices covered by $\mathcal{T}$.

Then we will show that the optimal cover exhibits nice structural properties.

We begin with the cover $\mathcal{C}_0 = (\mathcal{T}_0, \mathcal{M}_0, \mathcal{I}_0)$. Then in each step $i \geq 1$, if $\mathcal{C}_{i-1}$ is not optimal, we find another cover $\mathcal{C}_i = (\mathcal{T}_i, \mathcal{M}_i, \mathcal{I}_i)$ such that either $V(\mathcal{T}_i) \geq V(\mathcal{T}_{i-1}) + \eta^3 n$ or $V(\mathcal{T}_i) = V(\mathcal{T}_{i-1})$ but $V(\mathcal{M}_i) \geq V(\mathcal{M}_{i-1}) + \eta^3 n$ (for this we use the notation $\mathcal{C}_{i+1} > \mathcal{C}_i$). The size of a color class in each tripartite graph and bipartite graph in $\mathcal{C}_i$ is $t_i = \eta^i c \log n$. Thus in at most $1/\eta^3$ iterations we arrive at an optimal cover $\mathcal{C} = (\mathcal{T}, \mathcal{M}, \mathcal{I})$ while the size of the color classes in the complete tripartite and bipartite graphs is at least $(\eta^{1/\eta^3}) c \log n$, which is large enough when n is sufficiently large.

To get the initial cover $\mathcal{C}_0$, we use either Lemma 6 or Lemma 7 for for $\mathcal{T}_0$, depending on the value of γ. In case $\gamma \geq 4\eta$, the initial cover $\mathcal{C}_0$ is obtained by repeatedly applying

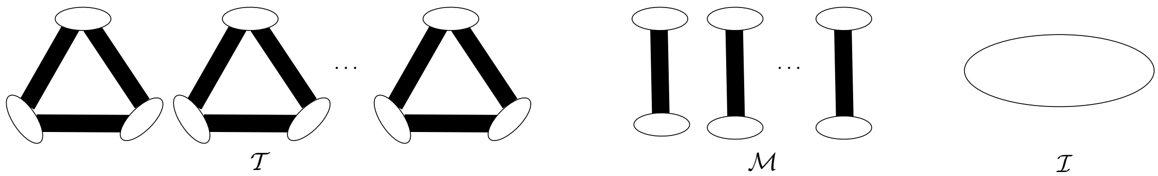


Figure 2.1: The optimal cover: the solid black rectangles indicate complete bipartite graphs

Lemma 6 with $s = 2$ and $\varepsilon = \eta$ in the leftover of G while the conditions of the lemma are satisfied and the number of leftover vertices is at least ηn to find complete tripartite graphs $K_3(t_0)$ where $t_0 = c \log n$. Then when the conditions of the lemma are not satisfied anymore with $s = 2$ (thus the density in the leftover of G is less than $1/2 + \eta$), we apply Lemma 6 repeatedly but this time with $s = 1$ and $\varepsilon = \eta$ while the condition of the lemma is satisfied and the number of leftover vertices is at least ηn to find complete bipartite graphs $K_2(t_0)$.

In case $\gamma < 4\eta$ then first we repeatedly apply Lemma 7 with $\varepsilon = 4\eta$ to get a collection of disjoint complete tripartite graphs $\mathcal{T}_0$ where each $K_3(t_0) \in \mathcal{T}_0$ has $t_0 = c \log n$ vertices in each color class. Note that we can apply Lemma 7 until we have $V(\mathcal{T}_0) \geq \eta n$. At which point we start applying Lemma 6 for $s = 1$, to get $\mathcal{M}_0$ and $\mathcal{I}_0$.

So in either case we have $|V(\mathcal{T}_0)| \geq \eta n$. At the ith step in this iterative procedure if $\mathcal{C}_i$ is not an optimal cover, then we get $\mathcal{C}_{i+1} > \mathcal{C}_i$ as follows: Assuming $|V(\mathcal{M}_i)|, |\mathcal{I}_i| \geq 2\eta n$

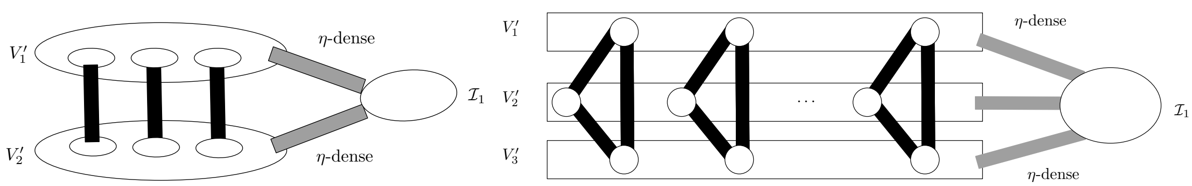
Observation 1: If $d(\mathcal{I}_i) \geq 2\eta$ then we can apply Lemma 6 with $s = 1$ in $\mathcal{I}_i$ to find a collection of complete bipartite graphs in $\mathcal{I}_i$ covering at least $\eta^2 n$ vertices. Adding these new bipartite graphs to $\mathcal{M}_i$, we get $\mathcal{C}_{i+1} > \mathcal{C}_i$.

Observation 2: If there is a subcover $\mathcal{M}' \subset \mathcal{M}_i$ such that if we denote by U_1' and U_2' the union of the two color classes of the bipartite graphs in this subcover, then we have

- $|U_1'| = |U_2'| \geq \eta |V(\mathcal{M}_i)|$ and
- for each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$ we have $d(\mathcal{I}_i, K_b^j) \geq (1/2 + 2\eta)$

then for each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$, by Lemma 5, we can find a disjoint balanced complete tripartite graph, T_j, with color classes of size ηt_i, as subsets of U_1^j, U_2^j and $\mathcal{I}_i$. We remove the vertices of T_j and add it to $\mathcal{T}_{i+1}$. The remaining part of K_b^j is added to $\mathcal{M}_{i+1}$.

We proceed in similar manner for all bipartite graphs in the subcover $\mathcal{M}'$. We add to $\mathcal{M}_{i+1}$ all the bipartite graphs in $\mathcal{M}_i \setminus \mathcal{M}'$. All the tripartite graphs in $\mathcal{T}_i$ are added to $\mathcal{T}_{i+1}$. We split each complete balanced bipartite graph and tripartite graph in $\mathcal{M}_{i+1}$ and $\mathcal{T}_{i+1}$ that has color classes of size more than t_{i+1} (the old ones) into disjoint complete balanced bipartite graphs and tripartite graphs each with a color class of size t_{i+1} (for simplicity we assume that t_i is a multiple of $1/\eta$). Note that, since each T_j uses $\eta t_i = t_{i+1}$ from $\mathcal{I}_i$ this together with the size of $\mathcal{M}'$ and $\mathcal{M}_i$ implies that $|V(\mathcal{T}_{i+1})| = |V(\mathcal{T}_i)| + \eta n^3$. Thus $\mathcal{C}_{i+1} > \mathcal{C}_i$.



(a) $\mathcal{I}_1$ is η-dense to both U_1' and U_2' (b) $\mathcal{I}$ is η-dense to both V_1', V_2' and V_3'

Observation 3: If there is a subcover $\mathcal{T}' \subset \mathcal{T}_i$ such that if we denote by V_1', V_2' and V_3' the union of the three color classes of tripartite graphs in this subcover, then we have

- $|V_1'| = |V_2'| = |V_3'| \geq \eta |V(\mathcal{T}_i)|$ and
- for each $K_t^j = (V_1^j, V_2^j, V_3^j) \in \mathcal{T}'$ we have $d(\mathcal{I}_i, K_t^j) \geq (2/3 + 3\eta)$

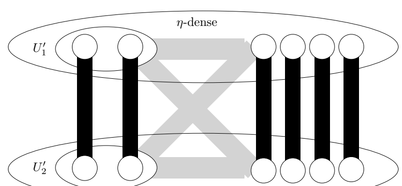
then for each $K_t^j = (V_1^j, V_2^j, V_3^j) \in \mathcal{T}'$, by Lemma 5, we can find a disjoint balanced complete 4-partite graph Q_j, with color classes of size $3\eta t_i$, as subsets of V_1^j, V_2^j, V_3^j and $\mathcal{I}_i$. As noted in the remark following Lemma 5 we split this Q_j into four disjoint balanced complete tripartite graphs, $T_1, \ldots, T_4$ each with color classes of size ηt_i. We remove the vertices of $T_1, \ldots, T_4$, and add them to $\mathcal{T}_{i+1}$. The remaining part of K_t^j is also added to $\mathcal{T}_{i+1}$.

We proceed in similar manner for all tripartite graphs in the subcover $\mathcal{T}'$. We add to $\mathcal{T}_{i+1}$ all the tripartite graphs in $\mathcal{T}_i \setminus \mathcal{T}'$ and let $\mathcal{M}_{i+1} = \mathcal{M}_i$. Then, similarly as above we make all the bipartite and tripartite graphs such that each has a color class of size t_{i+1}. Note that again, since each Q_j uses $3\eta t_i$ vertices from $\mathcal{I}_i$, we get $|V(\mathcal{T}_{i+1})| = |V(\mathcal{T}_i)| + \eta^3 n$ and hence $\mathcal{C}_{i+1} > \mathcal{C}_i$.

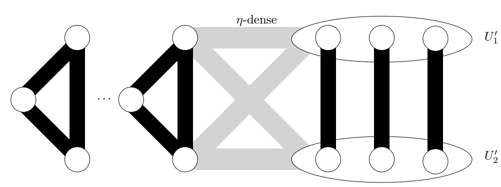
Observation 4: If there is a subcover $\mathcal{M}' \subset \mathcal{M}_i$ such that if we denote by U_1' and U_2' the union of the two color classes of bipartite graphs in this subcover, then we have

- $|U_1'| = |U_2'| \geq \eta |V(\mathcal{M}_i)|$ and
- for each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$ we have $d(V(\mathcal{M}_i), K_b^j) \geq (1/2 + 2\eta)$

then similarly as above for each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$, by Lemma 5, we can find a balanced complete tripartite graph T_j, with color classes of size ηt_i, as subsets of U_1^j, U_2^j and $V(\mathcal{M}_i)$ disjoint from each other (as $|V(\mathcal{M}_i)| \geq 2\eta n$). Adding these new tripartite graphs to $\mathcal{T}_{i+1}$ and repeating the same process as above, we get $|V(\mathcal{T}_{i+1})| = |V(\mathcal{T}_i)| + \eta^3 n$ and thus $\mathcal{C}_{i+1} > \mathcal{C}_i$.



(c) U_1' and U_2' are η-dense to both sides



(d) Two color classes of a tripartite graph are η-dense to both U_1', U_2' and V_3'

Observation 5: If there is a subcover $\mathcal{M}' \subset \mathcal{M}_i$ such that if we denote by U_1' and U_2' the union of the two color classes of bipartite graphs in this subcover, then we have

- $|U_1'| = |U_2'| \geq \eta |V(\mathcal{M}_i)|$ and
- for each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$ we have $d(V(\mathcal{T}_i), K_b^j) \geq (2/3 + 3\eta)$.

In this case for every $K_b^j \in \mathcal{M}'$, there must be a subcover of tripartite graphs $\mathcal{T}^j \subset \mathcal{T}_i$ covering at least $\eta |V(\mathcal{T}_i)|$ vertices, such that for each tripartite graph, $(V_1^l, V_2^l, V_3^l) \in \mathcal{T}^j$, there are at least two color classes (say V_1^l and V_2^l) with both $d(V_1^l, K_b^j) \geq 1/2 + 3\eta$ and $d(V_2^l, K_b^j) \geq 1/2 + 3\eta$. Because otherwise we have

$$e(V(\mathcal{T}_i), K_b^j) < \left(\eta + \frac{1}{3} + \frac{2}{3}\left(\frac{1}{2} + 3\eta\right)\right)|V(\mathcal{T}_i)||V(K_b^j)| = \left(\frac{2}{3} + 3\eta\right)|V(K_b^j)||V(\mathcal{T}_i)|$$

In particular every tripartite graph in $\mathcal{T}^j$ has at least ηt_i vertices in both V_1^l and V_2^l for which the degree in K_b^j is at least $(1/2 + 2\eta)|K_b^j|$, we refer to them as vertices *connected to* K_b^j and the tripartite graphs in $\mathcal{T}^j$ as *connected to* K_b^j.

Assume that for every $K_b^j \in \mathcal{M}'$ the connected vertices in each tripartite graph are in V_1^l and V_2^l. Note that this is true for at least a third fraction of $\mathcal{M}'$ and with the parameter $\eta/3$ (for simplicity we still denote it by $\mathcal{M}'$).

Now for each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$ by Lemma 5, we can find a balanced complete tripartite graph T_1^j with color classes of size at least ηt_i in the set U_1^j, U_2^j and the set of connected vertices to K_b^j. We can also find, another similar balanced complete tripartite graph, T_2^j with color classes in U_1^j, U_2^j and the set of connected vertices. We will repeat this process for every bipartite graph in $\mathcal{M}'$, but we have to pay attention to preserve the balance in the remaining complete tripartite graphs. We will guarantee that each of these new tripartite graphs, use exactly same number of vertices (if at all) in V_1^l and V_2^l from every tripartite graph $(V_1^l, V_2^l, V_3^l) \in \mathcal{T}_i$.

For this purpose we call a tripartite graph *full* if we used up an ηt_i vertices in its two color classes. When finding the next T_1^j and T_2^j we choose the third color class from the set Y_1 constructed as follows: We choose one vertex connected to K_b^j from each V_1^l that is not full yet. As long as an $\eta/2$-fraction of tripartite graphs are not full yet, Y_1 is of size at least $\eta|V(\mathcal{T}_i)|/6t_i \geq \eta^2 n/\log n$. Hence we can apply Lemma 5 on (Y_1, K_b^j), to get a complete tripartite graph. Note that in this complete tripartite graph, the third color class, Y_1' guaranteed by Lemma 5 is of size at least $\sqrt{n}$. Now we make another set, Y_2 by selecting one vertex connected to K_b^j, from each V_2^l if V_1^l has a vertex in Y_1'. Clearly $|Y_2| \geq \sqrt{n}$. Hence it is easy to see that applying Lemma 5 on (Y_2, K_j) we can get a balanced complete tripartite graph T_2^j. From Y_1' we can select an appropriate subset of vertices to make T_1^j such that we use exactly one vertex from a V_1^l and V_2^l.

After applying this procedure for each $K_b^j \in \mathcal{M}'$, from every remaining tripartite graph in $\mathcal{T}_i$ if V_3^l has more vertices than in V_l^1 and V_l^2, we discard arbitrarily some vertices from V_3^l to rebalanced the tripartite graphs. Note that, still the net increase in the number of vertices covered by tripartite graphs, is at least $\eta^3 n$. As above splitting (if necessary) the tripartite graphs and bipartite graphs to get their color classes of size t_{i+1}, we get $\mathcal{C}_{i+1} > \mathcal{C}_i$.

Observation 6: Assume we have the following setup: There is a subcover $\mathcal{M}' \subset \mathcal{M}_i$ and a subcover $\mathcal{T}' \subset \mathcal{T}_i$ such that if we denote by U_1' and U_2', and V_1', V_2' and V_3', the union of the color classes of the graphs in $\mathcal{M}'$ and $\mathcal{T}'$, respectively. Then we have

- $|U_1'| = |U_2'| \geq \eta |V(\mathcal{M}_i)|$
- $|V_1'| = |V_2'| = |V_3'| \geq \eta |V(\mathcal{T}_i)|$
- for each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$ we have $d(\mathcal{I}_i, U_2^j) \geq 2\eta$ and
- for each $K_t^j = (V_1^j, V_2^j, V_3^j) \in \mathcal{T}'$ we have $d(\mathcal{I}_i, (V_2^j, V_3^j)) \geq (1/2 + 2\eta)$

If, in addition, any of the following is true then we can increase the size of our cover.

i. $d(U_1') \geq 2\eta$.

ii. $d(V_1') \geq 2\eta$

iii. $d(\mathcal{I}_i, U_1') \geq 2\eta$

iv. $d(\mathcal{I}_i, V_1') \geq 2\eta$

v. $d(U_1', V_1') \geq 2\eta$

We will show how to increase the size of our cover when $d(U_1') \geq 2\eta$ and when $d(\mathcal{I}_i, U_1') \geq 2\eta$, the other cases are similar. For each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$, applying Lemma 5 we remove a complete bipartite graph with color class of size ηt_i as subset of U_2^j and $\mathcal{I}_i$. Note that this create an imbalance in the bipartite graphs in $\mathcal{M}'$. To re balance, we make a set of vertices Y by randomly selecting a subset of $2\eta t_i$ vertices from U_1^j of each $K_b^j = (U_1^j, U_2^j) \in \mathcal{M}'$. Now if $d(U_1') \geq 2\eta$, then with high probability, we have $d(Y) \geq \eta$. We apply Lemma 6 with $s = 1$ in Y to find a collection of complete bipartite graphs in Y which in total covers at least $\eta|Y|/2 \geq \eta^3 n$ vertices. Discarding the remaining part of Y, and adding these new bipartite graphs to $\mathcal{M}_{i+1}$ we get $|V(\mathcal{M}_{i+1})| = |V(\mathcal{M}_i)| + \eta^3 n$. On the other hand if $d(\mathcal{I}_i, U_1') \geq 2\eta$, then with high probability, we have $d(\mathcal{I}_i, Y) \geq \eta$, hence we can find a collection of complete bipartite graphs in the bipartite graph $(\mathcal{I}_i, Y)$ to increase the size of our cover.

2.3.2 The structure of the optimal cover

Using the above observations, in at most $1/\eta^3$ iterations we arrive at the optimal cover, $\mathcal{C} = (\mathcal{T}, \mathcal{M}, \mathcal{I})$ where a color class of each bipartite and tripartite graph is of size $l = \eta^{1/\eta^3} c \log n$. Let us collect the structural information that we have about this optimal cover. Denote by $\tau = |V(\mathcal{T})|/3n$, $\mu = |V(\mathcal{M})|/2n$, $\beta = |V(\mathcal{I})|/n$ and recall that $\gamma = k/n$, where k is the number of odd cycles to be embedded. As argued above, we have $\tau \geq \eta$.

From the fact that we are not in Extremal Case 1 we will derive our main lemma of the non-extremal case.

Lemma 12. $\tau \geq \min(\gamma + \beta + \alpha, \frac{1}{3} - 2\eta)$.

Proof. We may assume that $\tau < 1/3 - 2\eta$, since otherwise we are done. Then either $\mu \geq \eta$ or $\beta \geq \eta$ (or maybe both). We distinguish the following two cases to prove this lemma based on the size of the independent set $\mathcal{I}$.

Case 1: There is an independent set ($\beta \geq \eta$)

Since $\mathcal{C}$ is an optimal cover, none of the above observations holds in $\mathcal{C}$. By Observations 1,2 and 3, respectively, we have

$$d(\mathcal{I}) < \eta\,, \quad d(\mathcal{I}, V(\mathcal{M})) < \left(\frac{1}{2} + 4\eta\right) \quad \text{and} \quad d(\mathcal{I}, V(\mathcal{T})) < \left(\frac{2}{3} + 6\eta\right) \tag{2.4}$$

Note that this is true even if $\mu < \eta$. We are going to estimate the number of edges going between $\mathcal{I}$ and its complement. Using the minimum degree condition (2.3), (2.4) and the fact that $3\tau + 2\mu + \beta = 1$ we get

$$e(\mathcal{I}, V(\mathcal{T})) \geq |\mathcal{I}| \left(\frac{1+\gamma}{2} - \eta\right) n - 2\eta\binom{|\mathcal{I}|}{2} - |\mathcal{I}| \left(\frac{1}{2} + 4\eta\right) 2\mu n \geq \left(\frac{3\tau + \gamma + \beta}{2} - 20\eta\right) n|\mathcal{I}| \tag{2.5}$$

and

$$e(\mathcal{I}, V(\mathcal{M})) \geq |\mathcal{I}| \left(\frac{1+\gamma}{2} - \eta\right) n - 2\eta\binom{|\mathcal{I}|}{2} - |\mathcal{I}| \left(\frac{2}{3} + 6\eta\right) 3\tau n \geq \left(\frac{\gamma + \beta - \tau}{2} + \mu - 20\eta\right) n|\mathcal{I}| \tag{2.6}$$

We will show that if the lemma is not true then we get a contradiction to the non-extremality of G. To see that assume that $\tau < \gamma + \beta + \alpha$. Plugging this in (2.5) (in the form $\gamma + \beta > \tau - \alpha$) we get

$$e(\mathcal{I}, V(\mathcal{T})) \geq (2\tau - 2\alpha/3)\, n|\mathcal{I}| \tag{2.7}$$

Similarly plugging $\tau < \gamma + \beta + \alpha$ in (2.6) we get

$$e(\mathcal{I}, V(\mathcal{M})) \geq (\mu - 2\alpha/3)\, n|\mathcal{I}| \tag{2.8}$$

From (2.7), (2.8) and (2.4), for almost every vertex $x \in \mathcal{I}$ we have $d(x, V(\mathcal{T})) \geq (2/3 - 4\alpha/5)$ and $d(x, V(\mathcal{M})) \geq (1/2 - 4\alpha/5)$. Furthermore, since Observation 6 does not hold for $\mathcal{C}$ we must have that almost all vertices in $\mathcal{I}$ have the same neighborhoods in $V(\mathcal{M})$. Indeed otherwise, there must be an η -fraction of bipartite graphs in $\mathcal{M}$ for which both of its color classes are η-dense to $\mathcal{I}$, which by Observation 6 contradicts the optimality of $\mathcal{C}$. Similar argument gives us that almost all vertices in $\mathcal{I}$ have the same neighborhood in $V(\mathcal{T})$.

Therefore we have that almost all vertices $(1 - \eta$-fraction) in $\mathcal{I}$ are connected to two color classes, (say V_2^l and V_3^l) of almost every $(1 - 4\alpha/5$ -fraction) tripartite graph

$(V_1^l, V_2^l, V_3^l) \in \mathcal{T}$. Similarly almost all vertices in $\mathcal{I}$ are connected to one color class, (say U_2^l) of almost every bipartite graph $(U_1^l, U_2^l) \in \mathcal{M}$.

Denote by (V_1, V_2, V_3) the union of the three color classes in the complete tripartite graphs and by (U_1, U_2) the union of the two color classes in the complete bipartite graphs. By the above density information, the following remarks and observation 6, we have

$$d(V_1),\ d(U_1),\ d(V_1, U_1),\ d(\mathcal{I}, V_1),\ d(\mathcal{I}, U_1)\ \text{ and }\ d(\mathcal{I})\ \text{ are all }\ < \alpha/2$$

Now using the fact that $3\tau + 2\mu + \beta = 1$ and the assumption that $\tau < \gamma + \beta + \alpha$ we get that

$$|V_1 \cup U_1 \cup \mathcal{I}| = (\tau + \mu + \beta)\, n = \left(\frac{3\tau + 2\mu + 2\beta - \tau}{2}\right) n \geq \left(\frac{1-\gamma}{2} - \alpha\right) n$$

which together with the above density information implies that G is in extremal case 1, a contradiction.

Case 2: There is no independent set ($\beta < \eta$ but $\mu > \eta$)

Again, since $\mathcal{C}$ is an optimal cover, none of the above observations holds, hence by observation 4 and 5 we have

$$d\left(V(\mathcal{M})\right) < \left(\frac{1}{2} + 4\eta\right) \quad \text{and} \quad d\left(V(\mathcal{M}), V(\mathcal{T})\right) < \left(\frac{2}{3} + 6\eta\right) \tag{2.9}$$

Using the minimum degree condition (2.3), (2.9) and the fact that $3\tau + 2\mu + \beta = 1$ we get

$$e\left(V(\mathcal{M}), V(\mathcal{T})\right) \geq \left(\frac{1+\gamma}{2} - \eta - \left(\frac{1}{2} + 4\eta\right) 2\mu - 2\eta\right) n|V(\mathcal{M})| \geq \left(\frac{3\tau + \beta + \gamma}{2} - 20\eta\right) n|V(\mathcal{M})| \tag{2.10}$$

and

$$e\left(V(\mathcal{M})\right) \geq \left(\frac{1+\gamma}{2} - \eta - \left(\frac{2}{3} + 6\eta\right) 3\tau - 2\eta\right) n|V(\mathcal{M})| \geq \left(\frac{\gamma + \beta - \tau}{2} + \mu - 20\eta\right) n|V(\mathcal{M})| \tag{2.11}$$

Similarly as above, we will show that if the lemma is not true then we get a contradiction to the non-extremality of G. To see that assume that $\tau < \gamma + \beta + \alpha$. Plugging $\beta + \gamma > \tau - \alpha$ in (2.10) we get

$$e\left(V(\mathcal{M}), V(\mathcal{T})\right) \geq \left(2\tau - 2\alpha/3\right) n|V(\mathcal{M})| \tag{2.12}$$

Furthermore, we get the graph induces by $V(\mathcal{M})$ is an almost complete bipartite graph. Similarly plugging $\tau < \gamma + \beta + \alpha$ in (2.11) we get

$$e(\mathcal{I}, V(\mathcal{M})) \geq (\mu - 2\alpha/3)\, n|V(\mathcal{M})| \tag{2.13}$$

From (2.12) and (2.9) for almost every vertex $x \in V(\mathcal{M})$, we have $d(x, V(\mathcal{T})) \geq (2/3 - 4\alpha/5)$. In addition, since Observation 5 does not hold, , without loss of generality, we must have that the connectivity structure of almost every tripartite graph (V_1^l, V_2^l, V_3^l) and almost every bipartite graph (U_1^j, U_2^j) is as follows: (V_1^l, U_2^j) and (V_2^l, U_1^j) are almost complete bipartite graphs and both (V_3^l, U_1^j) and (V_3^l, U_2^j) are almost complete bipartite graphs. This is so because otherwise, we get the situation as in Observation 5 and get a contradiction to the optimality of the cover.

Denote by (V_1, V_2, V_3) the union of the three color classes in the complete tripartite graphs and by (U_1, U_2) the union of the two color classes in the complete bipartite graphs. From(2.13) and (2.9) for almost every vertex $x \in V(\mathcal{M})$, we have $d(x, V(\mathcal{M})) \geq (1/2 - 4\alpha/5)$. Furthermore, we must have that the graph induced by $V(\mathcal{M})$ is almost a complete bipartite graph, because otherwise by Lemma 7 we can find a collection of complete tripartite graphs in $V(\mathcal{M})$ contradicting the optimality of the cover.

In this case almost every complete bipartite graph (U_1^j, U_2^j) behaves exactly like the complete bipartite graph (V_1^l, V_2^l) of almost every tripartite graph (V_1^l, V_2^l, V_3^l). Because by the above connectivity structure, V_1^l and V_l^2 are essentially replaceable with U_1^j and U_2^j, respectively.

Denote by (V_1, V_2, V_3) the union of the three color classes in the complete tripartite graphs and by (U_1, U_2) the union of the two color classes in the complete bipartite graphs. By the replaceability property we have that the graph induced by (V_1, V_2), too is an almost complete bipartite graph.

Hence the graph induced by $U_1 \cup U_2 \cup V_1 \cup V_2$ is an almost complete bipartite graph with color classes $U_1 \cup V_1$ and $U_2 \cup V_2$, in particular $d(U_1 \cup V_1) < \alpha$. On the other hand, $|U_1 \cup V_1| = \frac{(1-\tau-\eta)n}{2} \geq \left(\frac{1-\gamma}{2} - \alpha\right) n$, where the last inequality uses the assumption that $\tau < \gamma + \beta + \alpha$ and $\beta < \eta$. But this implies that G is in extremal case 1, a contradiction. $\square$

2.3.3 Embedding

In this optimal cover, $\mathcal{C} = (\mathcal{T}, \mathcal{M}, \mathcal{I})$, we are ready to describe the embedding procedure. Let the size of a color class in a tripartite graph in $\mathcal{T}$ and a bipartite graph in $\mathcal{M}$ be l. Let us assume first that we have $\tau \geq \gamma + \beta + \alpha^2/2$ in Lemma 12. The other case when $\tau \geq 1/3 - 2\eta$ in Lemma 12 (the case of almost all triangles) is postponed until Section 2.3.5. Furthermore, we assume that in the cycle system H, all cycles are of length at most $\eta^2 l$. The embedding procedure for the case H has some cycles of length more than $\eta^2 l$ is given in Section 2.3.4.

By Lemma 12, it is easy to see that the total number of vertices of cycles in H, that are of length at least 4, (we refer to them as non-triangle cycle) is at least $(2\mu + 2\beta + \alpha^2/2)n$. Roughly speaking our embedding procedure is outlines as follows: We will start the embedding with cycles of length at least 4 and first we will eliminate the vertices in $\mathcal{M}$ and $\mathcal{I}$ (if the exist). We begin with embedding odd cycles of length at least 5, such that for each such cycle we use exactly three vertices from some $T_i \in \mathcal{T}$ (one triangle) and the remaining vertices are embedded into vertices of some $B_j \in \mathcal{M}$. If all the non-triangle odd cycles are embedded at this step and there are still some bipartite graphs left in $\mathcal{M}$, we embed even cycles to it. Having used almost all vertices of $\mathcal{M}$, we start embedding the remaining non-triangle cycles using vertices of $\mathcal{I}$. We use one triangle in $\mathcal{T}$ for each vertex in $\mathcal{I}$ and one additional triangle for each odd cycle.

From the above bound on the total number of vertices of non-triangle cycles, we still have some non-triangles left to embed. These and the triangles in H are embedded with an application of Lemma 11.

For each color class in the complete tripartite graphs in $\mathcal{T}$ we set aside a random α^3-portion as a *buffer zone*. At this stage will not embed any vertices into the buffer zones; they will be used later to finish the embedding.

Handling the vertices in $\mathcal{M}$

If $\mu < \eta$, then we add the vertices in $\mathcal{M}$ to our exceptional set V_0. Otherwise assuming $\mu > \eta$, we do the following. We add to V_0 all vertices of bipartite graphs that do satisfy Observation 4. Let $B_i = (U_1^i, U_2^i)$ be a bipartite graph in $\mathcal{M}$, Let (x, y) be a typical edge in B_i. By the minimum degree condition, Observation 2, Observation 4 we have

$$deg(x, V(\mathcal{T})) + deg(y, V(\mathcal{T})) \geq (3\tau + \gamma - \eta)n$$

Let $\mathcal{T}_{xy} = \{T_i \in \mathcal{T} : deg(x, T_i) + deg(y, T_i) \geq (3 + \eta)l\}$, by the above lower bound and Observation 5, we get that $|V(\mathcal{T}_{xy})| \geq (\gamma - \eta)n$. This implies that if we still have many unembedded odd non-triangle cycles in H, than there is an unused triangle, $T = (a, b, c)$ in $\mathcal{T}_{xy}$ such that $deg(x, T) + deg(y, T) \geq 4$. Hence, we embed the next odd cycle, C_j onto a path of length $j - 3$ in B_i from x to y and we close the cycle with T. We continue in this fashion until we have no more than $\eta^2 l$ vertices left in B_i. When the remaining part of B_i is less than $\eta^2 l$, we add it to V_0. Clearly we can embed odd cycles in this way, as long as there is some $B_i \in \mathcal{M}$ left. In case almost all $((\gamma - \eta)n)$ odd cycles are embedded, we embed even cycles to the remaining part of B_i in the straight forward greedy way. If a cycle doesn't fit the remaining part of B_i (that is number of vertices in B_i is less than $\eta^2 l$) we add B_i to V_0, and continue with the next bipartite graph in $\mathcal{M}$. Denote by $\gamma' n$ the number of odd cycles remaining in H, and let $\mathcal{T}'$ be the remaining unused triangles in $\mathcal{T}$. Let the remaining set of cycles to be embedded be H'

Handling the vertices in $\mathcal{I}$

Let a be the average length of odd non-triangle cycles in H'. We randomly partition these non-triangle odd cycles into groups, such that each group has t cycles (t to be determined later). Let $\mathcal{P} = \{P_1, P_2, \ldots, P_q\}$ be the groups. Let p_i be the average length of cycles in the group P_i. By the law of large numbers, we may assume that $p_i \sim a$.

