7-31-2007

The 3-color Ramsey number of a 3-uniform Berge-cycle

András Gyárfás

*Computer and Automation Research Institute, Hungarian Academy of Sciences, gyarfas@sztaki.hu*

Gábor N. Sárközy

*Worcester Polytechnic Institute, gsarkozy@cs.wpi.edu*

Follow this and additional works at: [https://digitalcommons.wpi.edu/computerscience-pubs](https://digitalcommons.wpi.edu/computerscience-pubs)

**Suggested Citation**


Retrieved from: [https://digitalcommons.wpi.edu/computerscience-pubs/157](https://digitalcommons.wpi.edu/computerscience-pubs/157)

This Other is brought to you for free and open access by the Department of Computer Science at Digital WPI. It has been accepted for inclusion in Computer Science Faculty Publications by an authorized administrator of Digital WPI. For more information, please contact digitalwpi@wpi.edu.
The 3-color Ramsey number of a 3-uniform Berge-cycle

András Gyárfás

Computer and Automation Research Institute
Hungarian Academy of Sciences
Budapest, P.O. Box 63
Budapest, Hungary, H-1518
gyarfas@sztaki.hu

Gábor N. Sárközy†

Computer Science Department
Worcester Polytechnic Institute
Worcester, MA, USA 01609
gsarkozy@cs.wpi.edu
and
Computer and Automation Research Institute
Hungarian Academy of Sciences
Budapest, P.O. Box 63
Budapest, Hungary, H-1518

July 31, 2007

Abstract

The asymptotics of 2-color Ramsey numbers of loose and tight cycles in 3-uniform hypergraphs have been recently determined ([16], [17]). We address here the same problem for Berge-cycles and for 3 colors. Our main result is that the 3-color Ramsey number of a 3-uniform Berge cycle of length $n$ is asymptotic to $\frac{5}{4}n$. The result is proved with the Regularity Lemma via the existence of a monochromatic connected matching covering asymptotically $4n/5$ vertices in the multicolored 2-shadow graph induced by the coloring of $K_n^{(3)}$.

1 Introduction

The investigations of Turán type problems for paths and cycles of graphs were started by Erdős and Gallai in [4]. The corresponding Ramsey problems first

---

*Research supported in part by the OTKA Grant No. K68322.
†Research supported in part by the National Science Foundation under Grant No. DMS-0456401 and by the OTKA Grant No. K68322.
for two colors and for paths have been looked at some years later in [8] and for three colors and for paths and cycles in [5], [12] and [19].

There are several possibilities to define paths and cycles in hypergraphs. In this paper we address the case of the Berge-cycle; probably it is the earliest definition of a cycle in hypergraphs in the book of Berge [1]. Turán type problems for Berge-paths and Berge-cycles of hypergraphs appeared perhaps first in [2]. Other types of hypergraph cycles, loose and tight, have been studied in [18] and [26]. The investigations of the corresponding Ramsey problems started quite recently with [16] and [17] where Ramsey numbers of loose and tight cycles have been determined asymptotically for two colors and for 3-uniform hypergraphs.

Let $H$ be an 3-uniform hypergraph (3-element subsets of a set). For vertices $x, y \in V(H)$ we say $x$ is adjacent to $y$, if there exists an edge $e \in E(H)$ such that $x, y \in e$. Let $K_n^{(3)}$ denote the complete 3-uniform hypergraph on $n$ vertices. A 3-uniform $\ell$-cycle, or Berge-cycle of length $\ell$, denoted $C^{(3)}_\ell$, is a sequence of distinct vertices $v_1, v_2, \ldots, v_\ell$, the core of the cycle, such that each $v_i$ is adjacent to $v_{i+1}$ and the edges $e_i$ that contain $v_i, v_{i+1}$ are all distinct for $1 \leq i \leq \ell$ where $\ell + 1 \equiv 1$. When 3-uniformity is clearly understood we may simply write $C_\ell$ for $C^{(3)}_\ell$. It is important to keep in mind that a 3-uniform Berge-cycle $C_\ell$ is not determined uniquely, it is considered as an arbitrary choice from many possible cycles with the same parameter. This is in contrast to the graph case or the case of loose and tight cycles in 3-uniform hypergraphs.