Again if $\beta < \eta$ then we add the vertices of $\mathcal{I}$ to V_0. Otherwise we do the following. We say that a vertex $x \in \mathcal{I}$ is *good* for a $T_j \in \mathcal{T}'$ if $deg(x, T_j) \geq (\frac{3}{2} + \frac{\beta+\gamma'}{6\tau} - \eta)l$. Note that by the minimum degree condition and Observation 1 and 2, this is the average degree of a typical vertex in $\mathcal{I}$ to a tripartite graph. Simple calculation using the minimum degree and these observations gives us that a typical vertex x is good for at least η-fraction of tripartite graphs. Counting from the other side we get that there is a tripartite graph $T_j \in \mathcal{T}'$, that is good for a set $\mathcal{I}^*$ of size at least $\eta|\mathcal{I}|$. Hence there are many vertices (at least $\sqrt{n}$) vertices that have the same neighborhood in T_j, (we still denote it by $\mathcal{I}^*$.

Let $T_j = (V_1^j, V_2^j, V_3^j)$ be such a tripartite graph in $\mathcal{T}'$. By observation 3 we may assume, without loss of generality, that $d(\mathcal{I}^*, V_1^j) \geq \eta$ and $d(\mathcal{I}^*, V_2^j) \geq \eta$.

We embed the next two even cycles $C_q, C_{q+1} \in H'$ as follows: For C_q we greedily map $\frac{|C_q|}{2}$ vertices of C_q each to unused vertices in $\mathcal{I}^*$ and V_1^j. For C_{q+1} we greedily map $\frac{|C_{q+1}|}{2}$ vertices of C_{j+1} to unused vertices in V_2^j and V_3^j. We similarly alternate the remaining even cycles, such that we use as equal as possible number of vertices from each color class of T_j. Using the fact that all cycles are of length at most $\eta^2 l$, we may assume that the remaining tripartite graph is still almost balanced.

Clearly with this procedure we can either exhaust almost all vertices of $\mathcal{I}$ (at least $(1-\eta)$-fraction), or we can embed almost all even cycles. In case there are still unused vertices in $\mathcal{I}$, (we keep the same notation for the remaining vertices). We will now embed odd cycles in the groups as follows. Again let T_j be a good tripartite graph and let $\mathcal{I}^*$ be the subset of $\mathcal{I}$ as above.

Let $m = \frac{t(a-3)}{4}$ and take the next m/t groups of odd cycles, $P_{i+1}, \ldots, P_{i+\frac{m}{t}}$. For $1 \leq q \leq \frac{m}{t}$ We embed $C_s \in P_{i+q}$ by mapping $\frac{|C_s|-1}{2}$ vertices to unused vertices in $\mathcal{I}^*$ and $\frac{|C_j|-1}{2}$ vertices to unused vertices in V_1^j, while the cycle is closed by embedding the last vertex onto an unused vertex in V_2^j. Take the next $\frac{m}{t}$ groups of odd cycles, $P_{i+\frac{m}{t}+1}, \ldots, P_{i+\frac{2m}{t}}$. Again, for $1 \leq q \leq \frac{m}{t}$, embed $C_s \in P_{i+\frac{m}{t}+q}$, by mapping $\frac{|C_s|-1}{2}$ vertices to unused vertices in $\mathcal{I}^*$ and $\frac{|C_s|-1}{2}$ vertices to unused vertices in V_2^j. Each cycle is closed using a vertex in V_1^j.

To restore the balance somewhat among the remaining part of T_j, take the next $\frac{m}{t}$ groups of odd cycles, $P_{i+\frac{2m}{t}+1}, \ldots, P_{i+\frac{3m}{t}}$. For $1 \leq q \leq \frac{m}{t}$, embed $C_s \in P_{i+\frac{2m}{t}+q}$, by mapping $\frac{|C_s|-1}{2}$ vertices to unused vertices in V_3^j and $\frac{|C_s|-1}{2}$ vertices to unused vertices in V_1^j, while the cycle is closed using a vertex in V_2^j. Similarly take the next $\frac{m}{t}$ groups of odd cycles, $P_{i+\frac{3m}{t}+1}, \ldots, P_{i+\frac{4m}{t}}$. For $1 \leq q \leq \frac{m}{t}$, Embed $C_s \in P_{i+\frac{3m}{t}+q}$, by mapping $\frac{|C_s|-1}{2}$ vertices to unused vertices in V_3^j and $\frac{|C_s|-1}{2}$ vertices to unused vertices in V_2^j, while the cycle is closed using a vertex in V_1^j.

Following this procedure we have used $\frac{m}{t} \cdot \frac{t(a-1)}{2} + \frac{m}{t} \cdot t + \frac{m}{t} \cdot \frac{t(a-1)}{2} + \frac{m}{t} \cdot t = ma + m$ vertices each in V_1^j and V_2^j. While the number of vertices used in V_3^j is $\frac{m}{t} \cdot \frac{t(a-1)}{2} + \frac{m}{t} \cdot \frac{t(a-1)}{2} = ma - m$, So the remaining part of V_3^j has $2m$ extra vertices compared to V_1^j and V_2^j.

To restore the balance completely, we map the next two groups as follows: For $C_s \in P_{i+\frac{4m}{t}+1}$ embed it by mapping $\frac{|C_s|-1}{2}$ vertices in V_3^j and $\frac{|C_s|-1}{2}$ vertices in V_1^j, closing each cycle with a vertex in V_2^j. While embed $C_s \in P_{i+\frac{4m}{t}+2}$ by mapping $\frac{|C_s|-1}{2}$ vertices in V_3^j and $\frac{|C_s|-1}{2}$ vertices in V_2^j, closing each cycle with a vertex in V_1^j. The number of vertices used in this step in V_3^j is $t(a-1)$ and that in V_1^j and V_2^j is $\frac{ta+t}{2}$. Using the value of m clearly the number of vertices used in all 3 color classes are the same hence the remaining part of T_j is balanced.

By the definition of goodness, if we choose t such that ma is less than $\min\{deg(x, V_1^j), deg(x, V_2^j)\}$, for any $x \in \mathcal{I}^*$ and such that $ma + m$ is much less than l, then clearly we can consume almost all ($\geq (1-\eta)$-fraction) of the vertices in $\mathcal{I}$.

Dealing with vertices in V_0

In the above procedure we assigned the vertices in $\mathcal{I}$ and $\mathcal{M}$ to cycles in such a way, that at most one triangle for each odd cycle and one triangle for each vertex in $\mathcal{I}$ will be used. Denote by τ' the corresponding quantity in the remainder of the complete tripartite graphs and let $\gamma'' = k'/n$ (where k' is the number of remaining odd cycles), since we had $\tau > \gamma + \beta + \alpha^2/2$ in Lemma 12, we still have $\tau' > \gamma'' + \alpha^2/2$. This implies in particular that the number of vertices in the cycles with length at least 4 is still at least $\alpha^2 n/2$.

We assign the triangles in H to the tripartite graphs in such a way that the number of remaining unassigned vertices to each tripartite graph is as equal as possible. The remaining tripartite graphs are balanced, and in every tripartite graph we still have at least α^3-portion of unassigned vertices (the buffer zone). In the remaining portion of each tripartite graph, we embed the remaining non-triangle cycles proportionally, i.e. such that the total number of vertices in cycles embedded in each tripartite graph is as equal as possible. For each $T_i(V_1^i, V_2^i, V_3^i)$, an η-fraction of these cycles are embedded in the bipartite graphs $(V_1^i, V_2^i), (V_2^i, V_3^i)$ and (V_1^i, V_3^i) (using a vertex from the third color class to close the odd cycles). We call this the special embedding. The rest most of the cycles are embedded using the Embedding Lemma (Lemma 11). If the cycles do not fit exactly, we add to V_0 some vertices T_i.

At this stage we will assign vertices in the exceptional set, V_0 to the tripartite graphs. Note that from the above procedure it is clear that $|V_0| < 10\eta n$ and all the remaining cycles in H are of length at least 4. We will assign these remaining cycles to the tripartite graphs and eliminating vertices from V_0. Let C_j be an unassigned cycle, $4 \leq j \leq \eta^2 l$.

For a vertex v we define R_v to be the set of vertices which can be replaced by v, i.e. R_v is the set of all vertices X, such that if a cycle is embedded using x then we can embed the cycle using v and freeing up x. We will prove that R_v is about $\frac{1+\gamma}{2}n$ for all $v \in V(G)$. To prove that we will show that if v has neighbors in any color class of a bipartite or tripartite graph, then there is a corresponding set of the same size as its neighborhood, that can be replaced with v.

Let $T_i = (V_1^i, V_2^i, V_3^i)$ be a tripartite graph. Let $d(v, V_1^i) \geq \eta$. Then with high probability we have that v is connected to at least an $\eta/10$-fraction of the buffer zone of V_1^i. And the since there are some cycles embedded through the special embedding and set of vertices could be used for this embedding, as it is complete bipartite graph. Say a particular instance of such a cycle is embedded as $x_1, y_1, x_2, y_2, \ldots, x_s, y_s, z_1, x_1$, where $x_j \in V_1^i, y_j \in V_2^i$ and $z_1 \in V_3^i$ (if the cycle is odd). Let x' and x'' be any two vertices of this cycle, adjacent to v. the above cycle can be embedded like $x', v, x'', y_2, \ldots, x_s, y_s, z_1, x'$, and we replace x' and x'' with x_1 and x_2 and hence y_1 is freed, that is y_1 can be moved out to the set R_v. Similarly vertices assigned to cycles embedded in T_i to consume the independent set $\mathcal{I}^*$ can be moved out to R_v. Furthermore the vertices assigned to cycles embedded to consume vertices in $\mathcal{M}$ be moved out.

If v is highly connected to some $\mathcal{I}^*$ then we will show that corresponding V_1^i and V_2^i can be moved out to R_v. This is clearly the case as the cycles are embedded by going around $\mathcal{I}^*, V_1^i, \mathcal{I}^*, V_1^i, \ldots.$ And any set of vertices in $\mathcal{I}^*$ can be used for such embedding. So we may assume that the neighbor of v are used in a cycle. Now similarly rearrangement of the cycles shows that any vertex in $N(\mathcal{I}^*, V_1^i)$ can be moved out to R_v. Similar argument gives us that if v is highly connected to U_1 of some bipartite graph (U_1, U_2) then the vertices in U_2 can be freed up.

So if v has neighborhood somewhere we have a corresponding set at least as big as the size of neighborhood, which can be replaced with v, (moved out). Therefore we have that $|R_v| \geq (1 + \gamma - 20\eta)n/2$. Indeed from the fact that G does not satisfy the conditions of Extremal Case 2 we will show that $|R_v| \geq (1 + \gamma + \alpha^2)n/2$. Indeed let $R_v = \{x_1, \ldots\}$, an easy calculation shows that there is a set $R'_v \subset \cup_{x_i} R_{x_i}$ such that $|R'_v| \geq (1/2 + \gamma - 40\eta)n$ and for all $v \in R'_v$, we have $v \in R_{x_i}$ for at least ηl vertices x_i. Note that any vertex in R'_v still is replaceable by v, we call it the expansion process. If $|R'_v| < (1+\gamma+\alpha^2)n/2$ then it must be that that the neighborhood of almost all vertices in R_v is in R_v only. Hence $d(R_v, V(G) \setminus R_v) < \alpha$ which implies that we are in Extremal Case 2. We apply the expansion process for all v and we still call the exchangeable set R_v.

Consider any j points $a_1, a_2, \ldots, a_j$ in V_0. Since $|V_0| \leq 10\eta n$, $deg(a_i, V(G) \setminus V_0) \geq (1 + \gamma - 20\eta)n/2$ and hence $|R_{a_i}| \geq (1 + \gamma - 20\eta)n/2 : 1 \leq i \leq j$. We will assign C_j to a tripartite graph such that each tripartite graph is assigned almost the same number of vertices.

Since $|R_{a_i}| \geq (1+\gamma+\alpha^2)n/2$ for all $1 \leq i \leq j$ every color class of any tripartite graph has at least $\alpha^2 n/2$ neighbors in each of R_{a_i}. Hence we can find j vertices $b_1, b_2, \ldots, b_j$, $b_i \in R_{a_i}$ such that $deg(b_i, V_1) \geq \eta l$. Replacing each b_i with a_i we can say that the j points outsided are highly connected to a color class of some tripartite graph (say V_1 of T_i). We will use the vertices $b_1, \ldots, b_j$ for parts of the cycles that are embedded by special embedding. So j vertices in V_2 and V_3 can be freed up as discussed above. We assign the cycle C_j to T_i.

For the next unassigned cycle and points outside we do the same exchanging, choosing some other color class of some other tripartite graph, so that the vertices assigned to each tripartite graph remains as balanced as possible.

We still have the unassigned vertices (freed-up vertices) in each color class of each tripartite graphs almost balanced, because V_0 is at most $10\eta n$ and we assigned the remaining cycles in a balanced way, the maximum imbalance among the color classes of any tripartite graph is bounded by a factor of $\sqrt{\eta}$. Furthermore the total length of cycles assigned in this phase to T_i is exactly the same as unused vertices in the tripartite graphs, as we freed exactly j vertices for a C_j. Hence in each tripartite graph the assigned cycles can be embedded by an application of Lemma 11.

2.3.4 Embedding Large Cycles

In this case we have some cycles larger than $\eta^2 l = c \log n$ for some small constant c. It is easy to see that all arguments in the previous section work even when the minimum degree is slightly less, i.e it is sufficient to use that $\delta(G) \geq (1/2 + \gamma - \eta)n$, since G is α-non extremal and $\eta \ll \alpha$.

We use the following lemma from [3] for splitting the graphs into two subgraphs such that the relative minimum degree and non-extremality in the induced subgraphs is roughly the same as the original graph.

Lemma 13. *For $0 < \varepsilon < 1$, there exist an n_0, such that if H is an α-non-extremal graph on $n \geq n_0$ vertices with $\delta(H) \geq \lambda n$ then for any random subset A of $V(H)$, such that $\varepsilon n \leq |A| \leq (1-\varepsilon)n$, (let $B = V(G) \setminus A$), with high probability we have* *that* $\delta(H|_A) \geq (\lambda - n^{-\frac{1}{3}})|A|$ *and* $\delta(H|_B) \geq (\lambda - n^{-\frac{1}{3}})|B|$ *and both* $H|_A$ *and* $H|_B$ *are* $(\alpha - n^{-\frac{1}{3}})$*-non-extremal.*

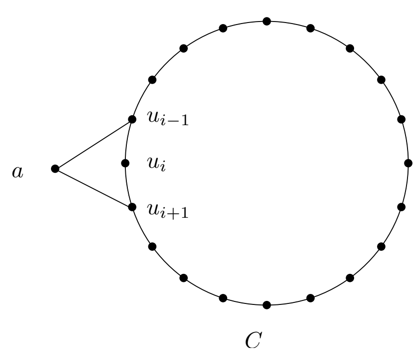
Furthermore we will make use of the following simple fact.

Fact 14 ([23])**.** *Every* 2*-connected graph* H *on* n *vertices has a cycle of length* $\min\{n, 2\delta(H)\}$*.*

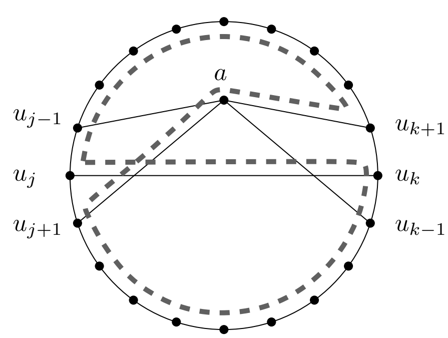
Using the fact that our graph G is α-non extremal we prove the following extensions of Dirac theorem [23] on Hamiltonian graphs and Bondy theorem [24] on pancyclic graph.

Lemma 15. *For* $0 < \eta \ll \alpha$ *there exist a constant* n_0 *such that if* H *is an* α*-non-extremal graph on* $n \geq n_0$ *vertices with* $\delta(H) \geq (1/2 - \eta)n$*, then* H *is Hamiltonian.*

Proof. Using Fact 14 we get a cycle $C = u_1, \ldots, u_q$; $q \geq (1 - 2\eta)n$. If $q < n$, we will insert the vertices outside C, to extend the cycle. Let a be an outside vertex. If a is connected to u_i, u_{i+1} then we can insert a between u_i and u_{i+1} to extend C. If not then by the minimum degree condition, a must be connected to many pairs like u_{i-1}, u_{i+1}. Let $R_a = \{u_i \in C : u_{i-1}, u_{i+1} \in N(a)\}$. Clearly any such u_i can be replaced by a without reducing length of the cycle (see Figure 2.2(e)). By the above observation $|R_a| \geq (1/2 - 4\eta)n$. Since we are not in Extremal case 1, $d(R_a) \geq \alpha$. Consider an edge (u_j, u_k) inside R_a. By definition a is connected to $u_{j-1}, u_{j+1}, u_{k-1}, u_{k+1}$ since there are many such edges we can assume $k > j + 5$. It is easy to see that we can make the following cycle $u_{j+1}, u_{j+2} \ldots, u_{k-1}, u_k, u_j, u_{j-1}, u_{j-2}, \ldots, u_{k+1}, a, u_{j+1}$. Hence C is extended (see Figure 2.2(f)) $\square$



(e) Replacing u_i with a



(f) The dashed-cycle extends C to include a

We now prove the following stronger statement.

Lemma 16. *For $0 < \eta \ll \alpha$ there exist a constant n_0 such that if H is an α-non-extremal graph on $n \geq n_0$ vertices with $\delta(H) \geq (1/2 - \eta)n$, then H is pancyclic i.e. H has a cycle of length q, $3 \leq q \leq n$.*

Proof. For $q \geq \varepsilon n$, $(0 < \varepsilon < \eta)$, we randomly choose a subset $A \subset V(H)$, $|A| = q$. By Lemma 13, $H|_A$ satisfies the conditions of Lemma 15, hence we are done. For $q < \varepsilon n$, We choose a path, $P = v_2, \ldots, v_{q-1}$ on $(q-2)$ vertices, (such a path obviously exits, as H is Hamiltonian). Since $N(v_2)$ and $N(v_{q-1})$ in $V(H) \setminus V(P)$ are both at least $(1/2 - 2\eta)n$. Hence by α-non extremality there are edges between $N(v_2)$ and $N(v_{q-1})$. Take such an edge (v_1, v_q) and $v_1, v_2, \ldots, v_{q-1}, v_q$ is a cycle on q vertices. $\square$

Now we prove our main theorem when the cycle system has larger cycles, i.e. at least one cycle is larger than $\eta^2 l = c \log n = m$, for some small constant c. We denote by H the given cycle system and by H_s the set of smaller cycles in H, (cycles of length at most m,) and let $H_l = H \setminus H_s$. Let M be the total number of vertices in cycles in H_s. We consider the following cases. Let ε be a positive constant such that $c \ll \varepsilon < \eta/10$

Case 1 $M < \varepsilon n$:

In this case we embed the cycles in H_s one by one using Lemma 16. The remaining induced graph still has minimum degree at least $(\frac{1+\gamma}{2} - 2\varepsilon)(n - M)$. The large cycles, cycles in H_l, are embedded in the remaining subgraph by the procedure as in Case 4.

Case 2 $\varepsilon n \leq M \leq (1 - \varepsilon)n$:

In this case we randomly partition $V(G)$ into A and B, $|A| = M$. By Lemma 13 we have $\delta(G|_A) \geq (\frac{1+\gamma}{2} - n^{-\frac{1}{3}})|A|$ and $\delta(G|_B) \geq (\frac{1+\gamma}{2} - n^{-\frac{1}{3}})|B|$ and both $G|_A$ and $G|_B$ are $(\alpha - n^{-\frac{1}{3}})$-non-extremal. We embed the cycles in H_s in $G|_A$ applying the procedure in the previous section and for embedding H_l in $G|_B$ we get Case 4.

Case 3 $M > (1 - \varepsilon)n$: Similarly as in Case 1, we embed the cycles in H_l one by one using Lemma 16. The remaining induced graph still has minimum degree at least $(\frac{1+\gamma}{2} - 2\varepsilon)M$ and is $(\alpha/2)$-non-extremal, so we use the procedure in the previous section to embed the cycles in H_s in the remaining subgraph.

Case 4 All cycles are of length at least m:

In this case since the order of graph could be $(n - M)$, for simplicity, we still say that G is a graph on n vertices and is α-non-extremal, and $\delta(G) \geq (1/2 - \eta)n$ and let H be the given cycle system. Let the cycles be $C_1, C_2, \ldots, C_r$ with length $n_1 \geq n_2 \geq \ldots \geq n_r$ respectively $n_i \geq m$. We consider two cases based on value of n_1.

Case 4.1 $n_1 \geq (1 - \eta)n$:

We embed all cycles C_i for $i > 1$, one by one using Lemma 16. It is easy to see that the remaining induced graph satisfies the conditions of Lemma 15, hence there is a Hamiltonian cycle C_1 in it.

Case 4.2 $n_i < (1 - \eta)n$ for $1 \leq i \leq r$:

We distributes the cycles in two sets H_1 and H_2, let n_A and n_B be the total number of vertices in cycles in H_1 and H_2 respectively. Since $\eta \ll 1$, we can distribute the cycles such that both n_A and n_B are at most $(1 - \eta)n$. We randomly partition $V(G)$ into two sets A and B such that $|A| = n_A$ and $|B| = n_B$. We will embed the cycles in H_1 and H_2 in $G|_A$ and $G|_B$ respectively.

We recursively apply the same splitting procedure until the condition of Case 4.1 is satisfied. Using the fact each $n_i > m$ we will show that minimum degree and non-extremality conditions hold till the very end of this process, hence the required cycles can be found. Define the normalized degree of a graph F as, $D(F) = \frac{\delta(F)}{|V(F)|}$. By definition $D(G) \geq (1/2 - \eta)$ therefore by Lemma 13 we have

$$D(G_A) \geq D(G) - n^{-\frac{1}{3}} \quad \text{and} \quad D(G_B) \geq D(G) - n^{-\frac{1}{3}}$$

Since the splitting process terminates before $n \leq m$ and each time the number of vertices is reduced by at least a factor of $(1 - \eta)$, using the fact that $m = c \log n$, for each final graph G_f we have

$$D(G_f) \geq D(G) - m^{-\frac{1}{3}} \sum_{i=0}^{\infty} (1 - \eta)^{i/3} \geq \frac{1}{2} - 10\eta$$

A similar computation shows that each final G_f is $\alpha/10$-non-extremal, hence the conditions of Lemma 16 are satisfied so we can apply the procedure of Case 4.1

2.3.5 The case of almost all triangles

We have $\tau \geq 1/3 - \eta$ in Lemma 12 so the cover consists almost entirely of complete tripartite graphs. If there are not too many odd cycles in the cycle system H, say $\gamma < 1/3 - 4\eta - \alpha^2/2$, then considering the set of vertices that are not covered by the tripartite graphs as our independent set $\mathcal{I}$, so $\beta \leq 3\eta$, the other inequality in Lemma 12 is also satisfied, since $\tau \geq 1/3 - \eta \geq \gamma + \beta + \alpha^2/2$, hence we can apply the previous embedding procedure. Therefore, we may assume that there are many odd cycles, $\gamma \geq 1/3 - 4\eta - \alpha^2/2$, which together with (2.3) and (4.2) imply that $\delta(G) \geq (2/3 - \alpha^2)n$. Furthermore, it follows that the H contains at least $(1/3 - 3\alpha^2)n$ triangles. Indeed, otherwise the total number of vertices is at least

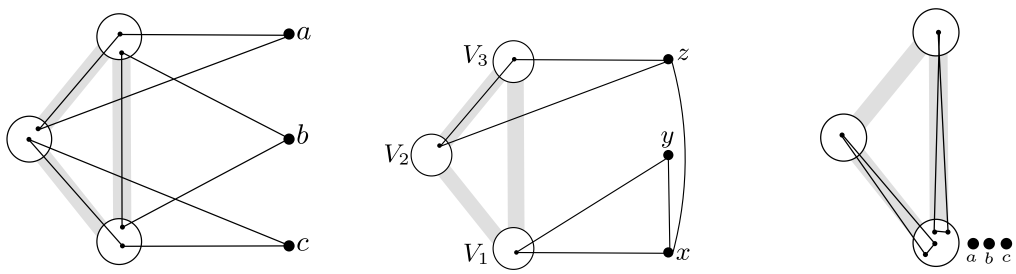
$$3(1/3 - 3\alpha^2)n + 5(2\alpha^2)n = n + \alpha^2 n > n,$$

a contradiction.

In the optimal cover we have at most $3\eta n$ points outside the complete tripartite graphs $\mathcal{T}$. By greedily embedding the non-triangles into the complete tripartite graphs (such that we do not embed too many vertices into each complete tripartite graph and then we keep the balance inside each tripartite graph) we may assume that we have only triangles left and thus the number of vertices outside the complete tripartite graphs is divisible by three. We will consider only three vertices outside $\mathcal{T}$ and extend the cover by one or more triangles to include these three vertices, such that the cover remains a balanced one. By repeating this procedure we eliminate all the vertices outside the complete tripartite graphs and then the remaining triangles of H can be embedded inside the complete tripartite graphs. Therefore, we consider only three vertices a, b and c outside $\mathcal{T}$ which do not make a triangle.

For $i \in \{1, 2, 3\}$ we say that a vertex v is *i-sided* to a tripartite graph $K_t = (V_1, V_2, V_3)$ if we have $d(v, K_t) \geq ((i-1)/3 + \eta)$, i.e. v has a large degree to at least i color classes. Denote by $s(v, K_t)$ the largest i for which v is i-sided to K_t.

If v is two-sided to K_t (say to the pair (V_2, V_3)) then we say that V_1 is *exchangeable* with v, i.e. any of the vertices in V_1 can be exchanged with v while keeping the cover balanced. Similarly, if v is three-sided to K_t then all three color classes are exchangeable with v. For a vertex v, define R_v to be the set of vertices that are exchangeable with v over all tripartite graphs. By the minimum degree condition, for any vertex v we have $|R_v| \geq (1/3 - \alpha)n$. Note that when we exchange a vertex v with a vertex from a color class V_i, then the new vertex has density 1 to the other two color classes, so we may assume that this is true already for v; a fact that will be important later. Furthermore, whenever a vertex v is exchanged with a vertex of a complete tripartite graph K_t, then we immediately cover it with a triangle in K_t to maintain the property that we still have a balanced complete tripartite graph.



(g) a, b and c are 2-sided to different pairs

(h) x is 1-sided to V_1 and y and z are 3-sided

(i) Replacing a,b and c in V_1

If $s(v, K_i) \leq 1$ for a $\sqrt{\alpha}$-fraction of the K_i's in $\mathcal{T}$, then there is at least a $\frac{\sqrt{\alpha}}{2}$-fraction of the other K_j's in $\mathcal{T}$ such that $s(v, K_j) = 3$. However, then from each such K_j all three color classes are exchangeable with v, hence $|R_v| > (1 + \sqrt{\alpha})n/3$. So R_v is small (smaller than $(1 + \sqrt{\alpha})n/3$) only if v is two-sided to most of the tripartite graphs.

Furthermore, from the minimum degree condition, it is easy to see that for any two vertices v and w, on average we have $s(v, K_t) + s(w, K_t) \geq 4$ for the complete tripartite graphs $K_t \in \mathcal{T}$. Note also that if there exists a $K_t \in \mathcal{T}$ such that a, b and c are two-sided to 3 different pairs of color classes in K_t (this happens for example if a, b and c are all three-sided to K_t), then we can easily expand the cover by one triangle such that we eliminate a, b and c and we keep the balance inside K_t (see Figure 2.2(g)) so we may assume that this is never the case.

We consider the following cases depending on the degree distribution of the vertices a, b and c.

Case 1: a, b and c are not all at least two-sided to most ($\geq (1 - \sqrt{\alpha})$-fraction) of the tripartite graphs in $\mathcal{T}$ (i.e. there exists a vertex in $\{a, b, c\}$, say c, such that for at least a $\sqrt{\alpha}$-fraction of the K_t's we have $s(c, K_t) \leq 1$).

We assume that there is no tripartite graph which has a color class exchangeable with all of a, b and c. Indeed suppose that there is some $V_i = \{x_1, \ldots, x_t\}$ exchangeable with all a, b and c. An easy calculation shows that there is a set $R_{abc} \subset \cup_{x_i} R_{x_i}$ such that $|R_{abc}| \geq (1/3 - \alpha)n$ and for all $v \in R_{abc}$, we have $v \in R_{x_i}$ for at least three vertices x_i. We call this process the *expansion process.* Then any three vertices in R_{abc} are exchangeable with a, b and c (we say that a, b and c *collapse*). Now for any $K_t \in \mathcal{T}$ there are at most two vertices in R_{abc} that are three-sided to K_t, because otherwise we get three vertices exchangeable with a, b and c that are all three-sided to K_t and we assumed that this is not the case. This implies that most vertices in R_{abc} are two-sided to almost all ($\geq (1 - \alpha)$-fraction) of the tripartite graphs and we are in Case 2.

It is easy to see that there exists a $K_t = (V_1, V_2, V_3)$ in $\mathcal{T}$, such that one of the V_i's (say V_1) is exchangeable with *both* a and b. Indeed, the condition in Case 1 implies that we can find a K_t such that $s(c, K_t) \leq 1$ but $s(a, K_t) + s(b, K_t) + s(c, K_t) \geq 6$. This implies $s(a, K_t) + s(b, K_t) \geq 5$ and the statement follows. We expand V_1 to R_{ab} as above such that $|R_{ab}| \geq (1/3 - \alpha)n$.

Consider a vertex $x \in R_c$ and a $K_t \in \mathcal{T}$. There is at most one vertex $y \in R_{ab}$ such that $d(x, K_t) + d(y, K_t) \geq (4/3 + \eta)$, because otherwise either a, b and c collapse or we have three vertices that are two-sided to three different pairs. This and the minimum degree condition imply that for most vertices $y \in R_{ab}$ we have the following density condition to almost all ($\geq (1 - \alpha)$-fraction) of the tripartite graphs K_t:

$$(4/3 - \alpha) \leq d(x, K_t) + d(y, K_t) \leq (4/3 + \eta). \tag{2.14}$$

Furthermore, there is no vertex x in R_c which is at least two-sided to at least a $(1-\sqrt{\alpha})$-fraction of the tripartite graphs. Otherwise, all but at most one vertices in R_{ab} must be three-sided to at least a $2\sqrt{\alpha}$-fraction of the tripartite graphs. Indeed, as we have no two vertices in R_{ab} that are two-sided to a $(1 - 4\sqrt{\alpha})$-fraction of the tripartite graphs (since we would get Case 2 with $4\sqrt{\alpha}$ instead of $\sqrt{\alpha}$), all but at most one vertices must be one-sided to at least a $4\sqrt{\alpha}$-fraction of the tripartite graphs and thus in fact they must be three-sided to at least a $2\sqrt{\alpha}$-fraction of the tripartite graphs. But then for this $x \in R_c$ there exist $y, z \in R_{ab}$ and a tripartite graph $K_t \in \mathcal{T}$ such that $s(x, K_t) \geq 2, s(y, K_t) = s(z, K_t) = 3$, which implies that a, b and c collapse. Note that by the above remark on $|R_v|$, this also implies that $|R_c| \geq (1 + \sqrt{\alpha})n/3$.