Let $R_t(C_n)$ denote the Ramsey number of a 3-uniform $n$ Berge-cycle using $t$ colors. It turns out that the case $t = 2$ can be easily solved: for $n > 4$, $R_2(C_n) = n$, i.e. there is a Hamiltonian Berge cycle in every 2-coloring of $K_n^{(3)}$, see [11]. In this paper we explore the 3-color Ramsey number of a Berge-cycle in 3-uniform hypergraphs, our main result is that $R_3(C_n) = (1 + o(1)) \frac{5n}{4} - \eta n$ - as far as we know this is the first 3-color Ramsey type result for cycles in hypergraphs. It seems purely incidental that our result has the same asymptotic as the 2-color Ramsey number of the loose $n$-cycle in 3-uniform hypergraphs, see ([16]).

**Theorem 1.** For all $\eta > 0$ there exists $n_0$ such that for every $n > n_0$, every coloring of the edges of $K_n^{(3)}$ with 3 colors contains a monochromatic Berge-cycle of length at least $\left(\frac{5}{4} - \eta\right)n$.

In fact we can prove the theorem in the following slightly stronger Ramsey formulation.

**Theorem 2.** For all $\eta > 0$ there exists $n_0$ such that for every $n > n_0$, we have the following

$$R_3(C_n) \leq \left(\frac{5}{4} + \eta\right)n.$$

Perhaps Theorem 1 can be extended as follows.

**Conjecture 3.** For all $\eta > 0$ and positive integer $r$ there exists $n_0 = n_0(\eta, r)$ such that for every $n > n_0$, every coloring of the edges of $K_n^{(3)}$ with $r$ colors contains a monochromatic Berge-cycle of length at least $\left(\frac{2r^2 - 2}{2r^2 - 1} - \eta\right)n$.  

2
Conjecture 3 (and thus Theorem 1) is best possible asymptotically as shown by the following construction. Let $A_1, \ldots, A_{r-1}$ be disjoint vertex sets of size $n/(2r-1)$ (for simplicity we assume that $n$ is divisible by $2r-1$). The $r$-edges not containing a vertex from $A_1$ are colored with color 1. The $r$-edges that are not colored yet and do not contain a vertex from $A_2$ are colored with color 2. We continue in this fashion. Finally the $r$-edges that are not colored yet with colors $1, \ldots, r-2$ and do not contain a vertex from $A_{r-1}$ are colored with color $r-1$. The $r$-edges that contain a vertex from all $r-1$ sets $A_1, \ldots, A_{r-1}$ get color $r$.

We claim that in this $r$-coloring of the edges of $K_{n}^{(r)}$ the longest monochromatic Berge-cycle has length $\leq \frac{2r-2}{2r-1}n$. This is certainly true for Berge-cycles in color $i$ for $1 \leq i \leq r-1$, since the subhypergraph induced by the edges in color $i$ leaves out $A_i$ (a set of size $n/(2r-1)$) completely. Finally, note that in a Berge-cycle in color $r$ from two consecutive vertices on the cycle one has to come from $A_1 \cup \ldots \cup A_{r-1}$ and thus the cycle has length at most $2(r-1)n/(2r-1)$.

The proofs of Theorems 1 and 2 use the following approach. To a given 3-uniform hypergraph $H$ consider the 2-shadow (or simply shadow) graph $\Gamma(H)$ on the same vertex set, with edge $(x, y) \in E(\Gamma(H))$ if and only if $x, y$ is covered by some hyperedge. To a given 3-coloring of the edges of the 3-uniform hypergraph, associate an edge multi-coloring of the shadow graph by coloring each edge with all colors appearing on hyperedges containing that pair. Edge (multi-)colorings of $\Gamma(H)$ defined this way will be called 3-uniform colorings of $\Gamma(H)$.

Then, following the method established in [24] and refined later in several papers ([5], [12], [13], [14], [15], [16], [17] and [19]), Theorems 1 and 2 can be reduced to finding a large (of size at least $\frac{2r-2}{2r-1}n$) monochromatic connected matching in any 3-uniform 3-coloring of $\Gamma(H)$ obtained from an almost complete hypergraph $H$ with $n$ vertices. Almost complete (or $(1-\epsilon)$-dense) means that $H$ has at least $(1-\epsilon)\binom{n}{3}$ edges. A monochromatic, say red matching is called connected, if its edges are in the same component in the graph defined by the red edges. Our key result is phrased as Lemma 4 and will be proved in Section 2.

**Lemma 4.** For all $\eta > 0$ there exist $\epsilon > 0$ and $n_0$ such that for every $n > n_0$ the following is true. In every 3-uniform 3-coloring of $\Gamma(H)$ obtained from an $(1-\epsilon)$-dense 3-uniform hypergraph $H$, there is a connected monochromatic matching of size at least $(\frac{2r-2}{2r-1} - \eta)n$.

In Section 3 we show how to use the Regularity Lemma to convert connected matchings into Berge-cycles, i.e. how to finish the proofs of Theorems 1 and 2. Although the approach outlined above is now becoming “standard”, there are several technical solutions to handle “almost complete” hypergraphs and their shadow graphs. We think that the following concept and the corresponding lemma (its straightforward proof is in [11]) are very convenient.

For $0 < \delta < 1$ fixed, we say that a sequence $L \subset V(H)$ of $k$ distinct vertices was obtained by a $\delta$-bounded selection if its elements are chosen in $k$ consecutive steps so that in each step there are at most $\delta n$ forbidden vertices that cannot be included as the next element. For simplicity, sometimes we shortly call the
sequence itself a $\delta$-bounded selection. Observe that a $\delta$-bounded selection $L$ is also a $\delta'$-bounded selection for any $\delta' > \delta$. A basic property of almost complete hypergraphs is expressed in the following lemma.

**Lemma 5.** Assume that $H$ is a $(1 - \epsilon)$-complete $r$-uniform hypergraph ($r \geq 2$) and set $\delta = \epsilon^{2^{-r}}$. There are forbidden sets such that for every $\delta$-bounded selection $L \subset V(H)$ of length at most $r$, at least $(1 - \delta) \frac{n^{r-|L|}}{(r-|L|)!}$ edges of $H$ contain $L$.

The case $|L| = r$ is very important, then we get that every $\delta$-bounded selection $L \subset V(H)$ of length at most $r$, at least $(1 - \delta)n^r$ edges of $H$ contain $L$.

To illustrate how to use Lemma 5, we generalize a result in [10] (more general form in [7]) from complete hypergraphs to almost complete ones. We start with a proposition (from [11]) about the connected components of a hypergraph.

**Proposition 6.** Assume $H$ is an arbitrary hypergraph and $0 < s < 1/3$. Then either there is a connected component $H'$ of $H$ with at least $(1 - s)n$ vertices or the connected components of $H$ can be partitioned into two groups so that each group contains more than $sn$ vertices.

**Proof.** Mark the connected components of $H$ until the union of them has at most $sn$ vertices. If one unmarked component remains, it can be $H'$. Otherwise, we form two groups from the unmarked components. The larger group has order at least $(n - sn)/2 > sn$, and the smaller one together with the marked components have a union containing more than $sn$ vertices as well. □

**Lemma 7.** Assume that $H$ is an $(1 - \epsilon)$-complete $r$-uniform hypergraph with $n$ vertices and $\delta = \delta(\epsilon, r) < \frac{1}{3}$. Then in every $r$-coloring of the edges of $H$ there exists a monochromatic connected component covering all but at most $\delta n$ vertices of $H$.

**Proof.** If the first possibility of Proposition 6 holds to any of the hypergraphs determined by the edges of the different color classes, we have nothing to prove. Otherwise the components of each color class can be partitioned into $X_i, Y_i$ so that both have at least $\delta n$ vertices. We shall reach a contradiction by defining a $\delta$-bounded selection of $r$ vertices as follows.

We want to select $x_1, x_2$ in the first two steps so that these vertices are in different partitions (one is in $X_i$, the other is in $Y_i$) for at least two values of $i$. This can be easily done as follows. Try an arbitrarily $y = x_1$ (apart from the $\delta n$ forbidden vertices), assume w.l.o.g that $y \in X_i$ for all $i$, $1 \leq i \leq r$. Try an $u = x_2$ such that $u \in Y_1$ (there is a choice since $|Y_1| > \delta n$). If $L = (y, u)$ does not work, it means that $u \in X_i$ for $i = 2, 3, \ldots, r$. Now select $z \in Y_2$ (there is a choice since $|Y_2| > \delta n$) and observe that either $L = (u, z)$ or $L = (y, z)$ satisfies the requirement.