Since R_c is large, we have $d(R_{ab}, R_c) > \alpha$. As argued above no vertex in R_c is two-sided to most of the tripartite graphs in $\mathcal{T}$ therefore for any $x \in R_c$ there is at least a $\sqrt{\alpha}$-fraction K_t's in $\mathcal{T}$ such that $s(x, K_t) \leq 1$. Using (2.14) consider $x \in R_c$ and $y, z \in R_{ab} \cap N(x)$ and a tripartite graph $K_t = (V_1, V_2, V_3)$ in $\mathcal{T}$, such that

$$s(x, K_t) \leq 1, d(x, K_t) + d(y, K_t) \geq (4/3 - \alpha) \text{ and } d(x, K_t) + d(z, K_t) \geq (4/3 - \alpha).$$

Therefore y and x have a common neighbor, say $u \in V_1$ so u, x, y is a triangle. Furthermore, z must be three-sided to K_t and thus we can find a triangle z, v, w with some vertices $v \in V_2, w \in V_3$, thus eliminating a, b and c and keeping the tripartite graphs balanced (see Figure 2.2(h)).

Case 2: a, b and c are all at least two-sided to most ($\geq (1 - \sqrt{\alpha})$-fraction) of the tripartite graphs in $\mathcal{T}$.

Consider a typical tripartite graph K_t. As noted above a, b and c are not two-sided to three different pairs, therefore, at least 2 vertices (say a and b) are two-sided to the same pair (say (V_1, V_2)) in K_t. Hence V_3 is exchangeable with both a and b. We expand V_3 to get R_{ab} of size at least $(1/3 - \alpha)n$.

Case 2.1: R_c and R_{ab} are disjoint (so a, b and c do *not* collapse).

In this case there are no three vertices $x \in R_c$, $y, z \in R_{ab}$, such that they are two-sided to the same pair of a tripartite graph, since otherwise a, b and c collapse.

Consider a typical tripartite graph $K_t = (V_1, V_2, V_3)$. Let $P_i \subset R_{ab}$, $1 \leq i \leq 3$, be the set of those vertices that are exchangeable with V_i. Note that at most two of the P_i's are non-empty. Indeed, otherwise if all three are non-empty, then we take a vertex in R_c that is exchangeable with a color class in K_t (say V_1) and one vertex each from P_2 and P_3 to have 3 vertices two-sided to three different pairs of K_t and then we are done. Assume that P_3 is empty and both P_1 and P_2 have at least 2 vertices. Similarly we group vertices in R_c into $Q_i, 1 \leq i \leq 3$. Since both P_1 and P_2 are non-empty, Q_3 must be empty. It is easy to see that if $|P_i| \geq 2$ and Q_i is non empty then a, b and c collapse.

Hence without loss of generality we may assume that only $|P_1| \geq 2$ and Q_2 and/or Q_3 are non empty (say Q_2 is non-empty). But now since we are not in Extremal Case 1, we have $d(P_1) > \alpha$. Thus there is an edge (x, y) inside P_1 and a common neighbor v in V_2, so (x, y, v) is a triangle. Also by the definition of Q_2, there are many triangles (z, u, w) with $z \in Q_2$, $u \in V_1$ and $w \in V_3$. Then removing these two triangles and replacing x, y and z with a, b and c, respectively, the remaining cover is balanced.

Case 2.2: R_c and R_{ab} are not disjoint.

In this case R_{abc} is non-empty, by expansion we have $|R_{abc}| \geq (1/3 - \alpha)n$ and any 3 vertices of R_{abc} are exchangeable with a, b and c. Assume first that $|R_{abc}| \leq (1+\sqrt{\alpha})n/3$. Then as noted above, any 3 vertices in R_{abc} are two-sided to the same pair of almost all tripartite graphs (say the pair is always (V_2, V_3)). This implies that all points in R_{abc} are exchangeable with each other. The fact that we are not in Extremal Case 1 implies $d(R_{abc}) \geq \alpha$, therefore by exchanging some vertices we can introduce edges inside V_1 of some tripartite graph K_t. Now for such a K_t, we take a triangle (x_1, y_1, z_1) with $x_1 \in V_2, y_1, z_1 \in V_1$ and another triangle (x_2, y_2, z_2) with $x_2 \in V_3, y_2, z_2 \in V_1$. Replacing three of the vertices used in these triangles from V_1 by a, b and c, we get the remaining cover balanced (see Figure 2.2(i)).

If $|R_{abc}| > (1 + \sqrt{\alpha})n/3$, and we can not finish the task, then we will show that $|R_{abc}| \geq (2/3 - \alpha^{1/3})n$. Indeed otherwise, say there is an $\alpha^{1/3}$-fraction of tripartite graphs, $\mathcal{T}' \subset \mathcal{T}$, where only one color class (say V_1) is part of R_{abc}. Most vertices in R_{abc} must be two-sided to (V_2, V_3) in most tripartite graphs in $\mathcal{T}'$ (as in Case 1, since otherwise we would get three vertices that are three-sided to the same K_t). Since now R_{abc} is large, by using the minimum degree condition there are many edges from the union of the V_1's in $\mathcal{T}'$ to R_{abc}; these edges can be brought inside V_1 for some tripartite graph in $\mathcal{T}'$, hence as above we can finish the task. Now if $|R_{abc}| \geq (2/3 - \alpha^{1/3})n$, then for any vertex $x \in R_{abc}$, $|N(x) \cap R_{abc})| \geq (1/3 - 2\alpha^{1/3})n$ and again by non-extremality there are edges inside $N(x) \cap R_{abc}$ which make triangles with x. We are done since we found a triangle in R_{abc} and those three vertices can be replaced by a, b and c. This finishes the proof of the case of almost all triangles.

2.4 The Extremal Cases

In the extremal cases we will repeatedly use the following simple fact.

Fact 17. *If $G(A, B)$ is a bipartite graph, with*

- $deg(a, B) \geq (1 - \eta)|B|, \ \ \forall\, a \in A$
- $deg(b, A) \geq (1 - \eta)|A|, \ \ \forall\, b \in B$
- *B has a matching M of size k and $|B| = |A| + k$*

then if $\mathcal{C}$ is a set of cycles, such that $\mathcal{C}$ has k odd cycles and the sum of the lengths of cycles in $\mathcal{C}$ is $|A| + |B|$, then $\mathcal{C}$ can be embedded into G.

The basic idea is to use one edge of M for each odd cycle and all other edges of the cycles will be found in the almost complete bipartite graph between A and B. We assign to each cycle C, $\lceil \frac{|C|}{2} \rceil$ vertices in B. To the next cycle $C_i \in \mathcal{C}$, with $|C_i| = 2s$, we assign s unassigned vertices $x_i^1, x_i^2, \ldots, x_i^s$ in B that are disjoint from M. For C_i, with $|C_i| = 2s + 1$, we assign $s - 1$ unassigned vertices $x_i^1, x_i^2, \ldots, x_i^{s-1}$ disjoint from M and an unassigned edge $e_i = (y_i^1, y_i^2)$ from M. Now we make an auxiliary bipartite graph $H = (A, B')$. Such that B' has a vertex corresponding to successive pairs of vertices assigned to C_i, i.e. $(x_i^1, x_i^2), (x_i^2, x_i^3), \ldots, (x_i^s, x_i^1)$ if $|C_i| = 2s$. For each odd cycles C_i, B' has a vertex for pairs $(y_i^1, x_i^1), (x_i^1, x_i^2), (x_i^2, x_i^3), \ldots, (x_i^{s-1}, y_i^2)$. In H, every vertex in B' is connected to common neighbors of the vertices in the corresponding pair in G.

Clearly we have $|A| = |B'|$ and the minimum degree of a vertex in H is at least $(1 - 2\eta)|A|$. Hence by König-Hall theorem H has a perfect matching. It is easy to see that any perfect matching in H corresponds to an embedding of $\mathcal{C}$ in G. $\square$

2.4.1 Extremal Case 1

Here our graph G satisfies (2.1) and we are in Extremal Case 1.

Extremal Case 1 (EC1) with parameter α: *There exists an $A \subset V(G)$ such that*

- $|A| \geq \frac{n-k}{2} - \alpha n$*, and*

- $d(A) < \alpha$.

By adding or deleting vertices to or from A we may achieve that $|A| = (n-k)/2$ and $|B| = (n+k)/2$ (note that these are always integers). Furthermore, an easy computation shows that we still have $d(A) < 10\alpha$ (for simplicity we keep the notation A, B). This and (2.1) imply that we have

$$d(A, B) > 1 - 10\alpha. \tag{2.15}$$

Thus roughly speaking, we have an almost complete bipartite graph between A and B, the basic idea is to find a matching of size k (using (2.1) and Lemma 8 as we have $\delta(G|_B) \geq k$) in B and then use Fact 17 to find the cycle. However, we have to deal with certain exceptional vertices first.

A vertex $v \in A$ (similarly in B) is called *exceptional* if it is *not* connected to most of the vertices in the other set, more precisely if we have

$$deg(v, B) \leq (1 - \sqrt{10\alpha})|B|.$$

Let us denote the set of exceptional vertices by E_A in A and by E_B in B. From (2.15) we get that we have few exceptional vertices

$$|E_A| \leq \sqrt{10\alpha}|A| \text{ and } |E_B| \leq \sqrt{10\alpha}|B|.$$

Next we further refine the definition of exceptional vertices: an exceptional vertex $v \in A$ (similarly in B) is called *strongly exceptional* if it is connected to few vertices in the other set, more precisely if we have

$$deg(v, B) \leq \alpha^{1/3}|B|.$$

Denote the set of strongly exceptional vertices by $SE_A(\subset E_A)$ in A and by $SE_B(\subset E_B)$ in B. From (2.15) it is clear that

$$|SE_A| \leq 20\alpha|A| \text{ and } |SE_B| \leq 20\alpha|B|.$$

If we have a $u \in SE_A$ and a $v \in SE_B$, then we can exchange the two vertices and they will not be strongly exceptional anymore in their new sets. Thus we may assume that one of the sets SE_A and SE_B is empty, since otherwise we may reduce the number of strongly exceptional vertices in both sets. Then we only have to eliminate the vertices in the non-empty set. For the remainder of this extremal case we will distinguish three subcases depending on the size of $\gamma = k/n$.

Case 1: $\gamma \leq \alpha^{1/3}$.

Without loss of generality we assume that $SE_B = \emptyset$ and $SE_A \neq \emptyset$ (the other case is similar). We may assume that we have no $u \in B$ with $deg(u, B) > \alpha^{1/3}|B|$, since otherwise we can exchange this vertex with a vertex $v \in SE_A$ and thus reducing the size of SE_A. Therefore any vertex $v \in B$ can be exchanged with any vertex $u \in SE_A$ without changing the degree conditions. First for each $v \in SE_A$ we want to find a path P_v of length 2 such that the center of P_v is v, the other two vertices are in B and the paths are vertex disjoint for different v. Consider the graph $H = G|_{SE_A \cup B}$, from (2.1) we have $\delta(H) \geq |SE_A|$ (in fact in this case the minimum degree is at least $k + |SE_A|$, but in the other case when $SE_B \neq \emptyset$ the minimum degree could be $|SE_B|$). From the size of SE_A and the fact that no vertex has high degree in B, we have $\Delta(H) \leq 2\alpha^{1/3}|H|$. Therefore by Lemma 8 there are at least SE_A vertex disjoint paths of length 2 in H. Since every vertex in B can be exchanged with any vertex in SE_A, we can assume that all these paths have the center vertex in SE_A and the two end points in B.

Furthermore since the minimum degree in B is at least k and no vertex has degree more than $\alpha^{1/3}|B|$, by Lemma 8 we can find a matching of size k in B. From the fact that SE_A is very small, there exists a matching of size k that is vertex disjoint from all P_v selected above.

Then first we eliminate the paths of length 2 by embedding cycle parts into them. Note that the endpoints of each P_v have very high degree in B, so the endpoints of any P_v and P_u can be connected in one step. Therefore, by a simple greedy procedure we can embed cycles in the bipartite graph between A and B, that use all P_v and use exactly one edge inside B for each odd cycle. The remaining exceptional vertices (EA and EB) can also be used similarly, using the fact that their degree across is much larger than their number. Finally in the leftover almost-complete bipartite graph we finish the embedding such that the k' matching edges are used for the remaining k' odd cycles using Fact 17.

Case 2: $\alpha^{1/3} < \gamma \leq (\frac{1}{3} - \alpha^{1/3})$.

In this case we know that $SE_A = \emptyset$, as for each $v \in A$ we have

$$deg(v, B) \geq k > \alpha^{1/3} n \geq \alpha^{1/3} |B|.$$

If $SE_B \neq \emptyset$, then similarly as in Case 1, we find $|SE_B|$ vertex disjoint paths of length 2 such that the middle vertices are in SE_B. We also find a matching M of size k, such that at least one of the endpoints of each edge is non-exceptional and M is disjoint from the length-2 paths. Then, as in Case 1, we first use the length-2 paths to embed a few cycles; the edges in M are used for the k odd cycles and we finish the embedding using Fact 17.

Case 3: $\gamma > (\frac{1}{3} - \alpha^{1/3})$.

In this case most of the cycles are triangles, indeed the number of vertices covered by cycles of length at least 4 is at most $8\alpha^{1/3} n$. Therefore in this case we will use triangles to use up the exceptional vertices. eliminating the exceptional vertices. Assume first that we are *not* in the following two subcases:

Subcase 3.1: *There is a partition* $B = B_1 \cup B_2$ *with* $\left(\frac{1}{3} - \alpha^{1/3}\right) n \leq |B_1| \leq \frac{n}{3}$ *and* $d(B_1) < \alpha^{1/4}$.

Subcase 3.2: *There is a partition* $B = B_1 \cup B_2$ *with* $\left(\frac{1}{3} - \alpha^{1/3}\right) n \leq |B_1| \leq \frac{n}{3}$ *and* $d(B_1, B_2) < \alpha^{1/4}$.

Let us consider an exceptional vertex $u \in A$ (note that we have $SE_A = \emptyset$) and its neighborhood in B of size at least $k \geq (\frac{1}{3} - \alpha^{1/3})n$. As we are not in Subcase 3.1, we have edges inside this set and thus we can cover u with a triangle where the other two vertices come from B, as desired. For an exceptional vertex $v \in (B \setminus SE_B)$ again we can easily find a triangle where one of the other two vertices comes from A, the other from B. Finally let us take a vertex $v \in SE_B$. We may assume that we have no $u \in A$ with $deg(u, A) > \alpha^{1/3}|A|$, since otherwise we can exchange this vertex with v and thus reducing the size of SE_B. As in Case 1, since the minimum degree in the graph induced by $A \cup SE_B$ is at least $|SE_B|$ we can find $|SE_B|$ vertex disjoint edges going between A and SE_B. For each such edge e, since both of its endpoints have very high degree in

B, we find a common neighbor in B for the endpoints to get a vertex disjoint triangle. Using the fact that we are not in Subcases 3.1 and 3.2, we can find a matching of size k' (the remaining number of odd cycles) in the leftover of B, and then Fact 17 finishes the embedding.

Finally let us assume that we are in Subcase 3.1 (Subcase 3.2 is similar). By greedily embedding the few cycles first that are not triangles we may assume that we have only triangles left and we have three sets A, B, C of equal size. We will have two types of strongly exceptional vertices in each set; $v \in A$ is called *strongly exceptional to* B if it is connected to few vertices in B, more precisely if we have

$$deg(v, B) \leq \alpha^{1/3}|B|.$$

Denote the set of these vertices by SE_A^B. SE_A^C and the strongly exceptional sets in B and C are defined similarly. We describe how to eliminate the vertices in SE_A^B; the others are similar. To eliminate vertices in SE_A^B, for each vertex v in SE_A^B we find a distinct triangle containing v and the other two vertices from B and C. This can be done similarly as above by an application of Lemma 8 since the degree conditions are satisfied. By repeating this procedure we eliminate all the strongly exceptional vertices, and then by Fact 17 we can finish the embedding. This finishes EC1.

2.4.2 Extremal Case 2

Here we have (2.1) and the following.

Extremal Case 2 (EC2) with parameter α: *There exists an $A \subset V(G)$ such that for $B = V(G) \setminus A$ we have*

- $\frac{n}{2} \geq |A| \geq \frac{n}{2} - \alpha n$, *and*
- $d(A, B) < \alpha$.

Thus roughly speaking, $G|_A$ and $G|_B$ are almost complete and the bipartite graph between A and B is sparse (note that k has to be small). By adding vertices to A we may achieve that $|A| = \lfloor n/2 \rfloor$ and $|B| = \lceil n/2 \rceil$. Furthermore, an easy computation shows that we still have $d(A, B) < 10\alpha$ (for simplicity we keep the notation A, B).

Again we define *exceptional* vertices $v \in A$ (and similarly for B), as

$$deg(v, B) \geq \sqrt{10\alpha}|B|.$$

Note that from the density condition $d(A, B) < 10\alpha$, the number of exceptional vertices in A is at most $\sqrt{10\alpha}|A|$ (and similarly for B). Let us denote the set of exceptional vertices by E_A in A and by E_B in B. Next again we further refine the definition of exceptional vertices: an exceptional vertex $v \in A$ (similarly in B) is called *strongly exceptional* if it is connected to few vertices in A, more precisely if we have

$$deg(v, A) \leq \alpha^{1/3}|A|.$$

Denote the set of strongly exceptional vertices by $SE_A(\subset E_A)$ in A and by $SE_B(\subset E_B)$ in B. If we have a $u \in SE_A$ and a $v \in SE_B$, then we can exchange the two vertices and they will not be strongly exceptional anymore in their new sets. Thus we may assume that one of the sets SE_A and SE_B is empty (say SE_B, the other case is similar). We first handle the vertices of SE_A.

We may assume that we have no $u \in B$ with $deg(u, A) > \alpha^{1/3}|A|$, since otherwise we can exchange this vertex with a vertex $v \in SE_A$ and thus reducing the size of SE_A. We remove the vertices in SE_A from A and add them to B, and denote the resulting sets by A' and B'. It is easy to see using (2.1) that in $G|_{A'}$ apart from at most $10\sqrt{\alpha}|A'|$ exceptional vertices all the degrees are at least $(1 - 10\sqrt{\alpha})|A'|$, and the degrees of the exceptional vertices are at least $\alpha^{1/3}|A'|/2$. In $G|_{B'}$ we have an even stronger degree condition; all the degrees are at least $(1 - 2\alpha^{1/3})|B'|$.

Suppose we have our cycles listed in increasing order of size, $C_1, C_2, \ldots C_r$. We assign cycles to A' until we have no room left and denote by C_m the last cycle, i.e. by adding this cycle we have at least $|A'|$ vertices assigned to $|A'|$, but without this cycle we have fewer than $|A'|$ assigned vertices. We refer to C_m as the *middle cycle*, let $n_m = |C_m|$, note that $n_m \leq 3$ only if we have many cycles of length 2 (i.e. *edges*) in the cycle system as we have few odd cycles and thus triangles. Denote

$$n_m^1 = |A'| - \sum_{i=1}^{m-1} |C_i| \text{ and } n_m^2 = n_m - n_m^1.$$

Note that part of C_m has to be embedded into A' (n_m^1 vertices) while the other part (n_m^2 vertices) into B'.

We may assume $n_m^2 > 0$ as well, since otherwise we are done. If $n_m^1, n_m^2 \geq 3$, then it is easy to embed the middle cycle C_m. Indeed we can find two bridge edges (u_i, v_i) with $u_i \in (A \setminus E_A)$, $v_i \in B$ for $i = 1, 2$ since from (2.1) we have $deg(u, B) \geq 1$ for each $u \in A$ and $deg(v, A) \leq \alpha^{1/3}|A|$ for each $v \in B$. Then we can connect u_1 and u_2 with a path of length $n_m^1 - 1$ in A and v_1 and v_2 with a path of length $n_m^2 - 1$ in B. Actually this argument also works for $n_m^1 = 2, n_m^2 \geq 3$ as well, since we can have two bridge edges where (u_1, u_2) is also an edge in $G|_A$. For $n_m^2 = 2$ note that if we can find two vertex disjoint paths of length 2 such that the center vertices are in B and the endpoints are in $A \setminus E_A$, then we move the center vertices to A' and now we have a perfect assignment. These length 2 paths can be used as part of some cycles and hence we are done. Thus we may assume that we have no two such paths. However, this fact and the degree conditions imply that we can find two bridges where (u_1, u_2) and (v_1, v_2) are *both* edges taking care of all cases $n_m^1, n_m^2 \geq 2$.

Finally let us assume $n_m^1 = 1, n_m^2 \geq 2$ (the other case is symmetric). If $k > 0$ or $SE_A \neq \emptyset$ we can clearly find a path of length 2 with its center vertex in A' and endpoints in B', and then we can move the center vertex to B' to have a perfect assignment. Thus we may assume that $k = 0$ (so n is even and $n_m^2 > 2$) and $SE_A = \emptyset$ and thus $|A| = |A'| = |B| = |B'|$ and we have no such path of length 2. Then $G|_A$ and $G|_B$ are both complete graphs with a perfect matching M between them and all of our cycles are even. Let $n_m = 2s, s \geq 2$, we will find a cycle $C_i : i < m, |C_i| = 2p, p < s$ or a cycle $C_j : j > m, |C_j| = 2q, q > s$. In case we have any such cycle (say C_i), we embed p vertices of C_i in A while the other p vertices in B, joining the corresponding end points using edges from M. Now in the remaining graph and cycle system we have $n_m^1 = p + 1 \geq 2$, while $n_m^2 = 2s - 1 - p \geq 2$, so C_m can be easily embedded using two edges from M. Note that we can always find either C_j or C_i, to see this, assume there are no such cycle, then all cycles are of length $2s$, hence $n \equiv 0 \pmod{2s}$ while $|A| = n/2 \equiv 1 \pmod{2s}$, which is a contradiction for $s > 1$.

It is easy to see that the other cycles apart from C_m (and possibly C_i or C_j in the last case) assigned to A' (and B') can be embedded in $G|_{A'}$ (and in $G|_{B'}$) by eliminating the few exceptional vertices first and then applying Fact 17. This finishes EC2.

Chapter 3

Perfect Matching in 3 uniform hypergraphs with large vertex degree

3.1 Introduction and Notation

For graphs we follow the notation in [6]. For a set T, we refer to all of its k-element subsets (k-sets for short) as $\binom{T}{k}$ and the number of such k-sets as $\binom{|T|}{k}$. We say that $H = (V(H), E(H))$ is an r-uniform hypergraph or r-graph for short, where $V(H)$ is the set of vertices and $E \subset \binom{V(H)}{r}$ is a family of r-sets of $V(H)$. We say that $H(V_1, \ldots, V_r)$ is an r-partite r-graph, if there is a partition of $V(H)$ into r sets, i.e. $V(H) = V_1 \cup \cdots \cup V_r$ and every edge of H uses exactly one vertex from each V_i. We call it a balanced r-partite graph if all V_i's are of the same size. Furthermore $H(V_1, \ldots, V_r)$ is a complete r-partite r-graph if every r-tuple that uses one vertex from each V_i belongs to $E(H)$. We denote a complete balanced r-partite r-graph by $K^{(r)}(t)$, where $t = |V_i|$. When the graph referred to is clear from the context we will use V instead of $V(H)$ and will identify H with $E(H)$. A matching in H is a set of disjoint edges of H and a perfect matching is a matching that contains all vertices. For $U \subset V$, $H|_U$ is the restriction of H to U.

For an r-graph H and a set $D = \{v_1, \ldots, v_d\} \in \binom{V}{d}, 1 \leq d \leq r$, the degree of D in H, $deg_H(D) = deg_r(D)$ denotes the number of edges of H that contains D. For $1 \leq d \leq r$, let

$$\delta_d = \delta_d(H) = \min \left\{ deg_r(D) \ : \ D \in \binom{V}{d} \right\}.$$

When H is an r-graph and A and B are disjoint subsets of $V(H)$, for a vertex $v \in A$ we denote by $deg_r(v, \binom{B}{r-1})$ the number of $(r-1)$-sets of B that make edges with v, while $d_r(v, \binom{B}{r-1}) = deg_r(v, \binom{B}{r-1})/\binom{|B|}{r-1}$ denotes the density. For such A and B, $e_r(A, \binom{B}{r-1})$

is the sum of $deg_r(v, \binom{B}{r-1})$ over all $v \in A$ while $d_r(A, \binom{B}{r-1}) = \frac{e_r(A,\binom{B}{r-1})}{|A|\binom{|B|}{r-1}}$. We denote by $H(A, \binom{B}{r-1})$ such a graph when all edges of H uses one vertex from A and $r-1$ vertices from B. When $A_1, \ldots, A_r$ are disjoint subsets of V, for a vertex $v \in A_1$ we denote by $deg_r(v, (A_2 \times \cdots \times A_r))$ the number of edges in the r-partite r-graph induced by subsets $\{v\}, A_2, \ldots, A_r$, and $e(A_1, (A_2 \times \cdots \times A_r))$ is the sum of $deg_r(v, (A_2 \times \cdots \times A_r))$ over all $v \in A_1$. Similarly

$$d_r(A_1, (A_2 \times \cdots \times A_r)) = \frac{e(A_1, (A_2 \times \cdots \times A_r))}{|A_1 \times A_2 \times \cdots \times A_r|}$$

An r-graph H on n vertices is η-*dense* if it has at least $\eta\binom{n}{r}$ edges. We use the notation $d_r(H) \geq \gamma$ to refer to an η-*dense* r-graph H. For $U \subset V$, for simplicity we refer to $d_r(H|_U)$ as $d_r(U)$ and to $E(H|_U)$ as $E(U)$. Throughout the thesis log denotes the base 2 logarithm. Moreover we will only deal with r-graphs on n vertices where $n = rk$ for some integer k, we denote this by $n \in r\mathbb{Z}$

Definition 1. *Let d, r and n be integers such that $1 \leq d < r$, and $n \in r\mathbb{Z}$. Denote by $m_d(r, n)$ the smallest integer m, such that every r-graph H on n vertices with $\delta_d(H) \geq m$ contains a perfect matching.*

For graphs ($r = 2$), by the Dirac theorem on Hamiltonicity of graphs, it s easy to see that $m_1(2, n) \leq n/2$, and since the complete bipartite $K_{n/2-1,n/2+1}$ does not have a perfect matching we get $m_1(2, n) = n/2$. For $r \geq 3$ and $d = r - 1$, it follows from a result of Rödl, Ruciński and Szemerédi on Hamiltonicity of r-graph [32] that $m_{r-1}(r, n) \leq n/2 + o(n)$. Kühn and Osthus [27] improved this result to $m_{r-1}(r, n) \leq n/2 + 3r^2\sqrt{n \log n}$. This bound was further sharpened in [31] to $m_{r-1}(r, n) \leq n/2 + C \log n$. In [33] the bound was improved to almost the true value, it was proved that $m_{r-1}(r, n) \leq n/2 + r/4$. Finally [34] settled the problem for $d = r - 1$.

The case $d < r - 1$ is rather hard, in [29] it is proved that for all $d \geq r/2$, $m_d(r, n)$ is close to $\frac{1}{2}\binom{n-d}{r-d}$. For $1 \leq d < r/2$ in [26] it was proved that

$$m_d(r, n) \leq \left(\frac{r-d}{r} + o(1)\right)\binom{n-d}{r-d}$$

A recent survey of these and other related results appear in [35]. In [26] the authors posed the following conjecture.

Conjecture 2 ([26]). *For all* $1 \leq d < r/2$,

$$m_d(r,n) \leq \max\left\{\frac{1}{2}, 1 - \left(\frac{r-1}{r}\right)^{r-d}\right\}\binom{n-d}{r-d}$$

Note that for $r = 3$ and $d = 1$ the above bound yields

$$m_1(3,n) \leq \frac{5}{9}\binom{n-1}{2}$$

The authors of [26] proved an approximate version of their conjecture for the case $r = 3$ and $d = 1$ they showed that $m_1(3,n) \leq \left(\frac{5}{9} + \eta\right)\binom{n}{2}$ for large n and some constant $\eta > 0$.

In this chapter we settle this conjecture for the case $r = 3$ and $d = 1$. Parallel to this work, independently Kühn, Osthus and Treglown [28] proved the same result. We believe our techniques are more general and have many other applications. In our subsequent work we use similar techniques to prove the conjecture for the case $r = 4$ and $d = 1$ as well. Our main result in this chapter is the following theorem.

Theorem 18. *There exist an integer* n_0 *such that if* H *is a* 3*-graph on* $n \geq n_0$ $(n \in 3\mathbb{Z})$ *vertices, and*

$$\delta_1(H) \geq \binom{n-1}{2} - \binom{2n/3}{2} + 1 \tag{3.1}$$

then H *has a perfect matching.*

On the other hand the following construction from [26] shows that the result is best possible.

Construction 1. *Let* $H = (V(H), E(H))$ *be a* 3*-graph on* n $(n \in 3\mathbb{Z})$ *vertices, such that* $V(H)$ *is partitioned into* A *and* B, $|A| = \frac{n}{3} - 1$ *and* $|B| = n - |A|$ *and* $E(H)$ *is the set of all* 3*-sets of* $V(H)$, T, *such that* $|T \cap A| \geq 1$ *(see Figure 4.1)*

We have $\delta_1(H) = \binom{n-1}{2} - \binom{2n/3}{2}$ (the degree of a vertex in B) but since every edge in a matching must use at least one vertex from A, the maximum matching in this graph is of size $|A| = \frac{n}{3} - 1$

3.1.1 Outline of the proof

We distinguish two cases to prove Theorem 25. When H is *far from* the extremal example as in Construction 2 we use the absorbing technique to find a perfect matching

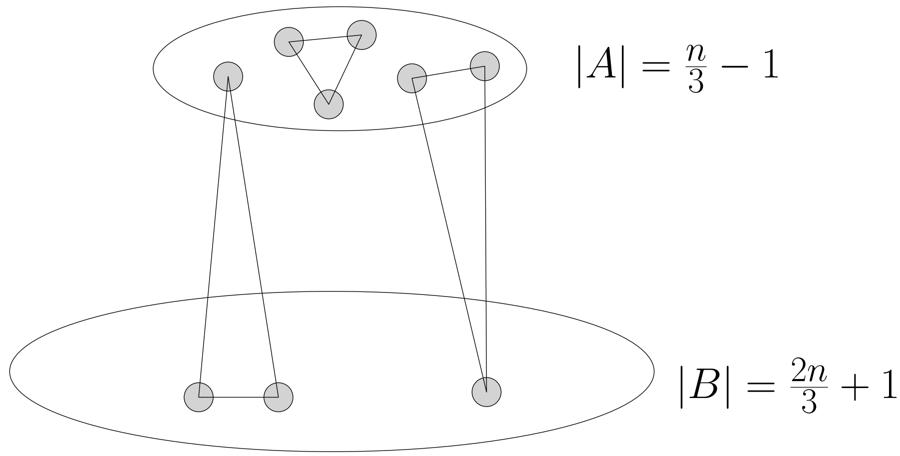


Figure 3.1: The extremal example: Every edge intersects the set A.

in H. On the other hand when H is *close to* to the extremal example we use a simple König-Hall type argument to find a perfect matching. We say that H is α-extremal for a constant $0 < \alpha < 1$, if the following is satisfied, otherwise it is non-extremal.