Having $x_1, x_2$ with the property required in the previous paragraph, say $x_1 \in X_i$ for $i = 1, 2, \ldots, r$, $x_2 \in Y_1 \cap Y_2$, we continue the $\delta$-bounded selection.
by picking $x_j$ from $Y_j$ for $j = 3, \ldots, r$. Now the vertex set of the sequence $L = (x_1, x_2, \ldots, x_r)$ is an edge of $H$ so it has a color $k$, $1 \leq k \leq r$. However, this is a contradiction since $L$ has elements in both $X_k, Y_k$. □

2 Large connected matchings in almost complete 3-uniform 3-colorings

In this section we shall prove Lemma 4. We need some basic facts about matchings. The size $|M|$ of a maximum matching is the matching number, $\nu(G)$. The following result is often referred to as the Tutte - Berge formula (see for example in [23] Theorem 3.1.14). We shall use $c(G)$ and $co(G)$ for the number of components and odd components of a graph $G$ and def($G$), the deficiency of $G$, is defined as $|V(G)| - 2\nu(G)$.

**Lemma 8.** For any graph $G$, $def(G) = \max\{co(G \backslash T) - |T|\}$ where the maximum is taken over all $T \subseteq V(G)$.

We also need the following obvious property of maximum matchings.

**Lemma 9.** Suppose $M = \{e_1, \ldots, e_k\}$ is a maximum matching in a graph $G$. Then $V(G) \backslash V(M)$ spans an independent set and one can select one endpoint $x_i$ of each $e_i$ - we call it strong point - so that for each $i$, $1 \leq i \leq k$, there is at most one edge in $G$ from $x_i$ to $V(G) \backslash V(M)$.

We assume that $n$ is sufficiently large and

$$0 < \epsilon << \delta = \epsilon^{1/8} << \eta$$

and we may also assume that $\eta$ is sufficiently small since the statement (existence of a monochromatic connected matching of size $(\frac{n}{2} - \eta)n$) of Theorem 4 from any fixed $\eta$ follows automatically for any larger $\eta$. Consider an arbitrary 3-uniform 3-coloring of $K = \Gamma(H)$, where $H$ is an $(1 - \epsilon)$-complete 3-uniform hypergraph with $n$ vertices. The edges of $K$, the complement of $W$, will be sometimes referred as the “missing edges”. These are the edges uncovered by the hyperedges of $H$. For convenience, we shall also consider the exceptional edges (from $x_i$ to $V(G) \backslash V(M)$) of Lemma 9 as missing edges.

We call color $i$ good if there is a $V' \subseteq V$ such that $|V'| \geq (1 - \delta)n$ and the edges of color $i$ in $V'$ form only one nontrivial component $C$. (We shall use that in a good color no edge of $V' \backslash V(C)$ has color $i$.) We select $M_1$ as the largest monochromatic matching among matchings in good colors. Note that $M_1$ is well defined since by Lemma 7 there is a color, say red, with a connected component of at least $(1 - \delta)n$ vertices in $K$. We may assume that $|M_1| = k_1 = (\frac{n}{2} - \eta - \rho_1)n$ with some $0 < \rho_1 \leq (\frac{n}{2} - \eta)$, otherwise we are done. Furthermore, by a result of [12], we may also assume that $|M_1| = k_1 \geq (\frac{1}{2} - \eta)n$ (indeed this is true for any 3-coloring of an almost complete graph). Apply Lemma 9 to select the strong endpoints in $M_1$ and denote the set of these strong endpoints by $B$, the set of other endpoints by $A$ and $C = V(K_n) \backslash V(M_1)$. Thus we have
\[ |A| = |B| = k_1 = (\frac{2}{5} - \eta - \rho_1)n \geq (\frac{1}{4} - \eta)n, \]  
\[ (\frac{1}{5} + 2\eta + 2\rho_1 - \delta)n \leq |C| \leq (\frac{1}{2} + 2\eta)n. \]  

Call an edge of \( K \) purely-\{blue, green\} (or simply \( p\{-\text{blue, green}\} \)) if this edge cannot be red, so it can only be blue and/or green. (We have a multicoloring!) Similarly a \( p\)-green edge can only be green. Notice that - using the convention that the exceptional red edges from each vertex of \( B \) are considered as missing edges - every edge of \( K \) in \( C \cup (B \times C) \) is \( p\{-\text{blue, green}\} \). We shall frequently use the following fact.