Definition 2 (Extremal Case with parameter α)**.** *There exists $B \subset V(H)$ such that*

- $|B| \geq (\frac{2}{3} - \alpha)n$
- $d_3(B) < \alpha$.

In the non-extremal case we use the absorbing lemma (Lemma 34) which roughly states that in H there exist a small matching M with the property that every "not too large" subset of vertices can be absorbed into this matching. In Section 4.4 using the tools developed in section 4.2 we find an *'almost perfect matching'* in $H|_{V \setminus V(M)}$. The left over vertices are absorbed into M to get a perfect matching in H. Finally in section 4.5 we build a perfect matching when our graph is α-extremal.

3.2 Tools

We use the following result from [25] which is a hypergraph extension of the Kővári-Sós-Turán theorem [17] to find complete balanced r-partite subhypergraphs of r-graphs.

Lemma 19. *For every integer $l \geq 1$ there is an integer $n_0 = n_0(r, l)$ such that: Every r-graph on $n > n_0$ vertices, that has at least $n^{r-1/l^{r-1}}$ edges, contains a $K^{(r)}(l)$.*

In particular Lemma 28 implies that for $\eta > 0$ and a sufficiently small constant β, if H is a 3-graph on $n > n_0(\eta, \beta)$ vertices with

$$|E(H)| \geq \eta \binom{n}{3}$$

then H contains a $K^{(3)}(t)$, where

$$t = \beta\sqrt{\log n}.$$

This is so because $\eta\binom{n}{3} \geq \frac{n^3}{n^{1/\beta^2 \log n}} = \frac{n^3}{2^{1/\beta^2}}$, when $\eta > \frac{6}{2^{1/\beta^2}}$.

We use the following two easy observations that will, nevertheless, be useful later on.

Fact 20. *If $G(A, B)$ is an η-dense bipartite graph, then there must be at least $\eta|B|/2$ vertices in B for which the degree in A is at least $\eta|A|/2$.*

Indeed, otherwise the total number of edges would be less than

$$\frac{\eta}{2}|A||B| + \frac{\eta}{2}|A||B| = \eta|A||B|$$

a contradiction to the fact that $G(A, B)$ is η-dense.

Fact 21. *If $G(A, B)$ is a bipartite graph, $|A| = c_1 \log n$, $|B| = c_2 n$ and for every vertex $b \in B$ deg$(b, A) \geq \eta|A|/2$, then we can find a complete bipartite subgraph $G'(A', B')$ of G such that $A' \subset A, B' \subset B, |A'| \geq \eta|A|/2$ and $|B'| \geq c_2 n^{(1-c_1)}$.*

To see this consider the neighborhoods in A, of the vertices in B. Since there can be at most $2^{|A|} = n^{c_1}$ such neighborhoods, by averaging there must be a neighborhood that appears for at least $\frac{c_2 n}{n^{c_1}} = c_2 n^{(1-c_1)}$ vertices of B. This means that we can find the desired complete bipartite graph.

The following two facts are repeatedly used in section 4.4.

Fact 22. *Let $H(X, Y, Z)$ be a 3-partite 3-graph with $|X| = |Y| = c_1 m$ and $|Z| = c_2 2^{m^2}$ for some constants $0 < c_1, c_2 < 1$. If $d_3(Z, (X \times Y)) \geq \eta$, then there exists a complete 3-partite 3-graph $H'(X', Y', Z')$ as a subgraph of H, such that $|X'| = |Y'| = |Z'| = c \log m$, where $c = c(\eta, c_1) > 0$.*

This can be seen by applying Fact 26 on H to get a subset Z_1 of Z such that for every vertex $z \in Z_1$, $deg_3(z, (X \times Y)) \geq \eta|X||Y|/2$ and $|Z_1| \geq \eta|Z|/2$. Now consider the auxiliary bipartite graph $G(A, Z_1)$, where $A = X \times Y$ and a vertex $z \in Z_1$ is connected to a pair $(a, b) \in A$ if (a, b, z) makes an edge of H. An application of Fact 27 on G gives a complete bipartite graph $G_2(A', Z_1')$ where $|A'| \geq \eta|X||Y|/2$ and $Z_1' \geq \frac{c_2\eta}{2} 2^{m^2(1-c_1^2)} > |X|$ when m is sufficiently large. Let G_3 be a bipartite graph where the color classes are X and Y and edges correspond to pairs in A'. Applying Lemma 28 on G_3 ($r = 2$), we get a complete bipartite graph in G_3 with color classes X' and Y', such that $|X'| = |Y'| \geq c \log m$. Clearly X', Y' and a subset of Z_1' of size $|X'|$, correspond to the color classes of required complete 3-partite 3-graph.

Fact 23. *Let $H = (A, \binom{B}{2})$ be a 3-graph such that $|A| = c_1 m$, $|B| = c_2 2^{m^2}$, for some constants $0 < c_1, c_2 < 1$. If $d_3(A, \binom{B}{2}) \geq \eta$ then there exists a complete 3-partite 3-graph $H'(A', B', B'')$, with $A' \subset A$, B' and B'' are disjoint subsets of B such that $|A'| = |B'| = |B''| = \eta|A|/2$.*

To see this first apply Fact 26 to get a subset of pairs of vertices in B, $P_1 \subset \binom{B}{2}$, such that every pair in P_1 makes edges with at least $\eta|A|/2$ vertices in A and $|P_1| \geq \eta\binom{|B|}{2}/2 \geq \eta|B|^2/5$. Next we find a $P_2 \subset P_1$, such that all pairs in P_2 make edges with the same subset of A (say $A' \subset A$). By Fact 27 we have

$$|P_2| \geq \frac{\eta|B|^2/5}{2^{c_1 m}} = \frac{\eta}{5}\frac{(c_2 2^{m^2})^2}{2^{c_1 m}} = \frac{c_2^2\eta}{5} 2^{m^2(2-c_1/m)} = \frac{c_2^2\eta}{5}\left(\frac{|B|}{c_2}\right)^{2-c_1/m} \geq |B|^{2-2/\eta|A|}$$

where the last inequality follows when m is sufficiently large and η, c_1 and c_2 are small constants. Now construct an auxiliary graph G_1 where $V(G_1) = B$ and edges of G corresponds to pairs in P_2. Applying lemma 28 on G_1 (for $r = 2$) we get a complete bipartite graph with color classes B' and B'' each of size $\eta|A|/2$. Clearly A', B', and B'' corresponds to color classes of a complete 3-partite 3-graph in H as in the statement of the fact.

Finally we use the following lemma from [26].

Lemma 24. *(Absorbing Lemma) For every $\eta > 0$, there is an integer $n_0 = n_0(\eta)$ such that if H is a 3-graph on $n \geq n_0$ vertices with $\delta_1(H) \geq (1/2 + 2\eta)\binom{n}{2}$, then there exist* *a matching M in H of size $|M| \le \eta^3 n$ such that for every set $W \subset V \setminus V(M)$ of size at size at most $\eta^6 n \ge |W| \in 3\mathbb{Z}$, there exists a matching covering exactly the vertices in $V(M) \cup W$.*

3.3 The Non Extremal Case

Throughout this section we assume that we have a 3-graph H satisfying (4.1) such that the extremal case does not hold for H. We shall assume that n is sufficiently large and besides our main parameter γ we use the parameters β and α such that following holds

$$\frac{6}{2^{1/\beta^2}} < \gamma = \alpha^3 \ll 1 \tag{3.2}$$

where $a \ll b$ means that a is sufficiently small compared to b. From (4.1) and (4.2), when n is large we have

$$\delta_1(H) \ge \binom{n-1}{2} - \binom{2n/3}{2} + 1 > \frac{5}{9}\binom{n-1}{2} - \frac{n}{3} > (1/2 + 2\sqrt{\alpha})\binom{n}{2}.$$

Hence our hypergraph H satisfies the conditions of Lemma 34 (the absorbing lemma). We remove from H an *absorbing matching* M of size at most $\alpha^{3/2} n = \sqrt{\gamma} n$. In the remaining hypergraph we find an almost perfect matching that leaves out a set of at most $\alpha^3 n = \gamma n$ vertices. As guaranteed by Lemma 34 the vertices that are left out from the almost perfect matching can be absorbed into M, therefore we get a perfect matching in H. In what follows we work with the remaining hypergraph (after removing $V(M)$). For simplicity we still denote the remaining hypergraph by H and assume that it is on n vertices. Since $\binom{|V(M)|}{2} \le \binom{3\sqrt{\gamma}n}{2} < 5\gamma n^2$ it is easy to see that in the remaining hypergraph we still have

$$\delta_1(H) \ge \left(\frac{5}{9} - 10\gamma\right)\binom{n}{2} \tag{3.3}$$

3.3.1 The optimal cover

Our goal is to find an almost perfect matching in H. In fact we are going to build a cover $\mathcal{T} = \{T_1, T_2, \ldots\}$ where each T_i is a disjoint balanced complete 3-partite 3-graphs in H (we refer to them as tripartite graphs). We say that such a cover is optimal if it covers at least $(1-\gamma)n$ vertices. We will show that either we can find an optimal cover or H is α-extremal. It is easy to see that such an optimal cover readily gives us an almost perfect matching.

We begin with a cover $\mathcal{T}$ obtained by repeatedly applying Lemma 28 in the leftover of H as long as there are at least γn vertices left and the condition of Lemma 28 is satisfied, to get disjoint $K^{(3)}(t)$'s where $t = \beta\sqrt{\log n}$. When the condition of Lemma 28 is no more satisfied then using the procedure outlined below we will increase the size of our cover. We will show that we can build an optimal cover unless H is α-extremal. We identify by $\mathcal{T}$ the set of tripartite graphs in the cover $\mathcal{T}$, while $V(\mathcal{T})$ denotes the union of vertices in tripartite graphs in $\mathcal{T}$ (the size of the cover). We refer to a subset of tripartite graphs in $\mathcal{T}$ as a subcover in $\mathcal{T}$. Let $\mathcal{I} = V(H) \setminus V(\mathcal{T})$ be the set of remaining vertices. Our goal is to show if H is non-extremal then either $|\mathcal{I}| < \gamma n$ (meaning the cover is optimal) or we can increase the size of our cover by at least $\gamma^2 n/8$ vertices. Assume that $|\mathcal{I}| \geq \gamma n$ then we must have

$$d_3(\mathcal{I}) < \gamma \tag{3.4}$$

Indeed otherwise by Lemma 28 we can find disjoint complete tripartite graphs in $H|_{\mathcal{I}}$ that cover at least $\gamma|\mathcal{I}|/2 > \gamma^2 n/8$ vertices, and adding these tripartite graphs to $\mathcal{T}$ increases the size of our cover by at least $\gamma^2 n/8$.

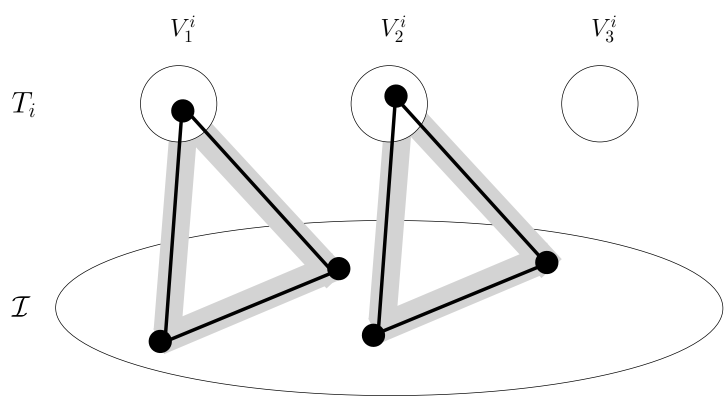


Figure 3.2: T_i is at least 2-sided to $\mathcal{I}$

Let $T_i = (V_1^i, V_2^i, V_3^i)$ be a tripartite graph in $\mathcal{T}$ we say that a color class, V_l^i of T_i, $(1 \leq l \leq 3)$, is connected to $\mathcal{I}$, if $d_3(V_l^i, \binom{\mathcal{I}}{2}) \geq \gamma$. For $1 \leq k \leq 3$, we say that T_i is k-sided if k color classes of T_i are connected to $\mathcal{I}$. We will show that most of the tripartite graphs in $\mathcal{T}$ are at most 1-sided to $\mathcal{I}$ or we can increase the size of our cover by at least $\gamma^2 n/8$ vertices. To see this assume that we have a subcover $\mathcal{T}' \subset \mathcal{T}$, such that each $T_i \in \mathcal{T}'$ is at least 2-sided to $\mathcal{I}$ and $|V(\mathcal{T}')| \geq \gamma n$. Let $\mathcal{T}' = \{T_1, T_2, \ldots\} \subset \mathcal{T}$ and without loss of generality, say V_1^i and V_2^i are the color classes in each T_i that are connected to $\mathcal{I}$. For each such T_i, since $|V_1^i| = |V_2^i| \leq \beta\sqrt{\log n} = \beta m$ and $|\mathcal{I}| \geq \gamma n \geq \gamma 2^{m^2}$ from Fact 29 we can find two disjoint balanced complete tripartite graphs (U_1^i, A_1^i, B_1^i) and (U_2^i, A_2^i, B_2^i) where U_1^i and U_2^i are subsets of V_1^i and V_2^i respectively, and A_1^i, A_2^i, B_1^i and B_2^i are disjoint subsets of $\mathcal{I}$. The size of each color class of these tripartite graphs is $\gamma|V_1^i|/8$ (see Figure 3.2) We remove the vertices of these new tripartite graphs from their respective sets and add the tripartite graphs to our cover. Removing these vertices from V_1^i and V_2^i creates an imbalance in the leftover part of T_i (V_3^i has more vertices). To restore the balance in the leftover of T_i we discard (add to $\mathcal{I}$) some arbitrary $|U_1^i| = |U_2^i| = \gamma|V_1^i|/8$ vertices from V_3^i. The new tripartite graphs use at least $2|A_1^i|+2|B_1^i| = \gamma|V_1^i|/2$ vertices from $\mathcal{I}$ even after discarding the vertices from V_3^i the net increase in the size of our cover is $3\gamma|V_1^i|/8$, while all the tripartite graphs in $\mathcal{T}$ are balanced. Repeating this procedure for all tripartite graphs in $\mathcal{T}'$ we will increase the size of our cover by at least $3\gamma^2 n/8$ vertices, because $|V(\mathcal{T}')| \geq \gamma n$. Therefore, we assume that the number of vertices in the at least 2-sided tripartite graphs is at most γn. Note that this implies that for a typical vertex $v \in \mathcal{I}$ we have $deg_3(v, \mathcal{I} \times V(\mathcal{T})) \leq (1/3 + 3\gamma)|\mathcal{I}||V(\mathcal{T})|$.

From (4.3), (4.4) and the fact that almost all tripartite graphs are at most 1-sided, we get that for a typical vertex $v \in \mathcal{I}$ we have

$$\begin{aligned} deg_3\left(v, \binom{V(\mathcal{T})}{2}\right) &\geq \left(\frac{5}{9} - 10\gamma\right)\binom{n}{2} - deg_3(v, \mathcal{I} \times V(\mathcal{T})) - deg_3\left(v, \binom{\mathcal{I}}{2}\right) \\ &\geq \left(\frac{5}{9} - 10\gamma\right)\binom{n}{2} - \left(\frac{1}{3} + 3\gamma\right)|\mathcal{I}||V(\mathcal{T})| - \gamma\binom{|\mathcal{I}|}{2} \\ &\geq \left(\frac{5}{9} - 38\gamma\right)\binom{|V(\mathcal{T})|}{2} \end{aligned}$$

where the last inequality holds when $|\mathcal{I}| \geq \gamma n$ and γ is a small constant. Similar calculation using (4.3), (4.4) and the fact that almost all tripartite graphs are at most 1-sided yields that $|V(\mathcal{T})| \geq 2n/3$.

For a vertex v, consider the edges that v makes with pairs of vertices within the tripartite graphs. Since the number of pairs of vertices of any tripartite graph $T_i \in \mathcal{T}$ is $O(\log n)$, the total number of pairs of vertices within the tripartite graphs in $\mathcal{T}$ is $O(n \log n) = o\binom{n}{2}$. Therefore we ignore these edges and for any vertex v we will only consider the edges that v makes with pairs of vertices $(x, y),\ x \in V(T_i),\ y \in V(T_j),\ i \neq j$. By the above observation, for the minimum degree of a typical vertex $v \in \mathcal{I}$ we still have

$$deg_3\left(v, \binom{V(\mathcal{T})}{2}\right) \geq \left(\frac{5}{9} - 40\gamma\right)\binom{|V(\mathcal{T})|}{2} \tag{3.5}$$

Let $T_i = (V_1^i, V_2^i, V_3^i)$ and $T_j = (V_1^j, V_2^j, V_3^j)$ be two tripartite graphs in $\mathcal{T}$, we say that $\mathcal{I}$ is *connected* to a pair of color classes (V_p^i, V_q^j), $1 \leq p, q \leq 3$, if $d_3(\mathcal{I}, (V_p^i \times V_q^j)) \geq \gamma$. For $k \in \{1, \ldots, 9\}$ we say $\mathcal{I}$ is k-*connected* to a pair of tripartite graphs $(T_i, T_j) \in \binom{\mathcal{T}}{2}$ if $\mathcal{I}$ is *connected* to k pairs of color classes (V_p^i, V_q^j), $1 \leq p, q \leq 3$. Denote by $s(\mathcal{I}, (T_i, T_j))$ the largest value of k for which $\mathcal{I}$ is k-*connected* to (T_i, T_j). For a constant $\eta > 0$ we say that $\mathcal{I}$ is (η, k)-*connected* to $\mathcal{T}$ if there exists a subcover $\mathcal{T}' \subset \mathcal{T}$ such that

- $|V(\mathcal{T}')| \geq \eta n$ and let $\mathcal{T}' = \{T_1', T_2', \ldots\}$
- for each $T_i \in \mathcal{T}'$ we have another subcover $\hat{\mathcal{T}_i} \subset \mathcal{T}$, (not necessarily disjoint from $\mathcal{T}'$ and other $\hat{\mathcal{T}_p}$'s) such that if $\hat{\mathcal{T}_i} = \{T_{i_1}, T_{i_2}, \ldots\}$ then for each $T_{i_j} \in \hat{\mathcal{T}_i}$ we have $s(\mathcal{I}, (T_i, T_{i_j})) = k$ and $|V(\hat{\mathcal{T}_i})| \geq \eta n$.

Note that if $\mathcal{I}$ is (η, k)-*connected* to $\mathcal{T}$ then by a simple greedy procedure we can find a set of disjoint pairs of tripartite graphs, $\mathcal{P} \subset \binom{\mathcal{T}}{2}$, such that for the each pair $(T_i^p, T_j^p) \in \mathcal{P}$, $(1 \leq p \leq |\mathcal{P}|)$, we have $s(\mathcal{I}, (T_i^p, T_j^p)) = k$ and if $m_p = \min\{|V(T_i^p)|, |V(T_j^p)|\}$ then $\sum_p m_p \geq \eta n/2$. We refer to the value $\sum_p m_p$ as the size of $\mathcal{P}$. Furthermore on average for a pair of tripartite graphs (T_i, T_j) we have $s(\mathcal{I}, (T_i, T_j)) \geq 5$, because by (4.5) on average for $(T_i, T_j) \in \binom{\mathcal{T}}{2}$ we have $d_3(\mathcal{I}, (V(T_i) \times V(T_j))) \geq (5/9 - \sqrt{\gamma}) > (4/9 + 5\gamma)$. This implies that if $\mathcal{I}$ is $(\sqrt{\gamma}, \leq 4)$-*connected* to $\mathcal{T}$, then we also have that $\mathcal{I}$ is $(2\gamma, \geq 6)$-*connected* to $\mathcal{T}$.

For $T_i = (V_1^i, V_2^i, V_3^i)$ and $T_j = (V_1^j, V_2^j, V_3^j)$, define v_1^{ij} (respectively v_1^{ji}) to be the number of color classes V_q^j's in T_j (respectively V_p^i's in T_i) such that $\mathcal{I}$ is *connected* to (V_1^i, V_q^j) (respectively (V_1^j, V_q^i)). Similarly we define v_2^{ij} and v_3^{ij} (v_2^{ji} and v_3^{ji}). For a fixed pair (T_i, T_j) we assume that $v_1^{ij} \geq v_2^{ij} \geq v_3^{ij}$ ($v_1^{ji} \geq v_2^{ji} \geq v_3^{ji}$). Note that when $s(\mathcal{I}, (T_i, T_j)) \geq 6$ then we must have $v_1^{ij} \geq v_2^{ij} \geq 2$. On the other hand when $s(\mathcal{I}, (T_i, T_j)) = 5$ then we could either have $v_1^{ij} \geq v_2^{ij} \geq 2$ or $v_1^{ij} = 3$ and $v_2^{ij} = v_2^{ij} = 1$. We will consider the following cases based on the way $\mathcal{I}$ is connected to $\mathcal{T}$ and show that either we can increase the size of our cover by at least $\gamma^2 n/8$ vertices or H is extremal.

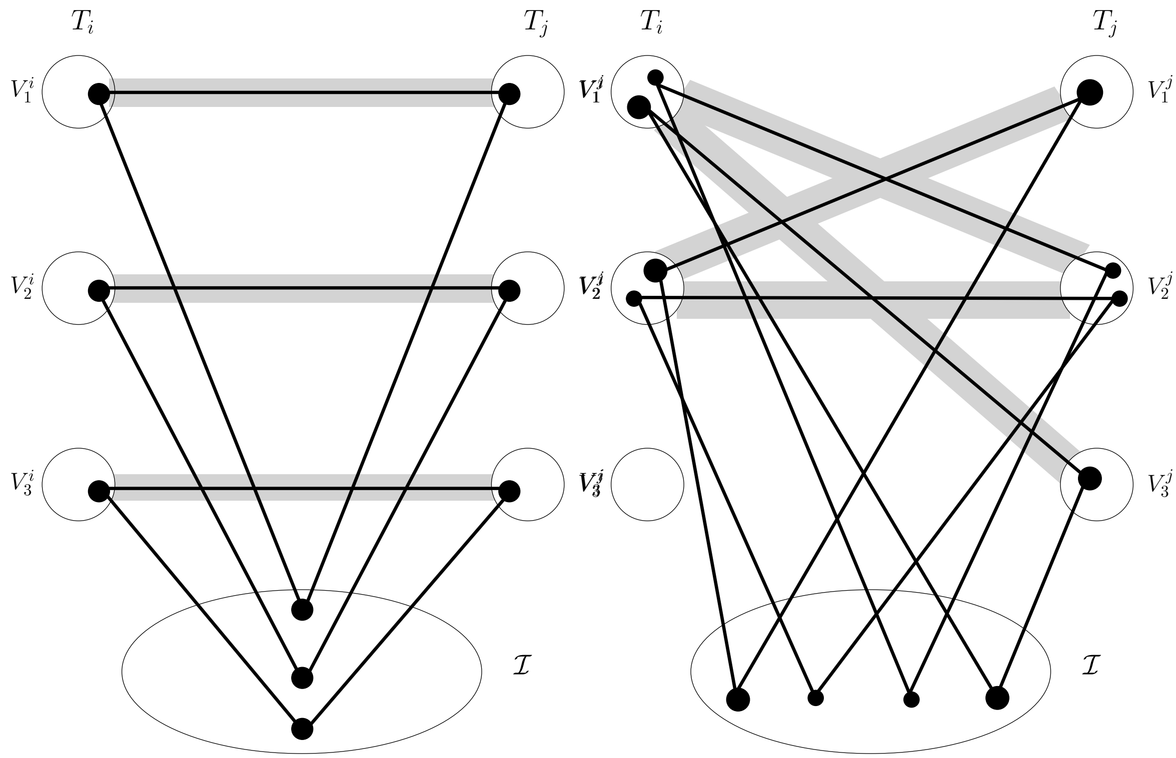
Case 1: There exists a set of disjoint pairs of tripartite graphs $\mathcal{P} \subset \binom{\mathcal{T}}{2}$ of size at least γn and for each pair $(T_i, T_j) \in \mathcal{P}$ we have $s(\mathcal{I}, (T_i, T_j)) \geq 5$ with $v_1^{ij}, v_2^{ij} \geq 2$.

In this case we increase the size of our cover as follows. For each pair $(T_i, T_j) \in \mathcal{P}$, if both $v_3^{ij} \geq 1$ and $v_3^{ji} \geq 1$ then it is easy to see that we can match each color class of T_i with a distinct color class of T_j such that the matched pairs are *connected* to $\mathcal{I}$. Without loss of generality assume that $\mathcal{I}$ is *connected* to (V_1^i, V_1^j), (V_2^i, V_2^j) and (V_3^i, V_3^j). Note that by the definition of *connectedness* the 3-partite subhypergraph of H induced by $(V_1^i, V_1^j, \mathcal{I})$ satisfies the conditions of Fact 22. Hence by Fact 22 we can find a complete balanced tripartite graph $T_1 = (U_1^i, U_1^j, \mathcal{I}_1)$, such that

$$\mathcal{I}_1 \subset \mathcal{I}, \;\; U_1^i \subset V_1^i \;, U_1^j \subset V_1^j \quad \text{and } |\mathcal{I}_1| = |U_1^i| = |U_1^j| = c \log m$$

where $m = \min\{|V_1^i|, |V_1^j|\}$ and c is as in Fact 22. Similarly, we can find such complete balanced tripartite graphs T_2 and T_3 in $(V_2^i, V_2^j, \mathcal{I})$ and $(V_3^i, V_3^j, \mathcal{I})$ respectively, that are disjoint from each other since $|\mathcal{I}| \geq \gamma n$ (see Figure 3.3(a)). We remove the vertices in T_1, T_2 and T_3 from their respective sets and add these three new tripartite graphs to our cover. In the remaining part of T_i and T_j we remove another such set of 3 disjoint tripartite graphs. Again by definition of *connectedness* and Fact 22 we can continue this process until we remove at least $\gamma m/2$ vertices from each color class of T_i and T_j. Note the new tripartite graphs use $3\gamma m/2$ vertices from $\mathcal{I}$. Therefore adding these new tripartite graphs to our cover increases the size of the cover by $3\gamma m/2$ vertices while keeping the cover balanced.

On the other hand, if for a pair $(T_i, T_j) \in \mathcal{P}$ either v_3^{ij} or $v_3^{ji} = 0$ (both cannot be 0 because otherwise $s(\mathcal{I}, (T_i, T_j)) \leq 4$) then we deal with this pair as follows. Say



(a) There is a perfect matching (b) There is a pseudo-matching

Figure 3.3: Shaded rectangles represent pairs connected to $\mathcal{I}$: Triangles represent complete tripartite graphs.

$v_3^{ij} = 0$, since $s(\mathcal{I}, (T_i, T_j)) \geq 5$ we must have $v_1^{ij} = 3$, $v_2^{ij} \geq 2$, $v_1^{ji}, v_2^{ji} \geq 2$ and $v_3^{ji} \geq 1$. Since $v_2^{ij} \geq 2$, assume without loss of generality that $\mathcal{I}$ is *connected* to (V_2^i, V_1^j) and (V_2^i, V_2^j), then as above applying Fact 22 we find the following four disjoint complete tripartite graphs: $(V_{11}^i, V_{31}^j, \mathcal{I}_1)$, $(V_{22}^i, V_{12}^j, \mathcal{I}_2)$, $(V_{13}^i, V_{23}^j, \mathcal{I}_3)$ and $(V_{24}^i, V_{24}^j, \mathcal{I}_4)$ such that for $1 \leq p \leq 3$ and $1 \leq q \leq 4$, $\mathcal{I}_q, V_{pq}^i$ and V_{pq}^j are disjoint subsets of $\mathcal{I}, V_p^i$ and V_p^j respectively. For the sizes of these new tripartite graphs we have

$$|\mathcal{I}_1| = |\mathcal{I}_2| = |V_{11}^i| = |V_{31}^j| = |V_{22}^i| = |V_{12}^j| = 2c \log m/3$$

and

$$|\mathcal{I}_3| = |\mathcal{I}_4| = |V_{13}^i| = |V_{23}^j| = |V_{24}^i| = |V_{24}^j| = c \log m/3$$

where $m = \min\{|V_1^i|, |V_1^j|\}$ and c is as in Fact 22 (see Figure 3.3(b)). In the remaining part of T_i and T_j we remove another such set of 4 disjoint tripartite graphs. Again by definition of *connectedness* and Fact 22 we can continue this process until we remove at least $\gamma m/2$ vertices from each V_1^i and V_2^i. Note that when we remove the vertices of these new tripartite graphs from their respective color classes in T_i and T_j, the remaining part of T_j is still balanced while it creates an imbalance in the remaining part of T_i, as V_3^i is larger than then the other two color classes. To restore the balance we discard (add to $\mathcal{I}$) an arbitrary subset of vertices in V_3^i (of size equal to the imbalance). Still in this process the net increase in the number of vertices in the cover is $\gamma m/2$.

We proceed in similar manner for all pairs in $\mathcal{P}$ and increase the size of the cover by at least $\gamma^2 n/2$ vertices (as the size of $\mathcal{P}$ is at least γn) while keeping all the tripartite graphs in the cover balanced. Note that since n is sufficiently large the size of each tripartite graph is still large enough.

Case 2: There is no such $\mathcal{P}$

In this case we must have that $\mathcal{I}$ is not $(\sqrt{\gamma}, 4)$-*connected* to $\mathcal{T}$ because otherwise, as noted above, there will be a set of disjoint pairs of tripartite graphs $\mathcal{P} \subset \binom{\mathcal{T}}{2}$ of size at least γn, such that for each pair $(T_i, T_j) \in \mathcal{P}$ we have $s(\mathcal{I}, (T_i, T_j)) \geq 6$ and such a $\mathcal{P}$ satisfies the conditions of case 1.

Therefore we must have that for 'almost all' pairs $(T_i, T_j) \in \binom{\mathcal{T}}{2}$ we have $s(\mathcal{I}, (T_i, T_j)) = 5$ and $v_1^{ij} = v_1^{ji} = 3, v_2^{ij} = v_2^{ji} = v_3^{ij} = v_3^{ji} = 1$. Call a tripartite graph $T_i \in \mathcal{T}$ good if for almost all other tripartite graphs T_j (covering $\geq (1 - 20\sqrt{\gamma})|V(\mathcal{T})|$ vertices), we have $s(\mathcal{I}, (T_i, T_j)) = 5$ such that $v_1^{ij} = v_1^{ji} = 3, v_2^{ij} = v_2^{ji} = v_3^{ij} = v_3^{ji} = 1$. Note that by the above observation and since the condition of Case 1 is not satisfied almost all tripartite graphs (covering $\geq (1 - 20\sqrt{\gamma})|V(\mathcal{T})|$ vertices) are good. Now by a simple greedy procedure we match every good tripartite graph T_i with another distinct tripartite graph T_j such that they have the above connectivity structure. Let the set of these disjoint pairs be $\mathcal{P}_g$ and the set of tripartite graphs in $\mathcal{P}_g$ be $\mathcal{T}_g$.

For every $(T_i, T_j) \in \mathcal{P}_g$ applying Fact 22 we find the following four disjoint complete tripartite graphs: $(V_{11}^i, V_{21}^j, \mathcal{I}_1)$, $(V_{22}^i, V_{12}^j, \mathcal{I}_2)$, $(V_{13}^i, V_{33}^j, \mathcal{I}_3)$, and $(V_{24}^i, V_{34}^j, \mathcal{I}_4)$ such that for $1 \leq p \leq 3$ and $1 \leq q \leq 4$, $\mathcal{I}_p$, V_{pq}^i and V_{pq}^j are disjoint subsets of $\mathcal{I}, V_p^i$ and V_p^j respectively. The size of each color class in these tripartite graphs is $c \log m$ where $m = \min\{|V_1^i|, |V_1^j|\}$ and c is as in Fact 22. As argued above we repeat this process until in total we use $\gamma^2 m$ vertices each from V_1^i and V_1^j while $\gamma^2 m/2$ vertices from each of the other color classes. We do this such that we use up as close as possible to $\gamma^2 m/2$ vertices each in V_2^i and V_3^i.

Note that in the new tripartite graphs in total we used $2\gamma^2 m$ vertices from $\mathcal{I}$. Removing these new tripartite graphs creates an imbalance among the color classes of the remaining part of T_i and T_j, to restore the balance we will have to discard $\gamma^2 m/2$ vertices from each color class of T_i and T_j except V_1^i and V_1^j. Which leaves us with no net gain in the size of the cover. Therefore we will not discard any vertices from these color classes at this time and say that these color classes have extra vertices. We proceed in similar manner for each pair in $\mathcal{P}_g$.

Let V_1^g, V_2^g, V_3^g be the union of the corresponding color classes of remaining parts of tripartite graphs in $\mathcal{T}_g$. By the above observation on the good tripartite graphs and since the number of vertices used from V_2^g and V_3^g in the newly made tripartite graphs above is at most $\gamma^2 n$ we have $|V_2^g| = |V_3^g| \geq (1 - 20\sqrt{\gamma})|V(\mathcal{T})|/3 - \gamma^2 n$. We will show that either we can increase the size of our cover or we have

$$d_3(V_2^g \cup V_3^g) \leq \sqrt{\gamma}. \tag{3.6}$$

For $d_3(V_2^g \cup V_3^g)$, we consider only those edges that use exactly one vertex from a tripartite graph T_i, as the number of edges of other types is at most $o(n^3)$. Assume that $d_3(V_2^g \cup V_3^g) \geq \sqrt{\gamma}$ then by Lemma 28 there exist balanced complete tripartite 3-graphs in $H|_{V_2^g \cup V_3^g}$ covering at least γn vertices. We remove some of these new tripartite graphs (possibly with splitting and discarding part of them) so that from no color class we remove more then the number of extra vertices in that color class. Now adding these new tripartite graphs to our cover, significantly increases the size of our cover. As we will not need to discard vertices from $V_2^g \cup V_3^g$ for rebalancing. Instead the extra vertices are part of these new tripartite graphs. In the remaining parts of V_2^g and V_3^g we arbitrarily remove some extra vertices to restore the balance in the tripartite graphs.

Similarly we will show that either we can increase the size of our cover or we have

$$d_3(V_2^g \cup V_3^g, \binom{\mathcal{I}}{2}) \leq \sqrt{\gamma}. \tag{3.7}$$

Indeed assume the contrary, i.e. $d_3(V_2^g \cup V_3^g, \binom{\mathcal{I}}{2}) \geq \sqrt{\gamma}$, then since both $|\mathcal{I}|$ and $|V_2^g \cup V_3^g|$ are at least γn, by Lemma 28 we can find disjoint complete tripartite graphs with one color class in $V_2^g \cup V_3^g$ and two color classes in $\mathcal{I}$ covering at least $\gamma^2 n/2$ vertices. And again as above we can add these tripartite graphs and increase the size of our cover as we have extra vertices in $V_2^g \cup V_3^g$.

From the above observations about the size of V_2^g and V_3^g and (4.2) we have that $|V_2^g \cup V_3^g \cup \mathcal{I}| \geq (2/3 - \alpha)n$. Therefore if we can not increase the size of our cover significantly (by at least $\gamma^n/8$ vertices), then by (4.4), (3.6) and (3.7) we get that $d_3(V_2^g \cup V_3^g \cup \mathcal{I}) < 10\sqrt{\gamma} < \alpha$. Hence H is α-extremal.

3.4 The Extremal Case

Here our graph H is in the Extremal Case, i.e. *there exists a $B \subset V(H)$ such that*

- $|B| \geq (\frac{2}{3} - \alpha)n$
- $d_3(B) < \alpha$.

We assume that n is sufficiently large and α is a sufficiently small constant < 1. Let $A = V(H) \setminus B$, by shifting some vertices between A and B we can have that $A = n/3$ and $B = 2n/3$ as $n \in 3\mathbb{Z}$ (we still keep the notation A and B). It is easy to see that we still have

$$d_3(B) < 6\alpha \tag{3.8}$$

Since we have

$$\delta_1(H) \geq \binom{n-1}{2} - \binom{2n/3}{2} + 1 = \binom{n-1}{2} - \binom{|B|}{2} + 1$$

together with (4.8) this implies that almost all 3-sets of $V(H)$ are edges of H except 3-sets of B. Thus roughly speaking we have that almost every vertex $b \in B$ makes edges with almost all pairs of vertices in $\binom{A}{2}$ and with almost all pairs of vertices in $B \setminus \{b\} \times A$

and vice versa. Therefore, we will basically match every vertex in A with a distinct pair of vertices in $\binom{B}{2}$ to get the perfect matching. However some vertices may be '*atypical*', in the sense that they may not have this connectivity structure hence we will first find a small matching that covers all such '*atypical*' vertices. For the remaining '*typical*' vertices we will show that they satisfy the conditions of König-Hall theorem, hence we will match every remaining vertex in A with a distinct pair of remaining vertices in B.

A vertex $a \in A$ is called *exceptional* if it does not make edges with almost all pairs of vertices in B, more precisely if

$$deg_3\left(a, \binom{B}{2}\right) < (1 - \sqrt{\alpha})\binom{|B|}{2}$$

A vertex $a \in A$ is called *strongly exceptional* if it makes edges with very few pairs in B, more precisely if

$$deg_3\left(a, \binom{B}{2}\right) < \alpha^{1/3}\binom{|B|}{2}$$

Similarly a vertex $b \in B$ is called *exceptional* if it does not make edges with almost all pairs of vertices $(b_i, a_j) : b \neq b_i \in B$ and $a_j \in A$, more precisely if

$$deg_3(b, (B \setminus \{b\} \times A)) < (1 - \sqrt{\alpha})|A|(|B| - 1)$$

A vertex $b \in B$ is called *strongly exceptional* if it makes edges with very few pairs of vertices $(b_i, a_j) : b \neq b_i \in B$ and $a_j \in A$, more precisely if

$$deg_3(b, (B \setminus \{b\} \times A)) < \alpha^{1/3}|A|(|B| - 1)$$

Denote the set of *exceptional* and *strongly exceptional* vertices in A (and B) by X_A and SX_A respectively (similarly X_B and SX_B). From (4.1) and (4.8) there are few vertices in X_A (and X_B) and very few vertices in SX_A (and SX_B). More Precisely we have that $|X_A| \leq 18\sqrt{\alpha}|A|$ and $|X_B| \leq 18\sqrt{\alpha}|B|$ and for the *strongly exceptional* sets we have $|SX_A| \leq 40\alpha|A|$ and $|SX_B| \leq 40\alpha|B|$. The constants are not the best possible but we choose them for ease of calculation. We will only show the bound on $|X_B|$ similar computation yields the bounds on $|X_A|$, $|SX_A|$ and $|SX_B|$. Assume that $|X_B| \geq 18\sqrt{\alpha}|B|$. Note that by (4.1) and the definition of X_B, for a vertex $b \in X_B$, we have $deg_3\left(b, \binom{B}{2}\right) \geq \sqrt{\alpha}|A|(|B|-1)/2$. Therefore for the number of edges inside B we have

$$3|E(B)| \geq |X_B| \cdot \sqrt{\alpha}|A|(|B|-1)/2 \geq 9\sqrt{\alpha}|B| \cdot \sqrt{\alpha}|A|(|B|-1) \geq 9\alpha|B|(|B|-1)|A| \geq 27\alpha\binom{|B|}{3}$$

where the last inequality uses $|A| = |B|/2$ and since an edge can be counted at most 3 times we use $3|E(B)|$. Note that this implies that $d_3(B) > 9\alpha$ a contradiction to (4.8).

If we have both SX_B and SX_A non empty, (say $b \in SX_B$ and $a \in SX_A$) then since by definition $deg_3(a, \binom{B}{2}) < \alpha^{1/3}\binom{|B|}{2}$, from (4.1) we must have $deg_3(a, A \setminus \{a\} \times B) \geq (1-2\alpha^{1/3})(|A|-1)|B|$, (similar bound holds for b). Therefore we can exchange a with b and reduce the size of both SX_B and SX_A, as both a and b are not *strongly exceptional* in their new sets. Hence one of the sets SX_A and SX_B must be empty.

Assume $SX_B \neq \emptyset$. As observed above by the minimum degree condition and definition of SX_B, for every vertex $b \in SX_B$, we have $deg_3(b, \binom{B}{2}) \geq (1-2\alpha^{1/3})\binom{|B|-1}{2}$. This together with the bound on the size of SX_B implies that we can greedily find $|SX_B|$ vertex disjoint edges in $H|_B$ each containing exactly one vertex of SX_B. We also select $|SX_B|$ other vertex disjoint edges such that each edge has a vertex in $B \setminus X_B$ and the two other vertices are in A. We can clearly find such edges because by (4.1) and (4.8) almost every vertex in $B \setminus X_B$ makes edges with at least $(1-2\sqrt{\alpha})\binom{|A|}{2}$ pairs of vertices in A (as otherwise $d_3(B)$ will be very large). We remove the vertices of these edges from A and B and denote the remaining set by A' and B'. Let $|A'| + |B'| = n'$, by the above procedure we have $n' = n - 6|SX_B|$, $|A'| = |A| - 2|SX_B|$ and $|B'| = |B| - 4|SX_B|$ hence we get $|B'| = 2|A'| = 2n'/3$.

In case $SX_A \neq \emptyset$ (and $SX_B = \emptyset$), we will first eliminate the vertices in SX_A. Note that in this case any vertex $b \in B$ is exchangeable with any vertex in SX_A, because if there is a vertex $b \in B$ such that $deg_3(b, \binom{B}{2}) \geq \alpha^{1/3}\binom{|B|}{2}$ then we can replace b with any vertex $a \in SX_A$ to reduce the size of SX_A (as the vertex b is not *strongly exceptional* and a can not be *strongly exceptional* in the set B). Therefore we consider the whole set $SX_A \cup B$. By (4.1) for any vertex $v \in SX_A \cup B$ we have

$$deg_3\left(v, \binom{SX_A \cup B}{2}\right) \geq (|SX_A|-1)|B| + 1 \geq \binom{3(|SX_A|-1)}{2} + 1$$

where the last inequality holds when n is large enough. So with a simple greedy procedure we find $|SX_A|$ disjoint edges in $H|_{SX_A \cup B}$ and remove these edges from H. Note that this is the only place where we critically use the minimum degree. We let $A' = A \setminus SX_A$ and B' has all other remaining vertices. Again as above we have $n' = n - 3|SX_A|$, $|A'| = |A| - |SX_A|$ and $|B'| = |B| - 2|SX_A|$ hence we get $|B'| = 2|A'| = 2n'/3$.

Having dealt with the *strongly exceptional* vertices, the vertices of X_A and X_B in A' and B' can be eliminated using the fact that their sizes are much smaller than the crossing degrees of vertices in those sets. For instance as observed above we have $|X_A| \leq 18\sqrt{\alpha}|A|$ while for any vertex $a \in X_A$, we have that $deg_3(a, \binom{B'}{2}) \geq \alpha^{1/3}\binom{|B'|}{2}/2$ (because $a \notin SX_A$). Therefore by a simple greedy procedure, using the fact that $\alpha^{1/3}\binom{|B'|}{2}/2$ is much larger than $54\sqrt{\alpha}|A||B'|$, for each $a \in X_A$ we delete an edge that contains a and two vertices from B'. Similarly for each $b \in X_B$ we delete an edge that contains b and uses one vertex from A' and the other vertex from B' distinct from b. Clearly we can find such disjoint edges, hence we removed a partial matching that covers all vertices in the *strongly exceptional* and *exceptional* sets.

Finally in the leftover sets of A' and B' (denote them by A'' and B'', by construction we still have $|B''| = 2|A''|$) we will find $|A''|$ disjoint edges each using one vertex in A'' and two vertices in B''. Note that for every vertex $a \in A''$ we have $deg_3(a, \binom{B''}{2}) \geq (1 - 2\alpha^{1/3})\binom{|B''|}{2}$ (as $a \notin X_A$). We say that two vertices $b_i, b_j \in B''$ are *good* for each other if $(b_i, b_j, a_k) \in E(H)$ for at least $(1 - 40\alpha^{1/4})|A''|$ vertices a_k in A''. We have that any vertex $b_i \in B''$ is *good* for at least $(1 - 40\alpha^{1/4})|B''|$ other vertices in B'' (again this is so because $b_i \notin X_B$).

We randomly select a set P_1 of $100\alpha^{1/4}|B''|$ vertex disjoint pairs of vertices in B''. By the above observation with high probability every vertex $a \in A''$ make edges in H with at least $3|P_1|/4$ pairs in P_1 and every pair in P_1 makes an edge with at least $3|A''|/4$ vertices in A''. In $B'' \setminus V(P_1)$ still every vertex is *good* for almost all other vertices. We pair up each vertex of $B'' \setminus V(P_1)$ with a distinct vertex in $B'' \setminus V(P_1)$ such that the paired vertices are *good* for each other. This can be done by considering a 2-graph with vertex set $B'' \setminus V(P_1)$ and all the *good* pairs as its edges. A simple application of Dirac theorem on this 2-graph gives such a perfect matching of vertices in $B'' \setminus V(P_1)$. Let the set of these pairs be P_2.

Now construct an auxiliary bipartite graph $G(L, R)$, such that $L = A''$ and vertices in R corresponds to the pairs in P_1 and P_2. A vertex in $a_k \in L$ is connected to a vertex $y \in R$ if the pair corresponding to y (say b_i, b_j) is such that $(b_i, b_j, a_k) \in E(H)$. We will show that $G(L, R)$ satisfies the König-Hall criteria. Considering the sizes of A'' and P_1 it is easy to see that for every subset $Q \subset R$ if $|Q| \leq (1 - 40\alpha^{1/4})|A''|$ then $|N(Q)| \geq |Q|$. When $|Q| > (1 - 40\alpha^{1/4})|A''|$ (using $|B''| = 2|A''|$) any such Q must have at least $6|P_1|/10$ vertices corresponding to pairs in P_1, hence with high probability $N(Q) = L \geq |Q|$. Therefore there is a perfect matching of R into L. This perfect matching in G readily gives us a matching in H covering all vertices in A'' and B'', which together with the edges we already removed (covering *strongly exceptional* and *exceptional* vertices) is a perfect matching in H. $\square$

Chapter 4

Perfect Matchings in 4-uniform hypergraphs

4.1 Introduction and Notation

For graphs we follow the notation in [6]. For a set T, we refer to all of its k-element subsets (k-sets for short) as $\binom{T}{k}$ and the number of such k-sets as $\binom{|T|}{k}$. We say that $H = (V(H), E(H))$ is an r-uniform hypergraph or r-graph for short, where $V(H)$ is the set of vertices and $E \subset \binom{V(H)}{r}$ is a family of r-sets of $V(H)$. When the graph referred to is clear from the context we will use V instead of $V(H)$ and will identify H with $E(H)$. For an r-graph H and a set $D = \{v_1, \ldots, v_d\} \in \binom{V}{d}, 1 \leq d \leq r$, the degree of D in H, $deg_H(D) = deg_r(D)$ denotes the number of edges of H that contain D. For $1 \leq d \leq r$, let

$$\delta_d = \delta_d(H) = \min\left\{deg_r(D) \;:\; D \in \binom{V}{d}\right\}$$

We say that $H(V_1, \ldots, V_r)$ is an r-partite r-graph, if there is a partition of $V(H)$ into r sets, i.e. $V(H) = V_1 \cup \cdots \cup V_r$ and every edge of H uses exactly one vertex from each V_i. We call it a balanced r-partite r-graph if all V_i's are of the same size. We say $H(V_1, \ldots, V_r)$ is a complete r-partite r-graph if every r-tuple that uses one vertex from each V_i belongs to $E(H)$. We denote a complete balanced r-partite r-graph by $K^{(r)}(t)$, where $t = |V_i|$. For $r = 3$, we refer to the balanced 3-partite 3-graph $H(V_1, V_2, V_3)$, where $|V_i| = 4$ as a $4 \times 4 \times 4$ 3-graph.

For an r-graph H, when A and B make a partition of $V(H)$, for a vertex $v \in A$ we denote by $deg_r\left(v, \binom{B}{r-1}\right)$ the number of $(r-1)$-sets of B that make edges with v while $e_r\left(A, \binom{B}{r-1}\right)$ is the sum of $deg_r\left(v, \binom{B}{r-1}\right)$ over all $v \in A$ and $d_r\left(A, \binom{B}{r-1}\right) = e_r\left(A, \binom{B}{r-1}\right) / |A|\binom{|B|}{r-1}$. We denote by $H\left(A, \binom{B}{r-1}\right)$, such an r-graph, when all edges use one vertex from A and $r-1$ vertices from B. Similarly $H\left(A, B, \binom{C}{r-2}\right)$ is an

r-graph where A, B and C make a partition of $V(H)$, and every edge in H uses one vertex each from A and B and $r-2$ vertices from C. Degrees of vertices in A and B are similarly defined as above. The density of $H\left(A, B, \binom{C}{r-2}\right)$ is

$$d_r\left(A, B, \binom{C}{r-2}\right) = \frac{\left|E\left(H\left(A, B, \binom{C}{r-2}\right)\right)\right|}{|A||B|\binom{|C|}{r-2}}$$

When $A_1, \ldots, A_r$ make a partition of $V(H)$, for a vertex $v \in A_1$ we denote by $deg_r(v, (A_2 \times \cdots \times A_r))$ the number of edges in the r-partite r-graph induced by subsets $\{v\}, A_2, \ldots, A_r$, and $e(A_1, (A_2 \times \cdots \times A_r))$ is the sum of $deg_r(v, (A_2 \times \cdots \times A_r))$ over all $v \in A_1$. Similarly

$$d_r(A_1, (A_2 \times \cdots \times A_r)) = \frac{e(A_1, (A_2 \times \cdots \times A_r))}{|A_1 \times A_2 \times \cdots \times A_r|}$$

An r-graph H on n vertices is η-*dense* if it has at least $\eta\binom{n}{r}$ edges. We use the notation $d_r(H) \geq \eta$ to refer to an η-*dense* r-graph H. For $U \subset V$, $H|_U$ is the restriction of H to U. For simplicity we refer to $d_r(H|_U)$ as $d_r(U)$ and to $E(H|_U)$ as $E(U)$. A matching in H is a set of disjoint edges of H and a perfect matching is a matching that contains all vertices. Moreover we will only deal with r-graphs on n vertices where $n = rk$ for some integer k, we denote this by $n \in r\mathbb{Z}$.

Definition 3. *Let d, r and n be integers such that $1 \leq d < r$, and $n \in r\mathbb{Z}$. Denote by $m_d(r, n)$ the smallest integer m, such that every r-graph H on n vertices with $\delta_d(H) \geq m$ contains a perfect matching.*

For graphs ($r = 2$), by the Dirac theorem on Hamiltonicity of graphs, it s easy to see that $m_1(2, n) \leq n/2$, and since the complete bipartite $K_{n/2-1, n/2+1}$ does not have a perfect matching we get $m_1(2, n) = n/2$. For $r \geq 3$ and $d = r - 1$, it follows from a result of Rödl, Ruciński and Szemerédi on Hamiltonicity of r-graph [32] that $m_{r-1}(r, n) \leq n/2 + o(n)$. Kühn and Osthus [27] improved this result to $m_{r-1}(r, n) \leq n/2 + 3r^2\sqrt{n \log n}$. This bound was further sharpened in [31] to $m_{r-1}(r, n) \leq n/2 + C \log n$. In [33] the bound was improved to almost the true value, it was proved that $m_{r-1}(r, n) \leq n/2 + r/4$. Finally [34] settled the problem for $d = r - 1$.

The case $d < r - 1$ is rather hard, in [29] it is proved that for all $d \geq r/2$, $m_d(r, n)$ is close to $\frac{1}{2}\binom{n-d}{r-d}$. For $1 \le d < r/2$ in [26] it was proved that

$$m_d(r,n) \le \left(\frac{r-d}{r} + o(1)\right)\binom{n-d}{r-d}$$

A recent survey of these and other related results appear in [35]. In [26] the authors posed the following conjecture.

Conjecture 3 ([26])**.** *For all* $1 \le d < r/2$,

$$m_d(r,n) \sim \max\left\{\frac{1}{2}, 1 - \left(\frac{r-1}{r}\right)^{r-d}\right\}\binom{n-d}{r-d}$$

Note that for $r = 4$ and $d = 1$ the above bound yields

$$m_1(3,n) \sim \frac{37}{64}\binom{n-1}{3}$$

For $r = 3$ and $d = 1$ this conjecture was proved in [36] and [28]. Markstöm and Ruciński in [37] improved the bound on $m_d(r,n)$ for $1 \le d < r/2$ slightly, by proving

$$m_d(r,n) \le \left(\frac{r-d}{r} - \frac{1}{r^{r-d}} + o(1)\right)\binom{n-d}{r-d}$$

Furthermore for $r = 4$ and $d = 1$ in [37] the authors proved that $m_1(4,n) \le \left(\frac{42}{64} + o(1)\right)\binom{n-1}{3}$. In this chapter we settle Conjecture 3 for the case $r = 4$ and $d = 1$. The main result in this chapter is the following theorem.

Theorem 25. *There exist an integer* n_0 *such that if* H *is a* 4*-graph on* $n \ge n_0$ *(*$n \in 4\mathbb{Z}$*) vertices, and*

$$\delta_1(H) \ge \binom{n-1}{3} - \binom{3n/4}{3} + 1 \tag{4.1}$$

then H *has a perfect matching.*

On the other hand the following construction from [26] shows that this result is tight.

Construction 2. *Let* A *and* B *be disjoint sets with* $|A| = \frac{n}{4} - 1$ *and* $|B| = n - |A|$*. Let* $H = (V(H), E(H))$ *be a* 4*-graph such that* $V(H) = A \cup B$ *and* $E(H)$ *is the set of all* 4*-tuples of vertices,* T*, such that* $|T \cap A| \ge 1$*.*

We have $\delta_1(H) = \binom{n-1}{3} - \binom{3n/4}{3}$ (the degree of a vertex in B) but since every edge in a matching must use at least one vertex from A, the maximum matching in this graph is of size $|A| = \frac{n}{4} - 1$

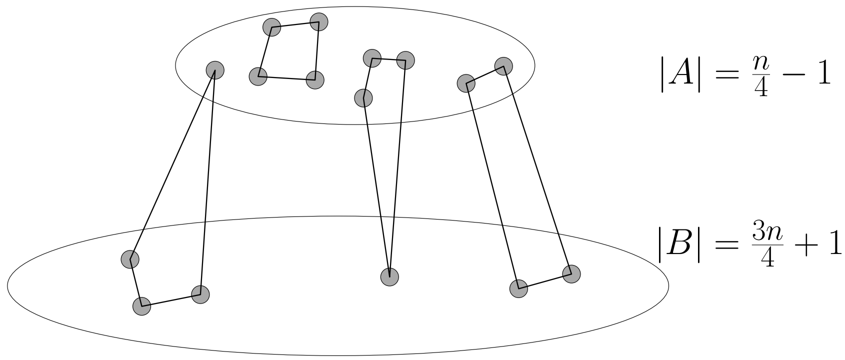


Figure 4.1: The extremal example: A Quadrilateral represent an edge. Every edge intersects the set A.

4.1.1 Outline of the proof

We distinguish two cases to prove Theorem 25. When H is *'far from'* the extremal example as in Construction 2 we use the *absorbing lemma* to find a perfect matching in H. The absorbing lemma (Lemma 34) roughly states, that any 3-graph H satisfying (4.1), contains a small matching M with the property that every "not too large" subset of vertices can be absorbed into M. In Section 4.4, we first remove an absorbing matching M from H. Then using the tools developed in section 4.2 and the auxiliary results of section 4.3, we find an *almost perfect matching* in $H|_{V \setminus V(M)}$. The left over vertices are absorbed into M to get a perfect matching in H. We will show that we can either find an almost perfect matching or H is in the extremal case.

For a constant $0 < \alpha < 1$, we say that H is *α-extremal*, if the following is satisfied, otherwise it is *α-non-extremal*.

Definition 4 (Extremal Case with parameter α)**.** *There exists a $B \subset V(H)$ such that*

- $|B| \geq \left(\frac{3}{4} - \alpha\right) n$
- $d_4(B) < \alpha$.

In section 4.5, using a König-Hall type argument, we build a perfect matching in H, when it is α-extremal.

4.2 Main tools

We use the following two easy observations that will, nevertheless, be useful later on.

Fact 26. *If $G(A,B)$ is a 2η-dense bipartite graph, then there must be at least $\eta|B|$ vertices in B for which the degree in A is at least $\eta|A|$.*

Indeed, otherwise the total number of edges would be less than

$$\eta|A||B| + \eta|A||B| = 2\eta|A||B|$$

a contradiction to the fact that $G(A,B)$ is 2η-dense.

Fact 27. *If $G(A,B)$ is a bipartite graph, $|A| = c_1 \log n$, $|B| = c_2 n$ and for every vertex $b \in B$ $deg(b,A) \geq \eta|A|$, then we can find a complete bipartite subgraph $G'(A',B')$ of G such that $A' \subset A, B' \subset B, |A'| \geq \eta|A|$ and $|B'| \geq c_2 n^{(1-c_1)}$.*

To see this consider the neighborhoods in A, of the vertices in B. Since there can be at most $2^{|A|} = n^{c_1}$ such neighborhoods, by averaging there must be a neighborhood that appears for at least $\frac{c_2 n}{n^{c_1}} = c_2 n^{(1-c_1)}$ vertices of B. This means that we can find the desired complete bipartite graph.

The main tool in this chapter, as in the previous one, is the following result of Erdös [25], on complete balanced r-partite subhypergraphs of r-graphs.

Lemma 28. *For every integer $l \geq 1$ there is an integer $n_0 = n_0(r,l)$ such that: Every r-graph on $n > n_0$ vertices, that has at least $n^{r-1/l^{r-1}}$ edges, contains a $K^{(r)}(l)$.*

In particular Lemma 28 implies that for $\eta > 0$ and a sufficiently small constant β, if H is an r-graph on $n > n_0(\eta,\beta)$ vertices with

$$|E(H)| \geq \eta \binom{n}{r}$$

then H contains a $K^{(r)}(t)$, where

$$t = \beta(\log n)^{1/r-1}.$$

This is so because $\eta\binom{n}{r} \geq \frac{n^r}{n^{1/\beta^{r-1}\log n}} = \frac{n^r}{2^{1/\beta^{r-1}}}$, when $\eta > \frac{2r!}{2^{1/\beta^{r-1}}}$.

The following three lemmas are repeatedly used in section 4.4.

Lemma 29. *Let $H(A, \binom{B}{3})$ be a 4-graph, $|A| = c_1 m$, $|B| = c_2 2^{m^3}$, for some constants $0 < c_1, c_2 < 1$, if*

$$d_4\left(A, \binom{B}{3}\right) \geq 2\eta$$

then there exists a complete 4-partite 4-graph $H'(A_1, B_1, B_2, B_3)$, with $A_1 \subset A$ and B_1, B_2 and B_3 are disjoint subsets of B such that $|A_1| = |B_1| = |B_2| = |B_3| = \eta|A|$.

Proof. First apply Fact 26 to get a subset of 3-sets of vertices in B, $T_1 \subset \binom{B}{3}$, such that every 3-set in T_1 makes edges with at least $\eta|A|$ vertices in A and $|T_1| \geq \eta\binom{|B|}{3} \geq \eta|B|^3/10$. Next we find a $T_2 \subset T_1$, such that all 3-sets in T_2 make edges with the same subset of A (say $A_1 \subset A$). By Fact 27 we have

$$|T_2| \geq \frac{\eta|B|^3/10}{2^{c_1 m}} = \frac{\eta}{10}\frac{\left(c_2 2^{m^3}\right)^3}{2^{c_1 m}} = \frac{c_2^3\eta}{10}2^{m^3(3-c_1/m^2)} = \frac{c_2^3\eta}{10}\left(\frac{|B|}{c_2}\right)^{3-c_1/m^2} \geq |B|^{3-1/(\eta|A|)^2}$$

where the last inequality follows when m is sufficiently large and η, c_1 and c_2 are small constants.

Now construct an auxiliary 3-graph H_1 where $V(H_1) = B$ and edges of H_1 corresponds to 3-sets in T_2. Applying lemma 28 on H_1 (for $r = 3$) we get a complete 3-partite 3-graph with color classes B_1, B_2 and B_3 each of size $\eta|A|$. Clearly A_1, B_1, B_2 and B_3 corresponds to color classes of a complete 4-partite 4-graph in H as in the statement of the lemma. □

Lemma 30. *Let $H(A, B, \binom{Z}{2})$ be a 4-graph with $|A| = |B| = c_1 m$, $|Z| = c_2 2^{m^3}$ for some constants $0 < c_1, c_2 < 1$, if*

$$d_4\left(A, B, \binom{Z}{2}\right) \geq 2\eta$$

then there exists a complete 4-partite 4-graph $H'(A', B', Z_1, Z_2)$, with $A' \subset A$, $B' \subset B$ and Z_1 and Z_2 are disjoint subsets of Z such that $|A'| = |B'| = |Z_1| = |Z_2| = c \log m$ for a constant $c = c(c_1, c_2, \eta)$.

Proof. First apply Fact 26 to get a subset of pairs of vertices in Z, $P_1 \subset \binom{Z}{2}$, such that every pair in P_1 makes edges with at least $\eta|A||B|$ pairs of vertices in $A \times B$ and $|P_1| \geq \eta\binom{|Z|}{2} \geq \eta|Z|^2/3$. Next we find a $P_2 \subset P_1$, such that all pairs in P_2 make edges with the same subset of pairs in $A \times B$ (say $Q_1 \subset A \times B$). By Fact 27 we have

$$|P_2| \geq \frac{\eta|Z|^2/3}{2^{(c_1 m)^2}} = \frac{\eta}{3}\frac{\left(c_2 2^{m^3}\right)^2}{2^{(c_1 m)^2}} = \frac{c_2^2\eta}{3}2^{m^3(2-c_1^2/m)} = \frac{c_2^2\eta}{3}\left(\frac{|Z|}{c_2}\right)^{2-c_1^2/m} \geq |Z|^{2-1/(\log m)}$$

where the last inequality follows when m is sufficiently large, and η, c_1 and c_2 are small constants.

Now construct an auxiliary 2-graph G_1 where $V(G_1) = Z$ and edges of G_1 corresponds to pairs in P_2. Applying lemma 28 on G_1 (for $r = 2$) we get a complete bipartite graph with color classes Z_1 and Z_2 each of size $\log m$.