**Fact 1.** Consider an edge \( e \in \mathcal{H} \) such that the triangle defined by \( e \) in \( K \) contains a \( p\{-\text{blue, green}\} \) edge. Then the other two edges of the triangle are also blue and/or green (however, they may also be red).

Indeed, \( e \in \mathcal{H} \) cannot be red, so it can only be blue or green.

Thus we have the following structural information about the 3-uniform 3-coloring of \( K \). Every edge of \( K \) is blue and/or green except perhaps the edges inside the set \( A \). In addition every edge of \( K \) in \( C \cup (B \times C) \) is \( p\{-\text{blue, green}\} \).

Define the subgraph \( H \) of \( K \) with vertex set \( A \cup B \cup C \) and with all edges of \( K \) in \( B \cup C \) and in \( A, B \cup C \). Now all edges of \( H \) have blue and/or green colors. Some of the edges of \( H \) might have a red color as well, we ignore that, i.e. consider \( H \) as a 2-multicolored graph. The pairs in \( B \cup C \) and in \( A, B \cup C \) that are not in \( K \) will be referred as the “missing edges” of \( H \).

**Proposition 10.** All but at most \( \delta n \) vertices of \( H \) has missing degree at most \( \delta n \).

**Proof.** Apply Lemma 5 for \( \delta \)-bounded selections such that \( x_1 \in A, x_2, x_3 \in B \cup C \). This shows that the missing degree of all but at most \( \delta n \) vertices of \( A \) is at most \( \delta n \). A similar selection starting with \( x_1 \in B \cup C \) shows that the statement is valid for vertices of \( B \cup C \) as well. \( \square \)

Next we establish some facts about the components of \( H \). We may assume that \( |A|, |B \cup C| > 2\delta n \). Applying Lemma 7 with \( r = 3 \) to the subgraph \( H[B \cup C] \) we get \( U_1 \subseteq B \cup C \) such that \( |(B \cup C) \setminus U_1| < \delta n \) and \( H[U_1] \) is connected in blue or green, say w.l.o.g in blue. In fact, we may assume that \( B \cup C = U_1 \) since deleting at most \( \delta n \) vertices does not influence the proof. Let \( U_2 \) be the set of vertices in \( A \) with at least one blue neighbor in \( U_1 \). Let \( K_1 \) be the the subgraph of \( H \) induced by the blue edges of \( U_1 \cup U_2 \). Observe that \( K_1 \) is the only nontrivial blue component of \( H \) so blue is a good color.

Case I. \( A \setminus U_2 \) is nonempty. Now all edges of \( [U_1, A \setminus U_2] \) are green. This implies that \( K_2 \), the subgraph of green edges of \( H \) is the only nontrivial green component of \( H \) so green is a good color. In this case we define \( M_2 \) as the larger of the maximum matchings of \( K_1, K_2 \), without loss of generality, \( M_2 \) is blue.
Case II. $A = U_2$. If there is only one nontrivial green component (i.e. if green is a good color) then we have the symmetry as in Case I and $M_2$ is defined as the larger of two maximal matchings, without loss of generality, $M_2$ is blue again. Otherwise - if there are more than one nontrivial green components, $M_2$ is defined as a maximum matching in $K_1$. (However, as it will turn out later we can find the required monochromatic matching without dealing with this possibility.)

Since $M_2$ is defined in a good color, $|M_1| \geq |M_2|$ (and no edge in $V(H) \setminus V(M_2)$ can be blue). We may assume that $|M_2| = k_2 = (\frac{2}{5} - \eta - \rho_2)n$ with some $0 < \rho_1 \leq \rho_2 \leq (\frac{2}{5} - \eta)$, otherwise we are done. In the remainder of the proof of Lemma 4 we will show in all cases that either directly we can find a green connected matching of size at least $\frac{2}{5}n$ or we show that there is only one nontrivial green component in Case II and it contains a matching $M_3$ with $|M_3| > |M_2|$, a contradiction.