Similarly construct an auxiliary bipartite graph G_2 with color classes A and B, and edges of G_2 corresponds to pairs in Q_1. Since we have $|Q_1| \geq \eta|A||B|$ applying lemma 28 on G_2 (for $r = 2$) we get a complete bipartite graph with color classes A' and B', each of size $c \log m$. Clearly A', B' and a subset of vertices each from Z_1 and Z_2 of size $|A'|$ corresponds to color classes of the balanced complete 4-partite 4-graph in H as in the statement of the lemma. □

Lemma 31. *Let $H(A, B, C, Z)$ be a 4-partite 4-graph with $|A| = |B| = |C| = c_1 m$ and $|Z| = c_2 2^{m^3}$ for some constants $0 < c_1, c_2 < 1$. If*

$$d_4(Z, (A \times B \times C)) \geq 2\eta$$

then there exists a complete 4-partite 4-graph $H'(A', B', C', Z')$ such that $|A'| = |B'| = |C'| = |Z'| = \beta\sqrt{\log|A|}$.

Proof. First apply Fact 26 on H to get a subset Z_1 of Z such that for every vertex $z \in Z_1$, $deg_3(z, (A \times B \times C)) \geq \eta|A|^3$ and $|Z_1| \geq \eta|Z|$. Now consider the auxiliary bipartite graph $G(D, Z_1)$, where $D = A \times B \times C$ and a vertex $z \in Z_1$ is connected to a 3-set $(a, b, c) \in D$ if (a, b, c, z) makes an edge of H. An application of Fact 27 on G

gives a complete bipartite graph $G_2(D', Z_1')$ where

$$|D'| \geq \eta |A|^3 \quad \text{and} \quad Z_1' \geq \eta c_2 2^{m^3(1-c_1^3)} > |A|$$

when m is sufficiently large.

Let G_3 be a 3-partite 3-graph where the color classes are A, B and C and edges correspond to 3-sets in D'. Since $|D'| \geq \eta |A|^3$ applying Lemma 28 on G_3 ($r = 3$), we get a balanced complete 3-partite 3-graph in G_3 with color classes A', B' and C', such that $|A'| = |B'| = |C'| \geq \beta \sqrt{\log |A|}$. Clearly A', B', C' and a subset of Z_1' of size $|A'|$, correspond to the color classes of the required complete 4-partite 4-graph. $\square$

We will frequently apply the following folklore statements (proofs are omitted):

Lemma 32. *Every graph H has a subgraph H' such that $\delta_1(H') \geq |E(H)|/|V(H)|$.*

Lemma 33. *Every 3-graph H on n vertices, with $\delta_1(H) \geq \eta \binom{n}{3}/n$, has a matching of size $\eta n/24$.*

Finally we use the following lemma from [26].

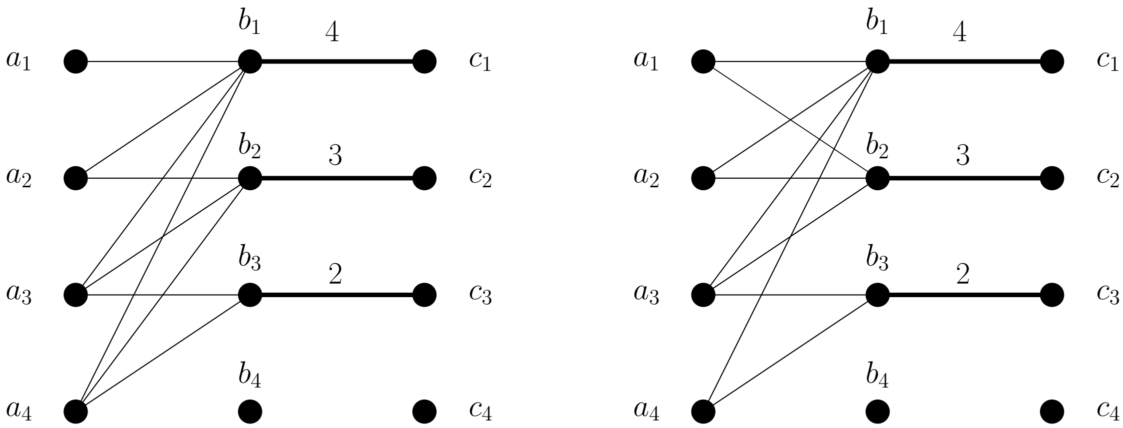
Lemma 34. *(Absorbing Lemma) For every $\eta > 0$, there is an integer $n_0 = n_0(\eta)$ such that if H is a 4-graph on $n \geq n_0$ vertices with $\delta_1(H) \geq (1/2 + 2\eta)\binom{n}{3}$, then there exist a matching M in H of size $|M| \leq \eta^4 n/4$ such that for every set $W \subset V \setminus V(M)$ of size at size at most $\eta^8 n \geq |W| \in 4\mathbb{Z}$, there exists a matching covering exactly the vertices in $V(M) \cup W$.*

4.3 Auxiliary results

For a $4\times4\times4$ 3-graph, denote by Q_1, Q_2, Q_3 its 3 color classes and let $Q_1 = \{a_1, a_2, a_3, a_4\}$, $Q_2 = \{b_1, b_2, b_3, b_4\}$ and $Q_3 = \{c_1, c_2, c_3, c_4\}$. For two vertices, $x \in Q_i,\ y \in Q_j,\ i \neq j$, $N(x, y)$ denotes the neighborhood of the pair x, y, i.e. the set of vertices in the third color class that make edges with the pair x, y. Let $deg(x, y) = |N(x, y)|$. For two disjoint pairs (x_1, y_1) and (x_2, y_2) in $Q_i \times Q_j$, $i \neq j$ the pairs (x_1, y_2) and (x_2, y_1) are called the crossing pairs and the value $deg(x_1, y_2) + deg(x_2, y_1)$ is referred to as the *crossing degree sum.* We define the following four special $4 \times 4 \times 4$ 3-graphs.

Definition 5. *H_{432} is a $4 \times 4 \times 4$ 3-graph, such that there exist 3 disjoint pairs in a $Q_i \times Q_j$, $i \neq j$, with degrees at least 4, 3 and 2 respectively.*

In the figures the triplets joined by a line represent an edge in the 3-graph. Let the H_{432} be as in Figure 4.2(a) or Figure 4.2(b). If the H_{432} does not have a perfect matching then we must have $deg(b_4, c_4) = 0$, (as otherwise by the König-Hall criteria we get a perfect matching).



(a) H_{432} with $|N(b_2, c_2) \cup N(b_3, c_3)| = 3$ (b) H_{432} with $|N(b_2, c_2) \cup N(b_3, c_3)| = 4$

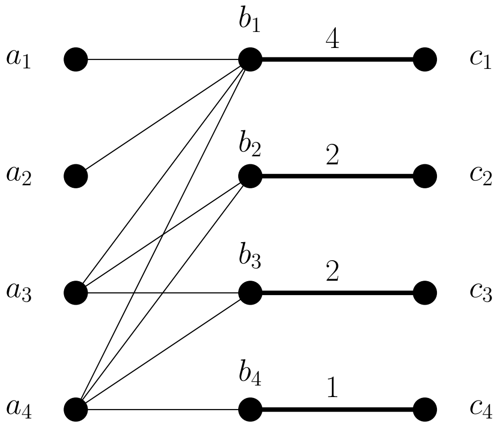
Figure 4.2: The two types of H_{432}. Labels on edges is the degree of the pair.

Definition 6. *H_{4221} is a $4 \times 4 \times 4$ tripartite 3-graph, such that there exist 4 disjoint pairs in a $Q_i \times Q_j$, $i \neq j$, with degrees at least 4, 2, 2 and 1 respectively.*

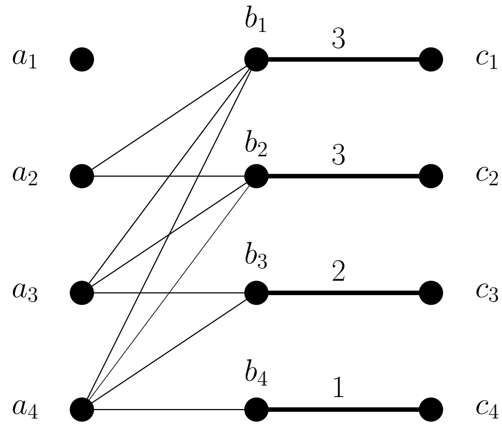
Let the H_{4221} be as in Figure 4.3(a). If $|N(b_2, c_2) \cup N(b_3, c_3) \cup N(b_4, c_4)| = 3$ then again by the König-Hall criteria we get that H_{4221} has a perfect matching. We only consider the H_{4221} that has no perfect matching.

Definition 7. *H_{3321} is a $4 \times 4 \times 4$ 3-graph, such that there exist 4 disjoint pairs in a $Q_i \times Q_j$, $i \neq j$, with degrees at least 3, 3, 2 and 1 respectively (see Figure 4.3(b)).*

Definition 8. *H_{ext} is a $4 \times 4 \times 4$ 3-graph with exactly 37 edges such that there are three vertices, one in each of Q_1, Q_2 and Q_3, and all edges are incident to at least one of these three vertices.*



(a) H_{4221} with no perfect matching



(b) H_{3321} with no perfect matching

The following lemma that classifies, $4 \times 4 \times 4$ 3-graphs with at least 37 edges, will be very useful in the subsequent section.

Lemma 35. *Let $H(Q_1, Q_2, Q_3)$ be a $4 \times 4 \times 4$ 3-graph. If $|E(H)| \geq 37$ then one of the following must be true*

1. *H has a perfect matching.*
2. *H has a subgraph isomorphic to H_{3321}*
3. *H has a subgraph isomorphic to H_{432}*
4. *H has a subgraph isomorphic to H_{4221}*
5. *H is isomorphic to H_{ext}.*

Proof. We consider the following cases based on degree of pairs in $Q_i \times Q_j$.

Case 1: There is Q_i and Q_j, $i \neq j$ such that no pair in $Q_i \times Q_j$ has degree 4.

Let Q_2, Q_3 be such a pair. Since $|E(H)| \geq 37$ and no pair has degree 4, at least 5 out of the 16 pairs in $Q_2 \times Q_3$ must be of degree 3. Which implies that there must be at least 2 disjoint pairs of degree 3. Consider the largest set of disjoint pairs of degree 3 in $Q_2 \times Q_3$. Assume that there are 3 disjoint pairs in $Q_2 \times Q_3$ of degree 3, say (b_1, c_1), (b_2, c_2) and (b_3, c_3).

If $deg(b_4, c_4) \geq 1$, then we have a H_{3321}. So assume that $deg(b_4, c_4) = 0$, The total number of edges incident to pairs in $\{b_1, b_2, b_3\} \times \{c_1, c_2, c_3\}$ is at most 27 (as there are 9 pairs and degree of every pair is at most 3), the remaining at least 10 edges are incident to either b_4 or c_4. Which implies that there must be at least one pair (say (b_3, c_3)), such that crossing degree sum of (b_3, c_3) and (b_4, c_4) is at least 4. Therefore we have that the degree of one crossing pair is at least 2 and that of the other crossing pair is at least 1. These two crossing pairs together with (b_1, c_1) and (b_2, c_2) gives us a subgraph isomorphic to H_{3321}.

On the other hand, if there are exactly 2 disjoint pairs of degree 3, say (b_1, c_1) and (b_2, c_2). Again the total number of edges incident to pairs in $\{b_1, b_2\} \times \{c_1, c_2\}$ is at most 12 (as there are 4 pairs and degree of every pair is at most 3). If there is a pair in $\{b_3, b_4\} \times \{c_3, c_4\}$ (say (b_3, c_3)) such that the *crossing degree sum* of (b_1, c_1) and (b_3, c_3) is at least 6, then since there is no degree 4 pair we must have that both $deg(b_1, c_3)$ and $deg(c_1, b_3)$ are 3. Now these crossing pairs together with (a_2, b_2) are 3 disjoint pairs of degree 3 which is a contradiction to the maximality of the set of disjoint pairs of degree 3. Therefore we must have that the sum of degrees of pairs in $\{b_1, b_2\} \times \{c_3, c_4\}$ and $\{c_1, c_2\} \times \{b_3, b_4\}$ is at most $4 \times 5 = 20$. Hence the number of edges of H incident to pairs in $\{b_3, b_4\} \times \{c_3, c_4\}$ is at least $37 - 12 - 20 = 5$ and no pair has degree 3. Therefore, in $\{b_3, b_4\} \times \{c_3, c_4\}$, we can find two disjoint disjoint pairs, (say (b_3, c_3) and (b_4, c_4)) with degree at least 2 and 1 respectively and we get a graph isomorphic to H_{3321}.

Case 2: There is a Q_i and Q_j, $i \neq j$ such that exactly one disjoint pair in $Q_i \times Q_j$ has degree 4.

Let Q_2, Q_3 be such a pair. Consider the largest set of disjoint pairs in $Q_2 \times Q_3$ with one pair of degree 4 and the remaining of degree 3. Note that if there are two disjoint pairs of degree 3 besides the degree 4 pair in the selected set, then clearly we have an H_{432}. So we consider the following two subcases based on whether or not there is a pair of degree 3 in the selected set. Let (b_1, c_1) be the degree 4 pair in the selected set.

Subcase 2.1 There is another pair in $Q_2 \times Q_3$ disjoint from (b_1, c_1) with degree 3.

Assume that $deg(b_2, c_2) = 3$. First observe that

1. If any pair in $\{b_3, b_4\} \times \{c_3, c_4\}$ has degree at least 2 then that pair together with (b_1, c_1) and (b_2, c_2) makes an H_{432}.

2. If both $deg(b_1, c_2)$ and $deg(b_2, c_1)$ are 4 then we get two disjoint degree 4 pairs. Therefore we have that the number of edges incident to pairs in $\{b_1, b_2\} \times \{c_1, c_2\}$ is at most $4 + 3 + 4 + 3 = 14$.

3. If there is a pair in $\{b_3, b_4\} \times \{c_3, c_4\}$ (say (b_3, c_3)) such that the *crossing degree sum* of (b_2, c_2) and (b_3, c_3) is at least 5. Then we must have that one crossing pair is of degree at least 3, and the other is of degree at least 2 (because none of them can be of degree 4). These 2 crossing pairs together with (b_1, c_1) makes the disjoint pairs of an H_{432}. Therefore we must have that (b_2, c_2) and any pair in $\{b_3, b_4\} \times \{c_3, c_4\}$ have their *crossing degree sum* at most 4.

4. Similarly (b_1, c_1) and any pair in $\{b_3, b_4\} \times \{c_3, c_4\}$ have their *crossing degree sum* at most 6.

Assume that (b_1, c_1) and (b_3, c_3) have their crossing degree sum, equal to 6. If the degrees of crossing pairs are 4 and 2, then these crossing pairs and (b_2, c_2) makes the disjoint pairs of an H_{432}. On the other hand if both the crossing pairs have degree 3. Then (b_1, c_3), (b_3, c_1) and (b_2, c_2) are three disjoint pairs of degree 3. From observation 1 we have $deg(b_3, c_3) \leq 1$ and from observation 3 the crossing degree sum of (b_3, c_3) and (b_2, c_2) is at most 4. Which together with observation 2 gives us that the total number of edges incident to pairs in $\{b_1, b_2, b_3\} \times \{c_1, c_2, c_3\}$ is at most $14+6+1+4 = 25$. Now if $deg(b_4, c_4) = 1$ then we have an H_{3321}, otherwise we have that the number of edges containing either b_4 or c_4 is at least $37 - 25 \geq 12$. By observation 1 we have that both $deg(b_4, c_3)$ and $deg(b_3, c_4)$ are at most 1 hence we must have that the *crossing degree sum* of (b_2, c_2) and (b_4, c_4) is 4 with one crossing pair of degree at least 1 and the other of degree at least 2. These crossing pairs together with (b_1, c_3) and (b_3, c_1) gives us an H_{3321}.

On the other hand if for any pair in $\{b_3, b_4\} \times \{c_3, c_4\}$ and (b_1, c_1) their *crossing degree* *sum* is be at most 5. Then the number of edges incident to pairs in $\{b_3, b_4\} \times \{c_3, c_4\}$ is at least $37 - 14 - 4(2) - 5(2) = 5$. Which implies that there must a degree 2 pair, and hence by observation 1, we get an H_{432}.

Subcase 2.2 There is no pair of degree 3 disjoint from (b_1, c_1).

In this case again as in observation 2, the *crossing degree sum* of (b_1, c_1) and any pair in $\{b_2, b_3, b_4\} \times \{c_2, c_3, c_4\}$ is at most 6 (as any other case results in two disjoint pairs of degree 4 and 3). This implies that the number of edges incident to pairs in $\{b_2, b_3, b_4\} \times \{c_2, c_3, c_4\}$ is at least $37 - 4 - 3(6) = 15$ and no pair has degree 3. Which implies that there are three disjoint pairs in $\{b_2, b_3, b_4\} \times \{c_2, c_3, c_4\}$ with degrees 2, 2 and at least 1 respectively. These pairs and (b_1, c_1) makes the 4 disjoint pairs of an H_{4221}.

Case 3: In every Q_i and Q_j, $i \neq j$ there are exactly two disjoint pairs in $Q_i \times Q_j$ with degree 4.

Consider Q_1, Q_2 and assume that (a_1, b_1) and (a_2, b_2) are the two disjoint pairs with degree 4. We make the following observations:

1. If any pair in $\{a_3, a_4\} \times \{b_3, b_4\}$ has degree at least 2 then that pair together with (a_1, b_1) and (a_2, b_2) makes the disjoint pairs of an H_{432}. Therefore the total number of edges spanned by pairs in $\{a_3, a_4\} \times \{b_3, b_4\}$ is at most 4.

2. For any of (a_1, b_1) and (a_2, b_2) and any pair in $\{a_3, a_4\} \times \{b_3, b_4\}$ their *crossing degree sum* can be at most 5. Indeed otherwise say the *crossing degree sum* of (a_1, b_1) and (a_3, b_3) is 6, then we must have that one crossing pair has degree at least 3, and the other has degree at least 2. These crossing pairs and (a_2, b_2) make the disjoint pairs of an H_{432}. Furthermore if any such *crossing degree sum* is 5 then by the same reasoning as above, it must be that one crossing pair is of degree 4 and the other is of degree 1.

3. If the total number of edges spanned by pairs in $\{a_1, a_2\} \times \{b_1, b_2\}$ is at most 12,

then the number of edges that uses one vertex from $\{a_1, a_2, b_1, b_2\}$ and one vertex from $\{a_3, a_4, b_3, b_4\}$ is at least $37 - 12 - 4 = 21$. Hence there will be a pair in $\{a_3, a_4\} \times \{b_3, b_4\}$ (say (a_3, b_3)) such that *crossing degree sum* of (a_3, b_3) and at least one of (a_1, b_1) and (a_2, b_2) is at least 6, and by observation 2 we get an H_{432}.

The above observations are true for any two disjoint pairs of degree 4. We choose two disjoint pairs of degree 4, (say (a_1, b_1) and (a_2, b_2)) pairs in $Q_1 \times Q_2$ such that (i) a_1 has the maximum vertex degree among all vertices that are part of some pairs of degree 4 and (ii) $deg(a_2, b_1)$ is as small as possible. Let $deg(a_2, b_1) = x$. Note that by observation 3, we have $x \geq 1$. We consider the following cases based on the value of x.

Subcase 3.1 $x = 4$

Note that that the number of edges in $\{a_1, a_2\} \times \{b_1, b_2\}$ is at most 16 ($deg(a_1, b_2) \leq 4$). First we will show that both $deg(a_3, b_1), deg(a_4, b_1) \leq 1$. Assume that $deg(a_3, b_1) \geq 2$, but then as in observation 2 both $deg(a_1, b_3)$ and $deg(a_1, b_4)$ can be at most 2. Hence by the maximality of the degree of a_1 we get $deg(a_3, b_1) + deg(a_4, b_1) \leq 4$. Now by observation 1, there must be at least $37 - 16 - 4 - 8 = 9$ edges containing one vertex from $\{a_3, a_4, b_3, b_4\}$ and one of $\{a_2, b_2\}$. Which implies that the degree of a_2 or b_2 is strictly larger than that of a_1, a contradiction. So we have $deg(a_3, b_1) = deg(a_4, b_1) = 1$.

Now we show that both $deg(a_3, b_2), deg(a_4, b_2) \leq 1$. To see this first assume that either $deg(a_1, b_3)$ or $deg(a_1, b_4)$ is equal to 4, (say $deg(a_1, b_3) = 4$) then by the minimality of x we must have that $deg(a_2, b_3) = 4$ too, because if $deg(a_2, b_3) < x$, then we can exchange b_1 with b_3 to get a smaller value of x. But if $deg(a_2, b_3) = 4$ then by observation 2 we must have $deg(a_3, b_2) = deg(a_4, b_2) = 1$. On the other hand if both $deg(a_1, b_3)$ and $deg(a_1, b_4)$ are at most 3 ($deg(a_1, b_3) + deg(a_1, b_4) \leq 6$) then again there must be at least $37 - 16 - 4 - 8 = 9$ edges containing some pair in $\{a_2\} \times \{b_3, b_4\}$ and $\{a_3, a_4\} \times \{b_2\}$. Which means at least one of these pairs must be of degree at least 3. Say $deg(a_2, b_3) \geq 3$, but then by observation 2 we have $deg(a_3, b_2), deg(a_4, b_2) \leq 1$ and we are done. In case say $deg(a_3, b_2) \geq 3$ then we have $deg(a_3, b_2) + deg(a_4, b_2) \geq 7$ and we get that the degree of b_2 is larger than degree of a_1, a contradiction.

So we have that $deg(a_3, b_1), deg(a_4, b_1), deg(a_3, b_2)$ and $deg(a_4, b_2) \leq 1$. This together with observation 1 implies that the number of edges containing some pair in $\{a_1, a_2\} \times \{b_1, b_2, b_3, b_4\}$ is at least $37 - 4 - 4 = 29$. Therefore the $2 \times 4 \times 4$ 3-graph $H'(\{b_1, b_2\}, Q_2, Q_3)$ has at least $37 - 4 - 4 = 29$ edges. There must be at least three vertices in Q_2 such that each one of them is part of at least three pairs in $(Q_2 \times Q_3)$ that are of degree at least 2. To see this assume that there are at most two such vertices in Q_2 (say b_1 and b_2). Then using the fact that the maximum degree of a pair in $Q_2 \times Q_3$ in H' is 2, we get that b_1 and b_2 can be contained in at most $2(4 \cdot 2) = 16$ edges. While at most 2 pairs containing either b_3 and b_4 can be of degree 2, we get that the number of edges containing either b_3 and b_4 is at most $2 \cdot (2 \cdot 2 + 2 \cdot 1) = 12$ which implies that $|E(H')| \leq 28$ a contradiction

Now since in Q_2 there are at least 3 vertices such that each one of them is part of at least 3 pairs in $(Q_2 \times Q_3)$ of degree at least 2. Which implies that there must a degree 2 pair disjoint from the two degree 4 pairs (guaranteed in Case 3) in $Q_2 \times Q_3$. Therefore we get an H_{432}.

Subcase 3.2 $x = 3$

Now we have that the number of edges in $\{a_1, a_2\} \times \{b_1, b_2\}$ is at most 15. Again we first show that both $deg(a_3, b_1), deg(a_4, b_1) \leq 1$. Assume that $deg(a_3, b_1) \geq 2$, but then by observation 2 both $deg(a_1, b_3)$ and $deg(a_1, b_4)$ can be at most 2. Hence by the maximality of the degree of a_1 we get $deg(a_3, b_1) + deg(a_4, b_1) \leq 5$. Now by observation 1, there must be at least $37 - 15 - 4 - 9 = 9$ edges containing one vertex from $\{a_3, a_4, b_3, b_4\}$ and one of $\{a_2, b_2\}$. Which implies that the degree of a_2 or b_2 is strictly larger than that of a_1, a contradiction. So we have $deg(a_3, b_1) = deg(a_4, b_1) = 1$.

Similarly as in the previous case we show that both $deg(a_3, b_2), deg(a_4, b_2) \leq 1$. To see this first assume that either $deg(a_1, b_3)$ or $deg(a_1, b_4)$ is equal to 4, (say $deg(a_1, b_3) = 4$) then by the minimality of x we must have that $deg(a_2, b_3) \geq 3$ too. But if $deg(a_2, b_3) = 3$ then by observation 2 we must have $deg(a_3, b_2), deg(a_4, b_2) \leq 1$. On the other hand if both $deg(a_1, b_3)$ and $deg(a_1, b_4)$ are at most 3 ($deg(a_1, b_3) + deg(a_1, b_4) \leq 6$) then again there must be at least $37-15-4-8=9$ edges containing some pair in $\{a_2\}\times\{b_3,b_4\}$ and $\{a_3,a_4\}\times\{b_2\}$. Which means at least one of these pairs must be of degree at least 3. Say $deg(a_2,b_3)\geq 3$, but then by observation 2 we have $deg(a_3,b_2), deg(a_4,b_2)\leq 1$ and we are done. In case say $deg(a_3,b_2)\geq 3$ then we have $deg(a_3,b_2)+deg(a_4,b_2)\geq 7$ and we get that the degree of b_2 is greater than the degree of a_1, a contradiction.

So we have that $deg(a_3,b_1), deg(a_4,b_1), deg(a_3,b_2)$ and $deg(a_4,b_2)\leq 1$. Again we get that the $2\times 4\times 4$ 3-graph $H'(\{b_1,b_2\},Q_2,Q_3)$ has at least $37-4-4=29$ edges. and we are done.

Subcase 3.3 $x=2$

Similarly as in the previous two subcases we have that $deg(a_3,b_1), deg(a_4,b_1), deg(a_3,b_2)$ and $deg(a_4,b_2)\leq$ 1 and the $2\times 4\times 4$ 3-graph $H'(\{b_1,b_2\},Q_2,Q_3)$ has at least $37-4-4=29$ edges and we are done.

Subcase 3.4 $x=1$

In this case observation 3 implies that the number of edges in $\{a_1,a_2\}\times\{b_1,b_2\}$ is exactly 13 ($deg(a_1,b_2)=4$). Using $|E(H)\geq 37|$ and observation 2 we get that every pair in $\{a_3,a_4\}\times\{b_3,b_4\}$ has degree exactly 1 and for any pair in $\{a_3,a_4\}\times\{b_3,b_4\}$ and any of (a_1,b_1) or (a_2,b_2) their *crossing degree sum* is exactly 5 $(4+1)$.

Therefore, we have that either

$$deg(a_1,b_3)=deg(a_1,b_4)=deg(a_3,b_2)=deg(a_4,b_2)=4 \quad \text{or}$$

$$deg(a_1,b_3)=deg(a_1,b_4)=deg(a_2,b_3)=deg(a_2,b_4)=4$$

In the latter case note that again we have that $H'(\{b_1,b_2\},Q_2,Q_3)$ has at least 29 edges and we are done as above.

So assume that $deg(a_1,b_3)=deg(a_1,b_4)=deg(a_3,b_2)=deg(a_4,b_2)=4$ and $deg(a_1,b_1)=deg(a_2,b_2)=deg(a_1,b_2)=4$ and every other pair in $Q_1\times Q_2$ is of degree exactly 1. This means that a_1 and b_2 are not part of any degree 1 pair. Now the neighborhoods of all the degree 1 pairs must be the same vertex in Q_3 (say c_3). Because otherwise we get a perfect matching in H, (using those two vertices in Q_3 for two disjoint degree 1 pairs, and the remaining two vertices of Q_3 are matched with two of the degree 4 pairs in $(Q_1 \times Q_2)$). But if all of these edges are incident to c_3, then all edges in this graph are incident to at least one vertex in $\{a_1, b_2, c_3\}$ and the number of edges is exactly 37, hence H is isomorphic to H_{ext}. $\square$

Let A, B and C be three disjoint balanced complete 4-partite 4-graphs with color classes $(A_1, \ldots, A_4)$, $(B_1, \ldots, B_4)$ and $(C_1, \ldots, C_4)$ respectively, and $|A_1| = |B_1| = |C_1| = m$. Let Z be a set of vertices disjoint from vertices in A, B and C. For a small constant $\eta > 0$, we say that Z is *connected* to a triplet of color classes (A_i, B_j, C_k), $1 \le i, j, k \le 4$, if $d_4(Z, (A_i \times B_j \times C_k)) \ge 2\eta$. For Z and (A, B, C) we define an auxiliary graph, (the *link graph*), L_{abc} to be a $4 \times 4 \times 4$ 3-graph where the vertex set of each color class of L_{abc} corresponds to the color classes in A, B and C. While a triplet of vertices (a_i, b_j, c_k) is an edge in L_{abc} iff Z is connected to the triplet of color classes (A_i, B_j, C_k).

Given three balanced complete 4-partite 4-graphs A, B and C and another set of vertices Z, as above, we say that we can *extend* (A, B, C) if we can build another set of balanced complete 4-partite 4-graphs using $V(A) \cup V(B) \cup V(C) \cup Z$ such that the total number of vertices in the new 4-partite 4-graphs is at least $12m + \eta m/16$ and the size of a color class in each new 4-partite 4-graph is $\beta\sqrt{\log m}$. In what follows we outline a procedure to extend (A, B, C) using the structure of L_{abc}.

Lemma 36. *For $\eta, c > 0$, let A, B and C be three balanced complete 4-partite 4-graphs such that $|A_1| = |B_1| = |C_1| = m$. If Z is a disjoint set of vertices with $|Z| \ge c2^{m^3}$. If the link graph L_{abc} has at least 37 edges and L_{abc} is not isomorphic to H_{ext} then we can extend (A, B, C).*

Proof. Since L_{abc} is a $4 \times 4 \times 4$ 3-graph with at least 37 edges and is not isomorphic to H_{ext}, for each of the other cases as in Lemma 35, we give the procedure to extend (A, B, C).

Case 1: L_{abc} has a perfect matching:

Without loss of generality assume that the perfect matching in L_{abc} corresponds to $\{(A_i, B_i, C_i)\ : 1 \leq i \leq 4\}$ i.e. Z is connected to the triplets $\{(A_i, B_i, C_i)\ : 1 \leq i \leq 4\}$. Note that by the definition of connectedness and the sizes of the sets the 4-partite 4-graph (A_1, B_1, C_1, Z) satisfies the conditions of Lemma 31. Hence we find a complete balanced 4-partite 4-graph $X_1 = (A_1^1, B_1^1, C_1^1, Z^1)$, such that

$$Z^1 \subset Z,\ \ A_1^1 \subset A_1\ ,\ B_1^1 \subset B_1 \text{ and } C_1^1 \subset C_1 \quad \text{and}$$

$$|Z^1| = |A_1^1| = |B_1^1| = |C_1^1| = \beta\sqrt{\log m} \quad \text{where } \beta \text{ is as in Lemma 31}$$

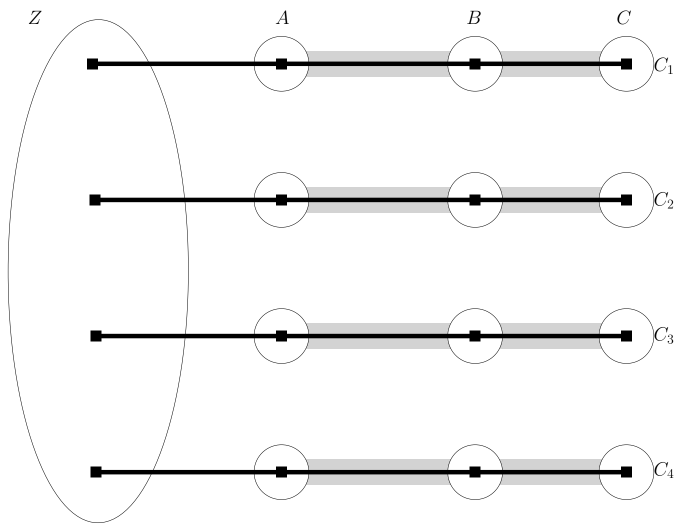


Figure 4.3: Extending (A, B, C) when L_{abc} has a perfect matching: The shaded boxes represent a triplet *connected* to Z, while solid thick lines represent a balanced complete 4 partite graphs

Similarly, we find such complete balanced 4-partite 4-graphs X_2, X_3 and X_4 in (A_2, B_2, C_2, Z), (A_3, B_3, C_3, Z) and (A_4, B_4, C_4, Z) respectively, that are disjoint from each other (as $|Z|$ is very large compared to m) (see Figure 4.3). We remove the vertices in X_1, X_2, X_3 and X_4 and make these four new 4-partite 4-graphs. In the remaining parts of A, B and C we remove another such set of 4 disjoint complete balanced 4-partite 4-graphs. Again by definition of connectedness and Lemma 31 we can continue this process until we remove at least $\eta m/8$ vertices from each color class of A, B and C.

Note that the new 4-partite 4-graphs use at least $4\eta m/8$ vertices from Z. Therefore these new 4-partite 4-graphs together with leftover parts of A, B and C have at least $3(4m)+\eta m/2$ vertices while all the 4-partite 4-graphs are balanced. Hence, we extended (A, B, C).

Case 2: L_{abc} has a subgraph isomorphic to H_{432}:

In this case we show in detail how to extend such an (A, B, C), while in the latter cases we will only briefly outline the procedure. First assume that the H_{432} in L_{abc} is as in Figure 4.2(a) and let the pairs corresponding to the degree 4, 3 and 2 pairs in this subgraph be (B_1, C_1), (B_2, C_2) and (B_3, C_3) respectively. Furthermore let the color classes corresponding to the neighbors of degree 3 and degree 2 pairs in H_{432} be $\{A_2, A_3, A_4\}$ and $\{A_3, A_4\}$ respectively.

Using the definition of connectedness and Lemma 31 we find two disjoint complete balanced 4-partite 4-graphs $X_1 = (A_4^1, B_3^1, C_3^1, Z^1)$ and $X_2 = (A_3^2, B_3^2, C_3^2, Z^2)$ such that Z^j, A_i^j, B_i^j, C_i^j are subsets of Z, A_i, B_i, C_i respectively.

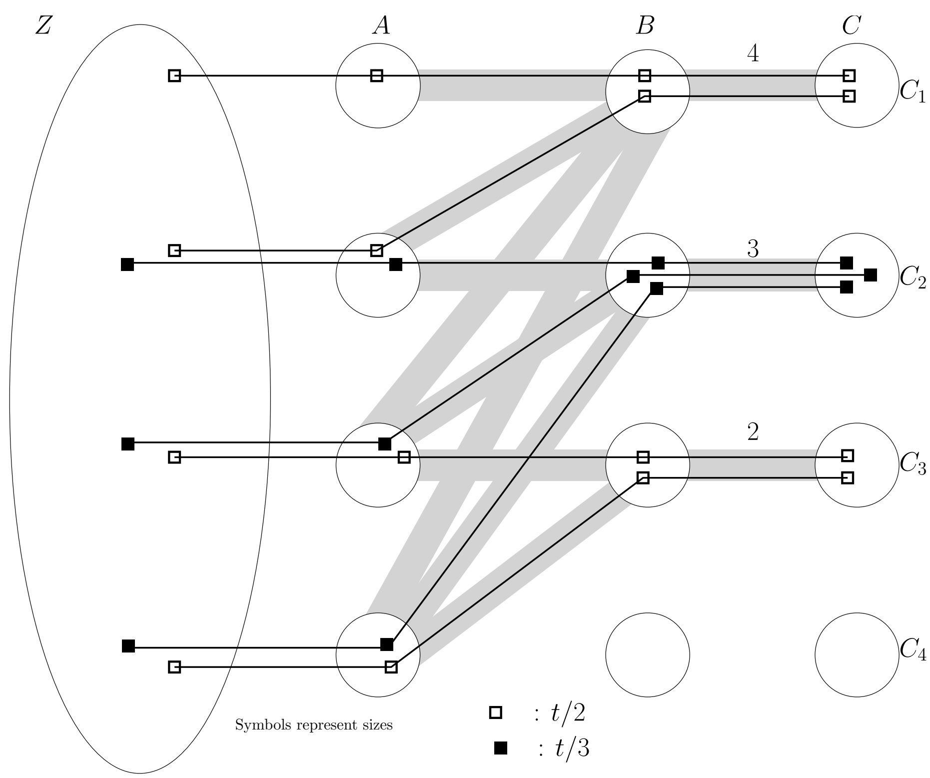


Figure 4.4: Extending (A, B, C) when L_{abc} has an H_{432} as in Figure 4.2(a): The shaded boxes represent a triplet *connected* to Z, while solid think lines represent a balanced complete 4 partite 4-graphs

Similarly we find three more disjoint balanced complete 4-partite 4-graphs X_3, X_4 and X_5 where vertices of three color classes in all of them are from B_2, C_2 and Z while vertices of the fourth color class are from A_2, A_3 and A_4 respectively. We build two more disjoint balanced complete 4-partite 4-graphs X_6 and X_7 such that vertices of three color classes in both of them are from B_1, C_1 and Z while vertices of the fourth color class are from A_1 and A_2 respectively.

The size of a color class in $X_1, \dots, X_7$ is $\beta\sqrt{\log m}$. We remove the vertices in $X_1, \dots X_7$ from their color classes to make these 7 new 4-partite 4-graphs

In the remaining parts of A, B and C we remove another such set of seven disjoint balanced complete 4-partite 4-graphs that are disjoint from the previous ones. Again by definition of connectedness and Lemma 31 we can continue this process until we remove $\eta m/8$ vertices each from B_i and C_i, $1 \leq i \leq 3$. By construction, if the number of vertices used from B_i and C_i, $1 \leq i \leq 3$ is t $(= \eta m/8)$ then the number of vertices used in A_2, A_3 and A_4 is $5t/6$, while that in A_1 is $t/2$ (see Figure 4.4).

Note that the new 4-partite 4-graphs use at least $3t \geq 3\eta m/8$ vertices from Z, but the remaining parts of A, B and C are not balanced (A_1, B_4 and C_4 have more vertices). To restore the balance in the remaining part of A we discard some arbitrary $t/3$ vertices from the remaining part of A_1. Similarly we discard some arbitrary t vertices from B_4 and C_4 to restore the balance in the remaining part of B and C. Therefore the new 4-partite 4-graphs together with leftover parts of A, B and C (after discarding the vertices) have at least $4(|A_1|+|B_1|+|C_1|)+3t-t/3-2t \geq 12m+\eta m/12$ vertices while all the 4-partite 4-graphs are balanced. Hence we extended (A, B, C).

On the other hand if the H_{432} in L_{abc} is as in Figure 4.2(b), then let the color classes corresponding to the neighbors of degree 3 and degree 2 pairs in H_{432} be $\{A_1, A_2, A_3\}$ and $\{A_3, A_4\}$ respectively. In this case extend (A, B, C) as follows.

For the pair (B_3, C_3) we remove two balanced complete 4-partite 4-graphs with the fourth color classes in A_3 and A_4 respectively. For the pair B_2, C_2 we remove two balanced complete 4-partite 4-graphs with the fourth color classes in A_1 and A_2 respectively. The size of of each color class in all of these new 4-partite graphs is $\beta\sqrt{\log m}$. Since the pair (B_1, C_1) has degree 4, we remove four balanced complete 4-partite 4-graphs with the fourth color class in A_1, A_2, A_3 and A_4 respectively.

Similarly as in the previous case we repeat this process so that we remove at least

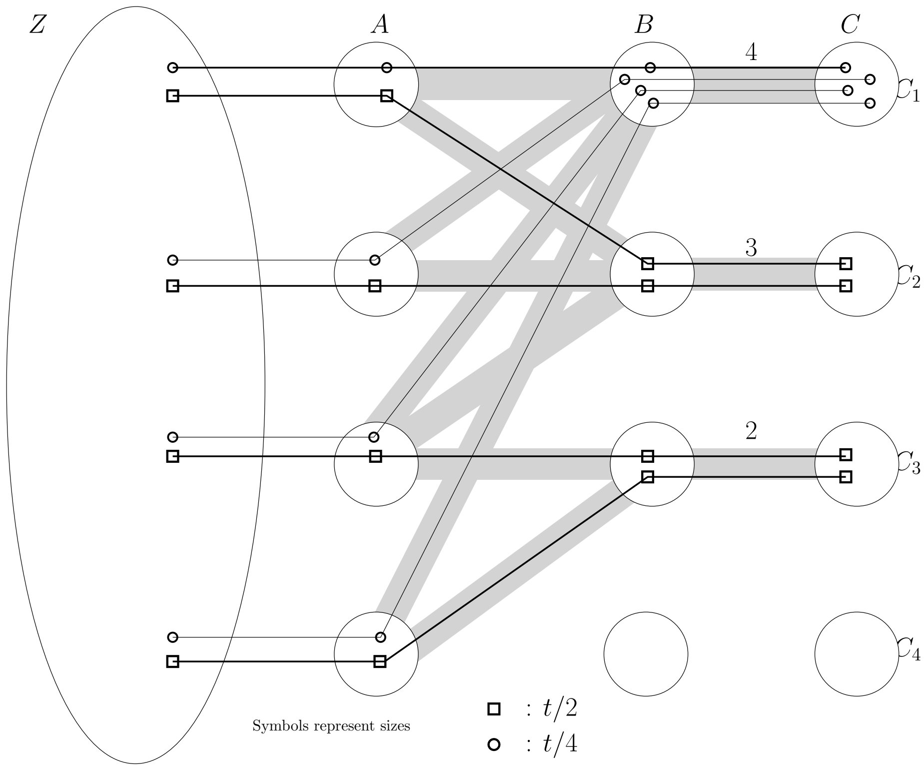


Figure 4.5: Extending (A, B, C) when L_{abc} has an H_{432} as in Figure 4.2(b): The shaded boxes represent a triplet *connected* to Z, while solid thick lines represent a balanced complete 4 partite 4-graphs

$t \geq \eta m/8$ vertices from each B_i and C_i, $1 \leq i \leq 3$. Note that by construction we have used $3t/4$ vertices in each color class of A (see Figure 4.5). So the remaining part of A is still balanced. While to restore balance in the remaining parts of B and C, we discard some arbitrary t vertices from each of B_4 and C_4. Again in total we added $3t$ vertices from Z, while we discarded $2t$ vertices form B_4 and C_4. Therefore the net increase in the number of vertices in the new set of complete 4-partite 4-graphs is $t \geq \eta m/8$, while all the 4-partite 4-graphs are balanced.

Case 3: L_{abc} has a subgraph isomorphic to H_{4221}:

Without loss of generality, assume that the pairs corresponding to the degree 4, 2, 2 and 1 pairs in this H_{4221} are (B_1, C_1), (B_2, C_2), (B_3, C_3) and (B_4, C_4) respectively. Furthermore let the color classes corresponding to the neighbors of degree 1 and the two degree 2 pairs in H_{4221} be $\{A_4\}$, $\{A_3, A_4\}$ and $\{A_3, A_4\}$ respectively (as in Figure 4.3(a)). By the definition of connectedness and Lemma 31 we build complete balanced 4-partite 4-graphs using (A_4, B_4, C_4, Z) and (A_3, B_3, C_3, Z) of size $\beta\sqrt{\log m}$. For the pair

(B_2, C_2) we make two more balanced complete 4-partite 4-graphs using (A_4, B_2, C_2, Z) and (A_3, B_2, C_2, Z). For the degree 4 pair (B_1, C_1) we remove balanced complete 4-partite 4-graphs in (A_1, B_1, C_1, Z) and (A_2, B_1, C_1, Z). Again we repeat this process

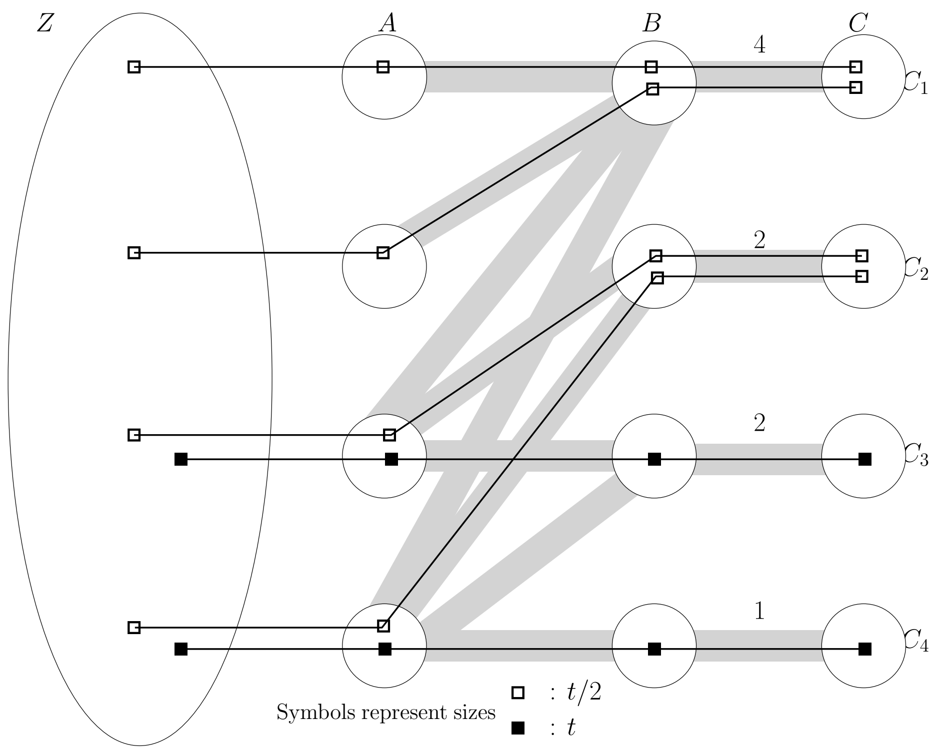


Figure 4.6: Extending (A, B, C) when L_{abc} has an H_{4221}: The shaded boxes represent a triplet *connected* to Z, while solid thick lines represent a balanced complete 4 partite 4-graphs

so that we remove $t \geq \eta m/24$ vertices from each color class of B and C. Note that with this process we have used $3t/2$ vertices in A_3 and A_4 while $t/2$ vertices each in A_1 and A_2 (see Figure 4.6). Furthermore only the remaining part of A is not balanced. The balance can be restored by discarding t vertices each from the remaining part of A_1 and A_2 which results in the net increase of $2t \geq \eta m/12$ vertices in all the balanced complete tripartite graphs.

Case 4: L_{abc} has a subgraph isomorphic to H_{3321}:

Assume that the pairs corresponding to the degree 3, 3, 2 and 1 pairs in the H_{3321} are (B_1, C_1), (B_2, C_2), (B_3, C_3) and (B_4, C_4) respectively. Let the color classes corresponding to the neighbors of these pairs in H_{3321} be $\{A_2, A_3, A_4\}$, $\{A_2, A_3, A_4\}$, $\{A_3, A_4\}$ and $\{A_4\}$. By the definition of connectedness and Lemma 31 we build three complete balanced 4-partite graphs in each of (A_4, B_4, C_4, Z), (A_3, B_3, C_3, Z) and (A_2, B_2, C_2, Z) of size $\beta\sqrt{\log m}$. In addition we build three more complete balanced 4-partite 4-graphs using (B_1, C_1) and Z, while the vertices of the fourth color classes are in A_2, A_3 and A_4 respectively. We repeat this process so as to remove at least $t \geq 3\eta m/64$ vertices each from each color class of B and C. Clearly the remaining part of B and C are still balanced, while in A we have used $4t/3$ vertices in each of A_2, A_3 and A_4. To restore the balance in remaining part of A we discard arbitrary $4t/3$ vertices from A_1. In the process the net increase in the number of vertices in the resultant balanced 4-partite 4-graphs is at least $8t/3 \geq 3\eta m/8$, hence (A, B, C) is *extended*. □

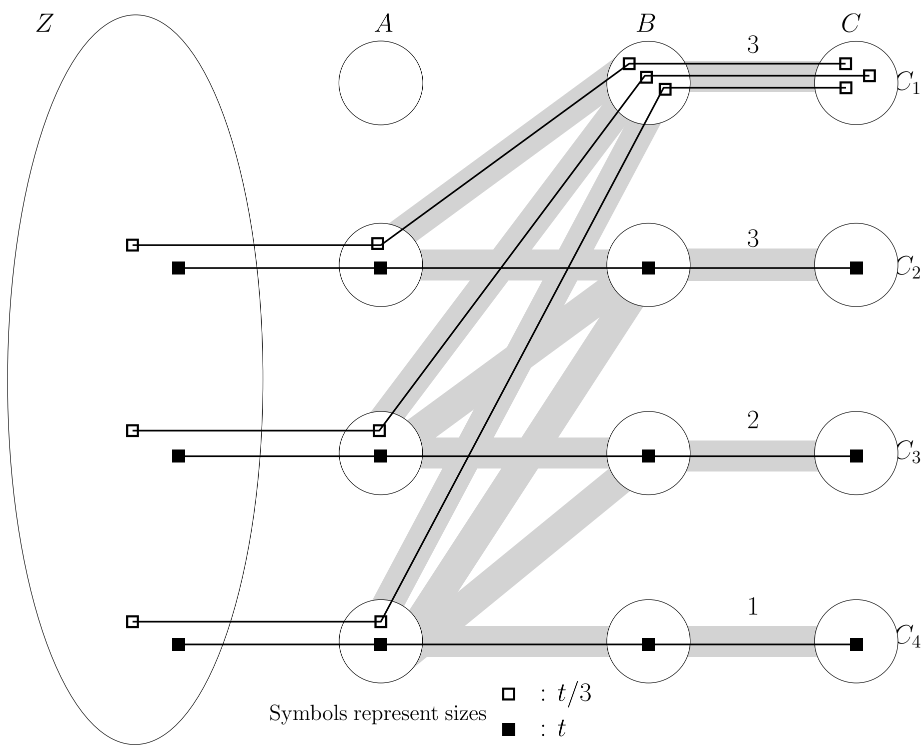


Figure 4.7: Extending (A, B, C) when L_{abc} has an H_{3321}: The shaded boxes represent a triplet *connected* to Z, while solid thick lines represent a balanced complete 4 partite 4-graphs

4.4 The Non Extremal Case

Throughout this section we assume that we have a 4-graph H satisfying (4.1) such that the extremal case does not hold for H. We shall assume that n is sufficiently large and besides our main parameter γ we use the parameters β and α such that the following holds

$$\frac{12}{2^{1/\beta^2}} < \gamma = \alpha^4 \ll 1 \tag{4.2}$$

where $a \ll b$ means that a is sufficiently small compared to b. From (4.1) and (4.2), when n is large we have

$$\delta_1(H) \geq \binom{n-1}{3} - \binom{3n/4}{3} + 1 > \frac{37}{64}\binom{n-1}{3} - \frac{9n^2}{64} > \left(1/2 + 2\sqrt{\alpha}\right)\binom{n}{3}$$

Hence our hypergraph H satisfies the conditions of Lemma 34 (the absorbing lemma). We remove from H an *absorbing matching* M of size at most $\alpha^2 n/4 = \sqrt{\gamma}n/4$. In the remaining hypergraph we find an almost perfect matching that leaves out a set of at most $\alpha^4 n = \gamma n$ vertices. As guaranteed by Lemma 34 the vertices that are left out from the almost perfect matching can be absorbed into M, therefore we get a perfect matching in H. In what follows we work with the remaining hypergraph (after removing $V(M)$). For simplicity we still denote the remaining hypergraph by H and assume that it is on n vertices. Since $|V(M)| \leq \sqrt{\gamma}n$, in the remaining hypergraph we still have

$$\delta_1(H) \geq \left(\frac{37}{64} - 6\sqrt{\gamma}\right)\binom{n}{3} \tag{4.3}$$

as for any vertex v in the remaining hypergraph, there can be at most $6\sqrt{\gamma}\binom{n}{3}$ edges containing v and at least one vertex in $V(M)$.

4.4.1 The optimal cover

Our goal is to find an almost perfect matching in H. We are going to build a cover $\mathcal{F} = \{Q_1, Q_2, \ldots\}$, such that, each Q_i is a disjoint balanced complete 4-partite 4-graph in H (we refer to them as 4-partite graphs). We say that such a cover is *optimal* if it covers at least $(1-\gamma)n$ vertices. We will show that either we can find an optimal cover or H is α-extremal. It is easy to see that such an optimal cover readily gives us an almost perfect matching.

Using the following iterative procedure, we either build an optimal cover or find a subset of vertices, which shows that H is α-extremal. We begin with a cover $\mathcal{F}_0$. Then in each step $t \geq 1$, if $\mathcal{F}_{t-1}$ is not optimal, we find another cover $\mathcal{F}_t$, such that

$|V(\mathcal{F}_t)| \geq |V(\mathcal{F}_{t-1})| + \gamma^2 n/16$ (for this we use the notation, $\mathcal{F}_t > \mathcal{F}_{t-1}$). The size of a color class in each 4-partite graph in $\mathcal{F}_t$ is m_t.

To get the initial cover, $\mathcal{F}_0$, we repeatedly apply Lemma 28 in the leftover of H, while the conditions of the lemma are satisfied and the number of leftover vertices are at least γn, to find 4-partite graphs, $K^{(4)}(m_0)$, where $m_0 = \beta(\log n)^{1/3}$. After the t^{th} step in this iterative procedure, if $\mathcal{F}_t$ is not an optimal cover, then we get $\mathcal{F}_{t+1}$. We will show that, unless H is α-extremal, we have $\mathcal{F}_{t+1} > \mathcal{F}_t$ and $m_{t+1} = \beta\sqrt{\log m_t}$. Let $\mathcal{I}_t = V(H) \setminus V(\mathcal{F}_t)$. Since $\mathcal{F}_t$ is not optimal and we cannot apply Lemma 28 in $H|_{\mathcal{I}_t}$, we must have that $|\mathcal{I}_t| \geq \gamma n$ and

$$d_4(\mathcal{I}_t) < \gamma. \tag{4.4}$$

By non-extremality of H, this implies that $|V(\mathcal{F}_0)| \geq n/4$.

In what follows we will show that if there are 'many' edges with three vertices in $\mathcal{I}_t$ and one vertex in some $Q_i \in \mathcal{F}_t$ then we get $\mathcal{F}_{t+1} > \mathcal{F}_t$. To that end let $Q_i = (V_1^i, V_2^i, V_3^i, V_4^i)$ be a 4-partite graph in $\mathcal{F}_t$, we say that a color class, V_p^i of Q_i, $(1 \leq p \leq 4)$ is *connected* to $\mathcal{I}_t$, if $d_4(V_p^i, \binom{\mathcal{I}_t}{3}) \geq 2\gamma$. We will show that if a 4γ-fraction of the 4-partite graphs in $\mathcal{F}_t$ have at least 2 color classes connected to $\mathcal{I}_t$, then we can we get $\mathcal{F}_{t+1} > \mathcal{F}_t$. To see this assume that we have a subcover $\mathcal{F}' \subset \mathcal{F}_t$, such that each $Q_i \in \mathcal{F}'$ has at least 2 color classes connected to $\mathcal{I}_t$ and $|V(\mathcal{F}')| \geq \gamma n$. Let $\mathcal{F}' = \{Q_1, Q_2, \ldots\} \subset \mathcal{F}_t$ and without loss of generality, say V_1^i and V_2^i are the color classes in each Q_i that are connected to $\mathcal{I}_t$. For each such Q_i, since $|V_1^i| = |V_2^i| = m_t \leq \beta(\log n)^{1/3}$ and $|\mathcal{I}_t| \geq \gamma n$, by Lemma 29 we can find two disjoint balanced complete 4-partite graphs $(U_1^i, A_1^i, B_1^i, C_1^i)$ and $(U_2^i, A_2^i, B_2^i, C_2^i)$ where U_1^i and U_2^i are subsets of V_1^i and V_2^i respectively, and A_k^i, B_k^i and C_k^i, $k \in \{1, 2\}$ are disjoint subsets of $\mathcal{I}_t$. The size of each color class of these new 4-partite graphs is at least $\gamma m_t/4$ (see Figure 4.8).

We remove the vertices of these new 4-partite graphs from their respective sets and add them to $\mathcal{F}_{t+1}$. Removing these vertices from V_1^i and V_2^i creates an imbalance in the leftover part of Q_i (V_3^i and V_4^i have more vertices). To restore the balance, we discard (add to $\mathcal{I}_t$) some arbitrary $|U_1^i| = |U_2^i|$ vertices each from V_3^i and V_4^i. The new 4-partite graphs use $6\gamma m_t/4$ vertices from $\mathcal{I}_t$. Therefore even after discarding the vertices from

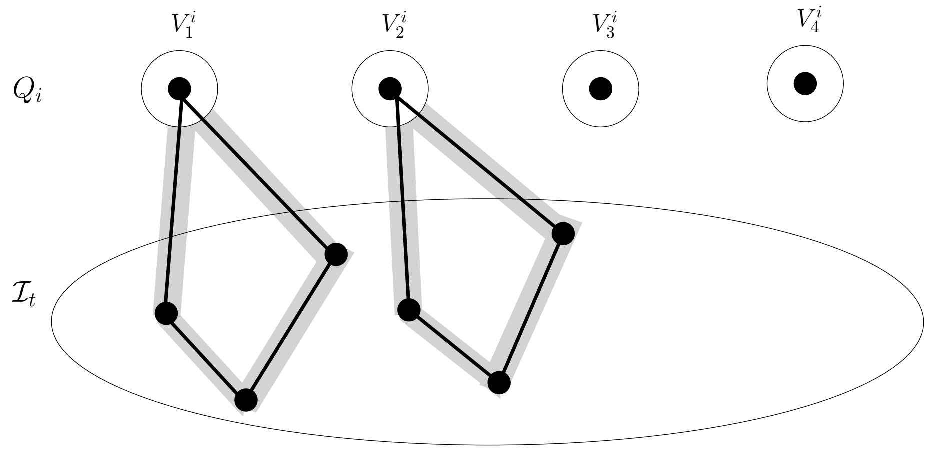


Figure 4.8: V_1^i and V_2^i of Q_i are connected to $\mathcal{I}_t$. The Shaded quadrilaterals represent density, while the dark black quadrilateral is balanced complete 4-partite 4-graph

V_3^i and V_4^i the net increase in the size of our cover is γm_t, while all the 4-partite graphs are balanced. We repeat this procedure for every $Q_i \in \mathcal{F}'$ and add the leftover part of Q_i and all 4-partite graphs in $\mathcal{F}_t \setminus \mathcal{F}'$, to $\mathcal{F}_{t+1}$. Now, we split each 4-partite graph in $\mathcal{F}_{t+1}$ into disjoint balanced complete 4-partite graphs, such that each has a color class of size $m_{t+1} = \beta\sqrt{\log m_t}$ (we assume divisibility). Since $|V(\mathcal{F}')| \geq \gamma n$, by the above observation, we have $|V(\mathcal{F}_{t+1})| \geq |V(\mathcal{F}_t)| + \gamma^2 n$, hence $\mathcal{F}_{t+1} > \mathcal{F}_t$.

Similarly, if there are 'many' edges that uses two vertices in $\mathcal{I}_t$ and two vertices from $V(\mathcal{F}_t)$ then we get $\mathcal{F}_{t+1} > \mathcal{F}_t$. First note that the number of pairs of vertices of any 4-partite graph $Q_i \in \mathcal{F}_t$ is $O(\log n)^{2/3}$, so the number of pairs of vertices within the 4-partite graphs in $\mathcal{F}_t$ is $O(n(\log n)^{2/3}) = o\binom{n}{2}$. Therefore the total number of edges, containing two vertices within a $Q_i \in \mathcal{F}_t$ and two vertices in $\mathcal{I}_t$ is $o\binom{n}{4}$, hence we ignore such edges. Let $Q_i = (V_1^i, V_2^i, V_3^i, V_4^i)$ and $Q_j = (V_1^j, V_2^j, V_3^j, V_4^j)$ be a pair of 4-partite graphs in $\mathcal{F}_t$ we say that $\mathcal{I}_t$ is *connected* to a pair of color classes (V_p^i, V_q^j), $(1 \leq p, q \leq 4)$, if $d_4\left(V_p^i, V_q^j, \binom{\mathcal{I}_t}{2}\right) \geq 2\gamma$. We say that $\mathcal{I}_t$ is *k-sided* to a pair $(Q_i, Q_j) \in \binom{\mathcal{F}_t}{2}$ if $\mathcal{I}_t$ is connected to k-pairs in $\{V_1^i, V_2^i, V_3^i, V_4^i\} \times \{V_1^j, V_2^j, V_3^j, V_4^j\}$. Assume that $\mathcal{I}_t$ is at least 9-sided to a 4γ-fraction of pairs of 4-partite graphs in $\binom{\mathcal{F}_t}{2}$. By a simple greedy procedure, (Lemma 32 and the 2-graph analog of Lemma 33) we get a disjoint set of pairs, $M' \subset \binom{\mathcal{F}_t}{2}$, such that for each pair $(Q_i, Q_j) \in M'$, $\mathcal{I}_t$ is at least 9-sided to (Q_i, Q_j) and the number of vertices covered by M' is at least γn.

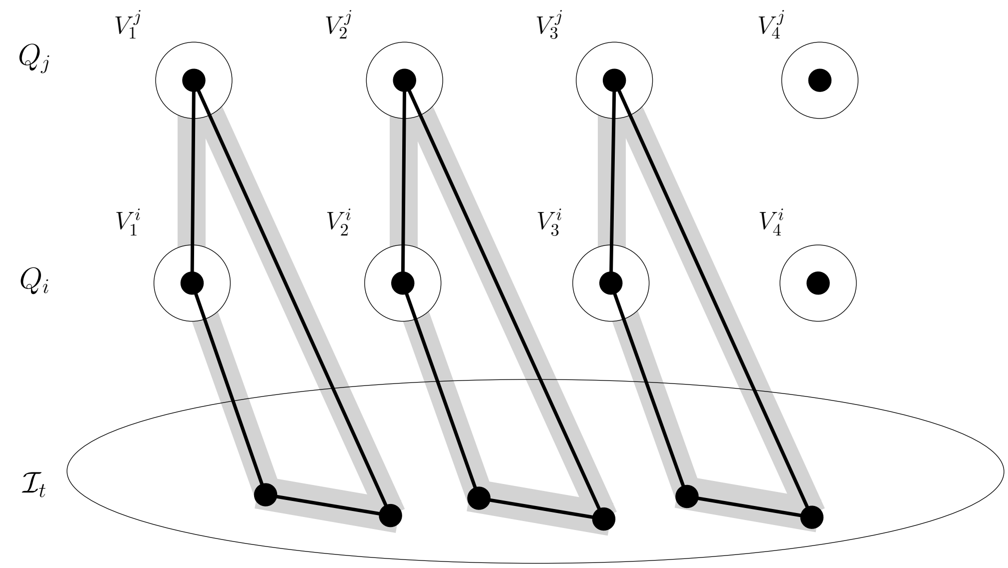


Figure 4.9: The pair (Q_i, Q_j) is at least 9-sided to $\mathcal{I}_t$. The Shaded quadrilaterals represent density, while the dark black quadrilateral is balanced complete 4-partite 4-graph

For every $(Q_i, Q_j) \in M'$ we proceed as follows: Since $\mathcal{I}_t$ is connected to at least 9 pairs of color classes in $\{V_1^i, V_2^i, V_3^i, V_4^i\} \times \{V_1^j, V_2^j, V_3^j, V_4^j\}$, it is easy to see that we can find 3 disjoint pairs of color classes such that $\mathcal{I}_t$ is connected to each of them (say $(V_1^i, V_1^j), (V_2^i, V_2^j)$ and (V_3^i, V_3^j) are such disjoint pairs of color classes). We have $|V_1^i| = |V_1^j| = m_t \leq \beta(\log n)^{1/3}$ and $|\mathcal{I}_t| \geq \gamma n$, so by definition of connectedness, the induced hypergraph, $H\left(V_1^i, V_1^j, \binom{\mathcal{I}_t}{2}\right)$ satisfies the conditions of Lemma 30. Therefore by Lemma 30 we remove balanced complete 4-partite 4-graphs, such that each of them has one color class in V_1^i, one in V_1^j and two color classes in $\mathcal{I}_t$. Note that the conditions of Lemma 30 are satisfied until in total we remove at least $\gamma m_t/2$ vertices each from V_1^i and V_1^j. We repeat the same process with (V_2^i, V_2^j) and (V_3^i, V_3^j) (see Figure 4.9). In total these new balanced complete 4-partite 4-graphs use $3\gamma m_t$ vertices from $\mathcal{I}_t$ and when we remove the vertices of the new 4-partite graphs, from Q_i and Q_j the remaining parts of Q_i and Q_j are not balanced (V_4^i and V_4^j have extra vertices). To restore the balance we remove an arbitrary set of vertices each from V_4^i and V_4^j (equal to the difference in color classes). Still the net increase in the number of vertices in the cover is at least $2\gamma m_t$.