Consider the set $R$ of remaining vertices that are not covered by $M_2$. Put $R_A = R \cap A$, $R_B = R \cap B$ and $R_C = R \cap C$. Apply Lemma 9 again to select the strong endpoints in $M_2$ and denote their set by $S$. Put $S_A = S \cap A$, $S_B = S \cap B$ and $S_C = S \cap C$. We have $R \cap S = \emptyset$. Denote the other (possibly weak) endpoints in $M_2$ by $W = W_A \cup W_B \cup W_C$. Thus we have $A = S_A \cup W_A \cup R_A$, $B = S_B \cup W_B \cup R_B$ and $C = S_C \cup W_C \cup R_C$ and these sets are all disjoint. We shall refer to these nine sets as atoms, and - by removing at most $9\delta n$ vertices - we may assume that every nonempty atom has order larger than $\delta n$. We have

$$|S| = |W| = k_2 = (\frac{2}{5} - \eta - \rho_2)n, \quad (4)$$

$$(\frac{1}{5} + 2\eta + 2\rho_2 - 10\delta)n \leq |R| \leq (\frac{1}{5} + 2\eta + 2\rho_2)n. \quad (5)$$

Notice that - considering at most one blue edge from each vertex of $S$ as a missing edge - every edge of $H$ in $R \cup (S \times R)$ is p-green.

Case 1: $R_C \neq \emptyset$ (then $|R_C| > \delta n$).

Consider a $\delta$-bounded selection starting with $v \in S$. Here we have the following claim.

Claim 1. All but at most $\delta n$ edges of $H$ incident to $v$ are green (they may be blue as well, so they are not necessarily p-green).

Indeed, assume by symmetry that $v \in S_A$. Let $u \in R_C$ be a second vertex in the $\delta$-bounded selection (this is possible, since $|R_C| > \delta n$) $(v, u)$ is p-green in $H$. From Lemma 5, for all but at most $\delta n$ choices of $w \in B \cup C$ the triple $\{u, v, w\}$ is an edge of $H$. The color of this edge cannot be red because of the edge $(u, w)$ that is p-$\{\text{blue, green}\}$ (since $u \in C$), it cannot be blue because of the edge $(v, u)$ that is p-green in $H$, so it must be green. Thus the edge $(v, u)$ is indeed green proving the claim (since the edge $(v, u)$ is also green).

Then clearly the green color is connected and we can span the vertices of the blue matching by a green matching (every blue matching edge is also green).
To get a larger green matching, we just add an arbitrary green edge in \( R_2 \), a contradiction. Thus in the rest we may assume that \( R_2 = \emptyset \).

**Case 2:** \( R_B \neq \emptyset \) (then \( \vert R_B \vert > \delta n \)).

We will define an auxiliary green subgraph \( H_1 \) of \( H \) having the green edges in the union of the following subgraphs

\[
H[S_A, R_B], H[S_B, R_A], H[S_B, R_B], H[R_A, R_B], H[R_B],
\]

\[
H[S_A, C], H[R_A, C], H[S_B, C], H[R_B, C].
\]

(6)

We show that \( H_1 \) contains almost all edges of \( H \) in the given subgraphs, so these subgraphs are almost totally green.

**Claim 2.** From all but at most \( \delta n \) vertices of \( H_1 \), all but at most \( \delta n \) edges of \( H \) are present in \( H_1 \) (so they are green).

Indeed, the claim is true for the subgraphs in the first line of (6) since \( R \cup (S \times R) \) is p-green in \( H \). For the subgraphs involving \( C \) we proceed similarly as in Claim 1. For \( H[S_A, C] \) and \( H[R_B, C] \) start a \( \delta \)-bounded selection with \( v \in S_A \). Continue with \( u \in R_B \) such that the edge \((u, v)\) is p-green in \( H \) (since \( v \) is a strong endpoint, at most one edge from \( v \) to \( R \) is blue in \( H \), all other edges are good). Let \( w \) be the third vertex of the selection from \( C \). Consider the color of the edge \((u, v, w)\) in \( H \). It cannot be red because of the edge \((u, w)\) that is p-blue, green} (since \( w \in C \)), it cannot be blue because of the edge \((u, v)\) that is p-green in \( H \), so it must be green. Thus the edges \((v, w)\) and \((u, w)\) are indeed green proving the claim for \( H[S_A, C] \) and \( H[R_B, C] \). For \( H[S_B, C] \) and \( H[R_A, C] \) it is similar. This proves the claim.