Applying this procedure for every $(Q_i, Q_j) \in M'$ and making all of the 4-partite graphs of the same size, by splitting, we get $\mathcal{F}_{t+1}$. Again we have all 4-partite graphs balanced with color classes of size $m_{t+1} = \beta\sqrt{\log m_t}$ and $\mathcal{F}_{t+1} > \mathcal{F}_t$.

Now, if we can not get $\mathcal{F}_{t+1} > \mathcal{F}_t$, we must have that, in almost all 4-partite graphs in $\mathcal{F}_t$, at most one color class is connected to $\mathcal{I}_t$ and almost all pairs of 4-partite graphs in $\binom{\mathcal{F}_t}{2}$ are at most 8-sided. In particular this implies that for a typical vertex $v \in \mathcal{I}_t$ we have

$$deg_4\left(v, \binom{\mathcal{I}_t}{2} \times V\left(\mathcal{F}_t\right)\right) \leq \left(\frac{1}{4} + 6\gamma\right) |V\left(\mathcal{F}_t\right)| \binom{|\mathcal{I}_t|}{2}$$

and

$$deg_4\left(v, \mathcal{I}_t \times \binom{V\left(\mathcal{F}_t\right)}{2}\right) \leq \left(\frac{1}{2} + 16\gamma\right) |\mathcal{I}_t| \binom{|V\left(\mathcal{F}_t\right)|}{2}.$$

From (4.3), (4.4) and the above degree bounds, for a typical vertex $v \in \mathcal{I}_t$ we have

$$\begin{aligned}
& deg_4\left(v, \binom{V(\mathcal{F}_t)}{3}\right) \\
\geq & \left(\frac{37}{64} - 6\sqrt{\gamma}\right)\binom{n}{3} - d_4\left(v, \mathcal{I}_t \times \binom{V(\mathcal{F}_t)}{2}\right) - d_4\left(v, \binom{\mathcal{I}_t}{2} \times V(\mathcal{F}_t)\right) - d_4\left(v, \binom{\mathcal{I}_t}{3}\right) \\
\geq & \left(\frac{37}{64} - 6\sqrt{\gamma}\right)\binom{n}{3} - \left(\frac{1}{2} + 16\gamma\right)|\mathcal{I}_t|\binom{|V(\mathcal{F}_t)|}{2} - \left(\frac{1}{4} + 6\gamma\right)\binom{|\mathcal{I}_t|}{2}|V(\mathcal{F}_t)| - \gamma\binom{|\mathcal{I}_t|}{3} \\
\geq & \left(\frac{37}{64} - 30\sqrt{\gamma}\right)\binom{|V(\mathcal{F}_t)|}{3}
\end{aligned}$$

where the last inequality holds when $|\mathcal{I}_t| \geq \gamma n$ and $|V(\mathcal{F}_t)| \geq n/4$.

For a vertex v, consider the edges that v makes with 3-sets of vertices within a $Q_i \in \mathcal{F}_t$. The number of 3-sets of vertices of any $Q_i \in \mathcal{F}_\sqcup$ is $O(\log n)$, (as the size of Q_i is at most $\beta \log^{1/3} n$), hence the total number of 3-sets of vertices within the 4-partite graphs in $\mathcal{F}_t$ is $O(n \log n) = o\binom{n}{3}$. Similarly the number of 3-sets of vertices which uses 2 vertices from a Q_i and one vertex from some other Q_j is $O(n^2 \log n) = o\binom{n}{3}$. Therefore, for any vertex v, we ignore these types of edges and we will only consider the edges that v makes with 3-sets of vertices $(x, y, z),\ x \in V(Q_i),\ y \in V(Q_j),\ z \in V(Q_k),\ i \neq j \neq k$. By the above observation, for the minimum degree of a typical vertex $v \in \mathcal{I}_t$ we still have

$$deg_4\left(v, \binom{V(\mathcal{F}_t)}{3}\right) \geq \left(\frac{37}{64} - 40\sqrt{\gamma}\right)\binom{|V(\mathcal{F}_t)|}{3} \tag{4.5}$$

Let $Q_i = (V_1^i, \ldots, V_4^i)$, $Q_j = (V_1^j, \ldots, V_4^j)$ and $Q_k = (V_1^k, \ldots, V_4^k)$ be three 4-partite graphs in $\mathcal{F}_t$, we say that $\mathcal{I}_t$ is *connected* to a triplet of color classes (V_p^i, V_q^j, V_r^k),

$1 \leq p, q, r \leq 4$, if $d_4(\mathcal{I}_t, (V_p^i \times V_q^j \times V_r^k)) \geq 2\gamma$. For $(Q_i, Q_j, Q_k) \in \binom{\mathcal{F}_t}{3}$ we consider the link graph L_{ijk} as defined above (with $\mathcal{I}_t$ playing the role of the set Z).

For a constant $\eta > 0$, we say that $\mathcal{I}_t$ is (η, s)-*connected* to $\mathcal{F}_t$, if there is a subset of triplets of 4-partite graphs, $T \subset \binom{\mathcal{F}_t}{3}$, such that for each triplet $(Q_i, Q_j, Q_k) \in T$, the link graph, L_{ijk} has s edges and $|T| \geq \eta\binom{|\mathcal{F}_t|}{3}$.

A simple calculation, using (4.5), implies that if $\mathcal{I}_t$ is $(\gamma^{1/3}, s)$-*connected* to $\mathcal{F}_t$ for some $s \leq 36$, then we also have that $\mathcal{I}_t$ is $(\sqrt{\gamma}, \geq 38)$-*connected* to $\mathcal{F}_t$.

We consider the following cases based on the way $\mathcal{I}_t$ is connected to $\mathcal{F}_t$ and show that either we get $\mathcal{F}_{t+1} > \mathcal{F}_t$ or H is extremal. First assume that $\mathcal{I}_t$ is $(32\gamma, \geq 37)$-*connected* to $\mathcal{F}_t$ such that for every triplet $(Q_i, Q_j, Q_k) \in T$, the link graph L_{ijk} is not isomorphic to H_{ext}. Then by lemma 32 and lemma 33 there exists a set a disjoint set of triplets of 4-partite graphs, $T' \subset T$, such that for each triplet $(Q_i, Q_j, Q_k) \in T'$ the link graph L_{ijk} has at least 37 edges and is not isomorphic to H_{ext}. Furthermore, the number of vertices covered by T' is at least γn.

Now, for each $(Q_i, Q_j, Q_k) \in T'$, since L_{ijk} has at least 37 edges and $L_{ijk} \neq H_{ext}$, using Lemma 36 we extend (Q_i, Q_j, Q_k) to add at least $\gamma m_t/16$ vertices to our cover. Clearly if we extend every triplet in T' the net increase in the size of our cover is at least $\gamma^2 n/16$ (as the size of T' is at least γn). Similarly as above we can split the 4-partite graphs to make them of the same size and get $\mathcal{F}_{t+1} > \mathcal{F}_t$.

On the other hand, if there is no such T, then we must have that $\mathcal{I}_t$ is not $(\gamma^{1/3}, s)$-*connected* to $\mathcal{F}_t$ for any $s \leq 36$, because otherwise, as observed above, we will get such a T. So, roughly speaking, for almost all triplets of 4-partite graphs in $\binom{\mathcal{F}_t}{3}$, we have that the link graph of the triplet has exactly 37 edges and is isomorphic to H_{ext}. Call a 4-partite graph $Q_i \in \mathcal{F}_t$ *good*, if for almost all pairs of other 4-partite graphs Q_j, Q_k, we have that the link graph L_{ijk} is isomorphic to H_{ext}.

By the above observation, almost all 4-partite graphs (covering $\geq (1 - 2\gamma^{1/3})|V(\mathcal{F}_t)|$ vertices) are good. By a simple greedy procedure we find a set of disjoint triplets of

4-partite graphs, T_g, such that for each triplet $(Q_i, Q_j, Q_k) \in T_g$, the link graph L_{ijk} is isomorphic to H_{ext} and all good 4-partite graphs are part of some triplet in T_g. Let the set of 4-partite graphs covered by T_g be $\mathcal{F}_g$, clearly $|V(\mathcal{F}_g)| \geq (1-2\gamma^{1/3})|V(\mathcal{F}_t)|$. With relabeling we may also assume that in each triplet $(Q_i, Q_j, Q_k) \in T_g$, V_1^i, V_1^j and V_1^k are the color classes corresponding to the vertices of the link graph L_{ijk} that intersect every edge of L_{ijk}.

For every $(Q_i, Q_j, Q_k) \in T_g$, by definition of connectedness and the sizes of the Q_i, Q_j and Q_k, the 4-partite hypergraph induced by $(V_1^i, V_2^j, V_2^k, \mathcal{I}_t)$ satisfies the conditions of Lemma 31. Hence applying Lemma 31 we find a balanced complete 4-partite graphs X_{i1} in $H(V_1^i, V_2^j, V_2^k, \mathcal{I}_t)$. We also find two more disjoint balanced complete 4-partite graphs, X_{i2} and X_{i3} in $H(V_1^i, V_3^j, V_3^k, \mathcal{I}_t)$ and $H(V_1^i, V_4^j, V_4^k, \mathcal{I}_t)$. Since $|\mathcal{I}_t| \geq \gamma n$, we can find these complete 4-partite graphs that are disjoint from each other. The sizes of a color class in each of X_{11}, X_{12} and X_{13} is $\beta\sqrt{\log m_t}$.

Similarly for each of V_1^j and V_1^k we find 3 disjoint balanced complete 4-partite graphs X_{jp} and X_{kp} in $H(V_p^i, V_1^j, V_p^k, \mathcal{I}_t)$ and $H(V_p^i, V_p^j, V_1^k, \mathcal{I}_t)$, $(2 \leq p \leq 4)$, respectively. All of these these balanced complete 4-partite graphs are disjoint from each other and the size of a color class in each one of them is $\beta\sqrt{\log m_t}$.

By the definition of connectedness, the structure of the link graph L_{ijk} and the fact that $|\mathcal{I}_t| \geq \gamma n$, clearly we can find these nine disjoint 4-partite graphs. And as argued above we repeat this process (remove another set of 9 such 4-partite graphs) until in total we remove $\gamma m_t/2$ vertices from each of V_1^i, V_1^j and V_1^k, while $\gamma m_t/3$ vertices from each of the other classes in Q_i, Q_j and Q_k.

Note that these tripartite graphs in total use $3\gamma m_t/2$ vertices from $\mathcal{I}_t$. But this creates an imbalance among the color classes of the remaining parts of Q_i, Q_j and Q_k (V_1^i, V_1^j and V_1^k have fewer vertices), to restore the balance we will have to discard $\gamma m_t/6$ vertices from each color class in Q_i, Q_j and Q_k except V_1^i, V_1^j and V_1^k. Which leaves us with no net gain in the size of the cover. Therefore we will not discard any vertices from these color classes at this time and say that these color classes have extra vertices. We proceed in similar manner for each triplet in T_g. So we have about $\gamma n/24$ extra vertices altogether.

Denote by V_1^g, V_2^g, V_3^g and V_4^g the union of the corresponding color classes of remaining parts of 4-partite graphs in $\mathcal{F}_g$. Clearly $|V_2^g| = |V_3^g| = |V_4^g| \geq (1-2\gamma^{1/3})|V(\mathcal{F}_t)|/4 - \gamma n/24 \geq (1/16 - 3\gamma^{1/3})n$. The last inequality follows from the lower bound on size of $V(\mathcal{F}_t)$ above, and the fact that γ is a small constant. We will show that either we can increase the size of our cover or we have

$$d_4\left(V_2^g \cup V_3^g \cup V_4^g\right) \leq \sqrt{\gamma}. \tag{4.6}$$

For $d_4(V_2^g \cup V_3^g \cup V_4^g)$, we only consider those edges that use exactly one vertex from a 4-partite graph Q_i, as the number of edges of other types is $o(n^4)$. Assume that $d_4(V_2^g \cup V_3^g \cup V_4^g) \geq \sqrt{\gamma}$ then by Lemma 28 there exist complete 4-partite graphs in $H|_{V_2^g \cup V_3^g \cup V_4^g}$ covering at least γn vertices. We remove some of these 4-partite graphs (possibly with splitting and discarding part of it) such that from no color class we remove more then the number of extra vertices in that color class. Adding these new 4-partite graphs to our cover increases the size of our cover by at least $\gamma^2 n$ vertices. As we will not need to discard vertices from $(V_2^g \cup V_3^g \cup V_4^g)$ for rebalancing. Instead the extra vertices are part of these new 4-partite graphs. In the remaining parts of V_2^g, V_3^g and V_4^g we arbitrarily discard some extra vertices to restore the balance in the 4-partite graphs.

Similarly we will show that either we can increase the size of our cover or we have

$$d_4\left(V_2^g \cup V_3^g \cup V_4^g, \binom{\mathcal{I}_t}{3}\right) \leq \sqrt{\gamma}. \tag{4.7}$$

Indeed assume the contrary, i.e. $d_4(V_2^g \cup V_3^g \cup V_4^g, \binom{\mathcal{I}_t}{3}) \geq \sqrt{\gamma}$, then since both $|\mathcal{I}_t|$ and $|V_2 \cup V_3^g \cup V_4^g|$ are at least γn, by Lemma 28 we can find disjoint complete 4-partite graphs with one color class in $V_2 \cup V_3^g \cup V_4^g$ and three in $\mathcal{I}_t$ covering at least $\gamma^2 n/2$ vertices. And again as above we can add these 4-partite graphs to our cover and increase the size of our cover as we have extra vertices in $V_2^g \cup V_3^g \cup V_4^g$. By the same reasoning we can prove that there are very few edges that uses two vertices from $V_2 \cup V_3^g \cup V_4^g$ and two from $\mathcal{I}_t$.

From the above observations about the size of $\mathcal{F}_g$ and (4.2) we have that $|V_2^g \cup V_3^g \cup V_4^g \cup \mathcal{I}_t| \geq (3/4 - \alpha)n$. Therefore if we can not increase the size of our cover significantly (by at least $\gamma^2 n/16$ vertices), then by (4.4), (4.6) and (4.7) we get that $d_4(V_2^g \cup V_3^g \cup V_4^g \cup \mathcal{I}_t) < \alpha$. Hence H is α-extremal.

4.5 The Extremal Case

Here our graph H is in the Extremal Case, i.e. *there is a* $B \subset V(H)$ *such that*

- $|B| \geq (\frac{3}{4} - \alpha)n$
- $d_4(B) < \alpha$.

We assume that n is sufficiently large and α is a sufficiently small constant < 1. Let $A = V(H) \setminus B$, by shifting some vertices between A and B we can have that $A = n/4$ and $B = 3n/4$ as $n \in 4\mathbb{Z}$ (we still keep the notation A and B). It is easy to see that we still have

$$d_4(B) < 6\alpha \tag{4.8}$$

Since we have

$$\delta_1(H) \geq \binom{n-1}{3} - \binom{3n/4}{3} + 1 = \binom{n-1}{3} - \binom{|B|}{3} + 1$$

this together with (4.8) implies that almost all 4-sets of $V(H)$ are edges of H except 4-sets of B. Thus roughly speaking we have that almost every vertex $b \in B$ makes edges with almost all 3-sets of vertices in $\binom{A}{3}$, with almost all 3-sets of vertices in $B \setminus \{b\} \times \binom{A}{2}$ and with almost all 3-sets of vertices in $\binom{B \setminus \{b\}}{2} \times A$ and vice versa. Therefore, we will basically match every vertex in A with a distinct 3-set of vertices in $\binom{B}{3}$ (disjoint from all 3-sets matched with other vertices in A) to get the perfect matching. However some vertices may be '*atypical*', in the sense that they may not have this connectivity structure hence we will first find a small matching that covers all such '*atypical*' vertices. For the remaining '*typical*' vertices we will show that they satisfy the conditions of König-Hall theorem, hence we will match every remaining vertex in A with a distinct 3-sets of remaining vertices in B.

A vertex $a \in A$ is called *exceptional* if it does not make edges with almost all 3-sets of vertices in B, more precisely if

$$deg_4\left(a, \binom{B}{3}\right) < (1-\sqrt{\alpha})\binom{|B|}{3}$$

A vertex $a \in A$ is called *strongly exceptional* if it makes edges with very few 3-sets in B, more precisely if

$$deg_4\left(a, \binom{B}{3}\right) < \alpha^{1/3}\binom{|B|}{3}$$

Similarly a vertex $b \in B$ is called *exceptional* if it makes edges with many 3-sets of vertices in B more precisely if

$$deg_4\left(b, \binom{B \setminus \{b\}}{3}\right) > \sqrt{\alpha}\binom{|B|}{3}$$

A vertex $b \in B$ is called *strongly exceptional* if it makes edges with almost all 3-sets of vertices in B more precisely if

$$deg_4\left(b, \binom{B \setminus \{b\}}{3}\right) > (1-\alpha^{1/3})\binom{|B|}{3}$$

Denote the set of *exceptional* and *strongly exceptional* vertices in A (and B) by X_A and SX_A respectively (similarly X_B and SX_B). Easy calculations using (4.1) and (4.8) yields that $|X_A| \leq 18\sqrt{\alpha}|A|$ and $|X_B| \leq 18\sqrt{\alpha}|B|$ and for the *strongly exceptional* sets we have $|SX_A| \leq 40\alpha|A|$ and $|SX_B| \leq 40\alpha|B|$. The constants are not the best possible but we choose them for ease of calculation.

If we have both SX_B and SX_A non empty, (say $b \in SX_B$ and $a \in SX_A$) then since then we can exchange a with b and reduce the size of both SX_B and SX_A, as it is easy to see that both a and b are not *strongly exceptional* in their new sets. Hence one of the sets SX_A and SX_B must be empty.

Assume $SX_B \neq \emptyset$. By definition of SX_B, for every vertex $b \in SX_B$, we have $deg_3(b, \binom{B}{3}) \geq (1-\alpha^{1/3})\binom{|B|}{3}$. This together with the bound on the size of SX_B implies that we can greedily find $|SX_B|$ vertex disjoint edges in $H|_B$ each containing exactly one vertex of SX_B. We also select $|SX_B|$ other vertex disjoint edges such that each edge has two vertices in $B \setminus X_B$ and the two other vertices are in A. We can clearly find such edges because by (4.1) and definition of X_B every vertex in $B \setminus X_B$ makes edges with at least $(1-3\sqrt{\alpha})$-fraction of 3-sets in $B \times \binom{|A|}{2}$. We remove the vertices of these edges from A and B and denote the remaining set by A' and B'. Let $|A'|+|B'| = n'$, by the above procedure we have $n' = n - 8|SX_B|$, $|A'| = |A| - 2|SX_B|$ and $|B'| = |B| - 6|SX_B|$ hence we get $|B'| = 3|A'| = 3n'/4$.

In case $SX_A \neq \emptyset$ (and $SX_B = \emptyset$), we will first eliminate the vertices in SX_A. Note that in this case any vertex $b \in B$ is exchangeable with any vertex in SX_A, because if there is a vertex $b \in B$ such that $deg_4(b, \binom{B}{3}) \geq \alpha^{1/3}\binom{|B|}{3}$ then we can replace b with any vertex $a \in SX_A$ to reduce the size of SX_A (as the vertex b is not *strongly exceptional* in A and a can not be *strongly exceptional* in the set B). Therefore we consider the whole set $SX_A \cup B$. By (4.1) for any vertex $v \in SX_A \cup B$ we have

$$\begin{aligned} deg_4\left(v, \binom{SX_A \cup B}{3}\right) \geq& (|SX_A|-1)\binom{|B|}{2} + \binom{(|SX_A|-1)}{2}|B| + \binom{(|SX_A|-1)}{3} + 1 \\ \geq& \binom{4(|SX_A|-1)}{3} + 1 \end{aligned}$$

where the last inequality holds when n is large enough and $|SX_A$ is small. So with a simple greedy procedure we find $|SX_A|$ disjoint edges in $H|_{SX_A \cup B}$ and remove these edges from H. Note that this is the only place where we critically use the minimum degree. We let $A' = A \setminus SX_A$ and B' has all other remaining vertices. Again as above we have $n' = n - 4|SX_A|$, $|A'| = |A| - |SX_A|$ and $|B'| = |B| - 3|SX_A|$ hence we get $|B'| = 2|A'| = 3n'/4$.

Having dealt with the *strongly exceptional* vertices, the vertices of X_A and X_B in A' and B' can be eliminated using the fact that their sizes are much smaller than the crossing degrees of vertices in those sets. For instance as observed above we have $|X_A| \leq 18\sqrt{\alpha}|A|$ while for any vertex $a \in X_A$, we have that $deg_4(a, \binom{B'}{2}) \geq \alpha^{1/3}\binom{|B'|}{3}/2$ (because $a \notin SX_A$). Therefore by a simple greedy procedure for each $a \in X_A$ we delete a disjoint edge that contains a and three vertices from B'. Similarly for each $b \in X_B$ we delete an edge that contains b and uses one vertex from A' and the other two vertices from B' distinct from b. Clearly we can find such disjoint edges, hence we removed a partial matching that covers all vertices in the *strongly exceptional* and *exceptional* sets.

Finally in the leftover sets of A' and B' (denote them by A'' and B'', by construction we still have $|B''| = 3|A''|$) we will find $|A''|$ disjoint edges each using one vertex in A'' and three vertices in B''. Note that for every vertex $a \in A''$ we have $deg_4(a, \binom{B''}{3}) \geq (1 - 2\alpha^{1/3})\binom{|B''|}{3}$ (as $a \notin X_A$). We say that a vertex b_i and a pair b_j, b_k in B'' are *good* for each other if $(b_i, b_j, b_k, a_l) \in E(H)$ for at least $(1 - 40\alpha^{1/4})|A''|$ vertices a_l in A''. We have that any vertex $b_i \in B''$ is *good* for at least $(1 - 40\alpha^{1/4})\binom{|B''|}{2}$ pairs of vertices in B'' (again this is so because $b_i \notin X_B$). We call such a (b_i, b_j, b_j) a good triplet.

We randomly select a set T_1 of $100\alpha^{1/4}|B''|$ vertex disjoint 3-sets of vertices in B''. By the above observation with high probability every vertex $a \in A''$ make edges in H with at least $3|T_1|/4$ triplets in T_1 and every triplet in T_1 makes an edge with at least $3|A''|/4$ vertices in A''. In $B'' \setminus V(T_1)$ still every vertex is *good* for almost all pairs (as the size of T_1 is very small).

We cover vertices in $B'' \setminus V(T_1)$ with disjoint good triplets (i.e. the triplet makes an edge in H with at least $(1-40\alpha^{1/4})|A''|$ vertices in A''. This can be done by considering a 3-graph with vertex set $B'' \setminus V(T_1)$ and all the *good* triplets as its edges. As argued above every vertex is good for almost all pairs. We can find a perfect matching in this 3-graph (see [36]). Let the set of triplets in this perfect matching be T_2.

Now construct an auxiliary bipartite graph $G(L, R)$, such that $L = A''$ and vertices in R corresponds to the triplets in T_1 and T_2. A vertex in $a_l \in L$ is connected to a vertex $y \in R$ if the triplet corresponding to y (say b_i, b_j, b_k) is such that $(b_i, b_j, b_k, a_l) \in E(H)$. We will show that $G(L, R)$ satisfies the König-Hall criteria. Considering the sizes of A'' and T_1 it is easy to see that for every subset $Q \subset R$ if $|Q| \leq (1 - 40\alpha^{1/4})|A''|$ then $|N(Q)| \geq |Q|$. When $|Q| > (1 - 40\alpha^{1/4})|A''|$ (using $|B''| = 2|A''|$) any such Q must have at least $6|T_1|/10$ vertices corresponding to pairs in T_1, hence with high probability $N(Q) = L \geq |Q|$. Therefore there is a perfect matching of R into L. This perfect matching in $G(L, R)$ readily gives us a matching in H covering all vertices in A'' and B'', which together with the edges we already removed (covering *strongly exceptional* and *exceptional* vertices) is a perfect matching in H. $\square$

References

[1] S. Abbasi, Spanning cycles in dense graphs, Ph.D. thesis, Rutgers University, 1999.

[2] M. Aigner, S. Brandt, Embedding arbitrary graphs of maximum degree two, *J. Lond. Math. Soc.*, Ser. II. 48 (1993), pp. 39-51.

[3] N. Alon, E. Fischer, 2-Factors in dense graphs, *Discrete Mathematics* 152 (1996), pp. 13-23.

[4] N. Sauer, J. Spencer, Edge disjoint placements of graphs, *J. Combin. Theory Ser. B* 25 (1978), pp. 295-302

[5] N. Alon, R. Yuster, H-factors in dense graphs, *J. Comb. Theory, Ser. B* 66 (1996), pp. 269-282.

[6] B. Bollobás, *Extremal Graph Theory*, Academic Press, London (1978).

[7] B. Bollobás, P. Erdős, M. Simonovits, On the structure of edge graphs II, *J. London Math. Soc.*, 12 (2) (1976), pp. 219-224.

[8] K. Corrádi, A. Hajnal, On the maximal number of independent circuits in a graph, *Acta Math. Acad. Sci. Hung.* 14 (1963), pp. 423-439.

[9] M.H. El-Zahar, On circuits in graphs, *Discrete Mathematics* 50 (1984), pp. 227-230.

[10] P. Erdős, R. Faudree, Some recent combinatorial problems, Technical Report, University of Bielefeld, 1990.

[11] P. Erdős, L. Pósa, On the maximal number of disjoint circuits of a graph, *Publ. Math. Debrecen* 9 (1962), pp. 3-12.

[12] A. Gyárfás, G.N. Sárközy, E. Szemerédi, Stability of the path-path Ramsey number, Discrete Mathematics 309 (2009), pp. 4590-4595.

[13] R. Johansson, An El-Zahar type condition ensuring path-factors, *J. Graph Theory* 28 (1998), no. 1, pp. 39-42.

[14] R. Johansson, On the bipartite case of El-Zahar conjecture, *Discrete Mathematics* 219 (2000), pp. 23-134.

[15] J. Komlós, G. N. Sárközy and E. Szemerédi, Blow-up Lemma, *Combinatorica* 17 (1997), pp. 109-123.

[16] J. Komlós, G. N. Sárközy, E. Szemerédi, Proof of the Alon-Yuster conjecture, *Discrete Mathematics* 235 (2001), pp. 255-269.

[17] P. Kővári, V.T. Sós, P. Turán, On a problem of Zarankiewicz, *Colloq. Math.*, 3 (1954), pp. 50-57.

[18] B. Randerath, I. Schiermeyer, H. Wang, On quadrilaterals in a graph, *Discrete Mathematics* 203 (1999), pp. 229-237.

[19] E. Szemerédi, Regular partitions of graphs, Colloques Internationaux C.N.R.S. № 260 - Problèmes Combinatoires et Théorie des Graphes, Orsay (1976), pp. 399-401.

[20] H. Wang, Covering a graph with cycles, *J. Graph Theory* 20 (1995), pp. 202-211.

[21] H. Wang, Vertex-disjoint quadrilaterals in graphs, *Discrete Mathematics* 288 (2004), pp. 149-166.

[22] J. Yan, Disjoint triangles and quadrilaterals in a graph, *Discrete Mathematics* 308 (2008), pp. 3930-3937.

[23] G.A. Dirac, Some theorems on abstract graphs, *Proc. Lond. Math. Soc.* 2 (1952), pp. 69-81.

[24] J. A. Bondy, Pancyclic graphs I, *Journal of Combinatorial Theory*, Series B, Volume 11, (1971), pp. 80-84.

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

[26] H. Hàn, Y. Person and M. Schacht, On Perfect Matchings in Uniform Hypergraphs with Large Minimum Vertex Degree, *SIAM J. Discret. Math.*, 23(2), (2009), pp. 732–748.

[27] D. Kühn and D. Osthus, Matchings in hypergraphs of large minimum degree, *J. Graph Theory*, 51(4),(2006), pp. 269–280.

[28] D. Kühn, D. Osthus and A. Treglown, Matchings in 3-uniform hypergraphs, (submitted).

[29] O. Pikhurko, Perfect matchings and K3 4-tilings in hypergraphs of large codegree, *Graphs Combin.* 24(4) (2008) 391-404.

[30] V. Rödl, A. Ruciński and E. Szemerédi, A Dirac-Type Theorem for 3-Uniform Hypergraphs, *Comb. Probab. Comput.*, Volume 15, (2006), pp. 229–251.

[31] V. Rödl, A. Ruciński and E. Szemerédi, Perfect matchings in uniform hypergraphs with large minimum degree, *Europ. J. Combin.* 27 (2006) 1333-1349.

[32] V. Rödl, A. Ruciński and E. Szemerédi, An approximate Dirac-type theorem for k-uniform hypergraphs, *Combinatorica* 28(2) (2008) 229-260.

[33] V. Rödl, A. Ruciński and E. Szemerédi, A note on perfect matchings in uniform hypergraphs with large minimum collective degree, *Commentationes Mathematicae Universitatis Carolinae* 49(4) (2008) 633-636.

[34] V. Rödl, A. Ruciński and E. Szemerédi, Perfect matchings in large uniform hypergraphs with large minimum collective degree, *J. Comb. Theory Ser. A*, 116(3), (2009), pp. 613–636.

[35] V. Rödl and A. Ruciński, Dirac-type questions for hypergraphs – a survey (or more problems for Endre to solve), An Irregular Mind (Szemerédi is 70), Bolyai Soc. Math. Studies 21 (2010)

[36] I. Khan, Perfect matching in 3 uniform hypergraphs with large vertex degree, *submitted*

[37] A. Ruciński and K. Markström Perfect matchings and Hamilton cycles in hypergraphs with large degrees *Submitted*

Vita

Imdadullah Khan

2011 Ph.D. in Computer Science from Rutgers – The State University of New Jersey, NJ, USA.

2006 M.S. in Computer Science from Lahore University of Management Sciences, Lahore, Pakistan.

2002 BS in Computer Science from P.A.F. Degree College, University of Peshawar, Pakistan.