**Subcase 2.1:** \( S_C = \emptyset \).

In this case we will prove that \( H_2 \) has a matching \( M_3 \) leaving out at most \( \delta n \) vertices. This will be enough as this matching basically leaves out only those weak endpoints of \( M_2 \) which are not in \( C \), so altogether only \( k_2 - \vert C \vert + 10 \delta n \leq \frac{1}{5} n \) vertices using (1), (3) and (4) and thus

\[
\vert M_3 \vert \geq \frac{2}{5} n.
\]

Furthermore, \( H_2 \) is clearly connected, and thus the matching \( M_3 \) is connected.

For this purpose we need to estimate the maximum of \( c_o(H_2 \setminus T) - \vert T \vert \) in the Tutte-Berge formula. An argument of this type will be used several times later in the proof, at those places we will omit some of the details. In fact, we will estimate the maximum of a larger quantity, \( cr(T) = c(H_2 \setminus T) - \vert T \vert \) since in our special graph most odd components will be isolated points (where \( c \) denotes the number of components). Indeed, our graphs are of a very special type, they can be derived from a “base graph” \( G = G(H_1) \) with a constant number of vertices (at most eight) by replacing \( v \in V(G) \) with independent sets \( I_v \) of size at least \( \delta n \) and replacing their edges with almost complete bipartite graphs. In such graphs it is not hard to see that one can determine the maximum of \( cr(T) \) with a \( \delta n \) error term by looking at the maximum of expressions in the following form:

\[
cr^*(S) = \sum_{v \in S} |I_v| - \sum_{v \in N_G(S)} |I_v|,
\]
where $S$ is an independent set of $G$ and $N_G(S)$ is the neighborhood of $S$ in $G$. Indeed, here $\cup_{v \in N_G(S)} I_v$ corresponds to $T$ and removing $T$ from $H_1$ we get isolated points $\cup_{v \in S} I_v$ and some nontrivial components.

Following this approach in the next claim we estimate these expressions $cr^*(S)$ for the independent sets $S$ of $G(H_1)$. With a slight abuse of notation $R_A$ will denote the vertex of $G(H_1)$ as well as the set of vertices $R_A$ in $H_1$ (and similarly for the other sets).

**Claim 3.** Set $S_1 = \{S_A, R_A\}$, $S_2 = \{S_A, S_B\}$, $S_3 = \{C\}$, $S_4 = \{S_A\}$. Then

$$cr^*(S_1) = (|S_A| + |R_A|) - (|S_B| + |R_B| + |C|) \leq -\eta n,$$

$$cr^*(S_2) = (|S_A| + |S_B|) - (|R_A| + |R_B| + |C|) \leq -\eta n,$$

$$cr^*(S_3) = |C| - (|R_A| + |R_B| + |S_A| + |S_B|) \leq -\eta n,$$

$$cr^*(S_4) = |S_A| - (|R_B| + |C|) \leq -\eta n,$$

Indeed, we get these estimates from the following inequalities.

$$|S_A| + |R_A| \leq |A| = |B| = |S_B| + |R_B| + |W_B| < |S_B| + |R_B| + \frac{1}{5} n <$$

$$< |S_B| + |R_B| + |C| - \eta n$$

(here we used (3), (4) and the fact that in this subcase all vertices in $C$ are weak endpoints, and thus $C = W_C$ and $|W_B| < \frac{1}{5} n$),

$$|S_A| + |S_B| = |S| \leq \left(\frac{2}{5} - \eta \right) n < |R_A| + |R_B| + |C| - \eta n,$$

$$|C| \leq |W| = |S| = |S_A| + |S_B| \leq |S_A| + |S_B| + |R_A| + |R_B| - \eta n,$$

and

$$|S_A| + |R_A| \leq |A| \leq \left(\frac{2}{5} - \eta \right) n < |R| + |C| - \eta m =$$

$$= |R_A| + |R_B| + |C| - \eta m.$$

Furthermore, note that for $S_5 = \{R_A\}$ and $S_6 = \{S_B\}$ we have $cr^*(S_5) \leq cr^*(S_1) \leq -\eta n$ and $cr^*(S_6) \leq cr^*(S_2) \leq -\eta n$ since $N_{G}(S_5) = N_{G}(S_1)$ and $N_{G}(S_6) = N_{G}(S_2)$. We do not have to look at $\{R_B\}$ as this is not an independent set in $H_1$. Thus for each nonempty independent set $S$ of $G(H_1)$ we have $cr^*(S) \leq -\eta n$. This fact and a moment of reflection on the structure of $H_1$ implies the following.

**Claim 4.** For all $T \subseteq V(H_1)$, $cr(T) \leq \delta n$.

Then the Tutte-Berge formula (Lemma 8) implies that $H_1$ has an almost perfect matching, finishing the subcase.

**Subcase 2.2:** $|S_C| > 0$ (then $|S_C| > \delta n$).
In this subcase we extend $H_1$ with the green edges in the subgraphs

$$H[R_A, W_B, H[R_B, W_A], H[S_C, W_B], H[S_C, W_A].$$

Thus $H_1$ contains now all the vertices of $H$. Again, we can show that almost all edges in these subgraphs are in $H_1$, so they are green. Indeed, let us take a $\delta$-bounded selection starting with $v \in S_C$. Continue with $v \in R_A$ such that the edge $(v, u)$ is $p$-green in $H$ (since $v$ is a strong endpoint, at most one edge from $v$ to $R$ is blue in $H$, all other edges are good). Finish the selection with $w \in W_B$. Consider the color of the triple $(v, w)$ in the original $H$. This cannot be red because of the edge $(v, w)$ that is $p$-green (since $v \in C$), it cannot be blue because of the edge $(v, u)$ that is $p$-green in $H$, so it must be green. Thus the edges $(v, w)$ and $(u, w)$ are indeed green.

In this subcase we shall show again that $H_1$ has a matching of size at least $\frac{2}{5}n$. For this purpose we apply the Tutte-Berge formula again similarly as above so we are not going to present all the details. As above we have to show that for each nonempty independent set $S$ of $G(H_1)$ here we have $cr^*(S) \leq (\frac{4}{5} - \eta)n$. We have to check this inequality for all the subsets of the maximal independent sets $\{W_A, S_A, W_B, S_B\}, \{W_A, S_A, R_A\}, \{W_A, W_B, W_C\}, \{W_C, S_C\}$ and $\{W_B, R_B\}$.

For example for the independent sets $S_1 = \{W_A, S_A, W_B, S_B\}$ and $S_2 = \{W_C, S_C\}$ we get

$$cr^*(S_1) = (|S_A| + |W_A| + |S_B| + |W_B|) - (|R_A| + |R_B| + |C|) \leq \left(\frac{1}{5} - \eta\right)n,$$

$$cr^*(S_2) = (|W_C| + |S_C|) - (|A| + |B|) \leq \left(\frac{1}{5} - \eta\right)n,$$

from the following inequalities (using (2), (3) and (5))

$$|S_A| + |W_A| + |R_A| + |S_B| + |W_B| + |R_B| = |A| + |B| \leq \left(\frac{4}{5} - \eta\right)n \leq

\leq 2|R| + |C| + \left(\frac{1}{5} - \eta\right)n = 2(|R_A| + |R_B|) + |C| + \left(\frac{1}{5} - \eta\right)n,$$

and

$$|W_C| + |S_C| = |C| \leq \left(\frac{3}{5} - \eta\right)n \leq |A| + |B| + \left(\frac{1}{5} - \eta\right)n.$$ 

Similarly, the other independent sets can be checked, and we get that here for each nonempty independent set $S$ of $G(H_1)$ we have $cr^*(S) \leq (\frac{1}{5} - \eta)n$. Then from the Tutte-Berge formula again we get a matching $M_3$ in $H_1$ of size at least $\frac{2}{5}n$. This finishes Case 2, we may assume in the rest of the proof that $|R_B| = |R_C| = 0$ holds. Thus $M_2$ covers all the vertices in $B \cup C$.

At this point we have to refine the strong-weak structure of $M_2$. Any endpoint of any edge of $M_2$ is strong, if it has at most one blue edge to $R = V(H) \setminus V(M_2)$ in $H$ and it is weak if it has at least two blue edges to $R$ in $H$. By Lemma 9 every edge of $M_2$ has at least one (now maybe two) strong
endpoint. As above we define \( S = S_A \cup S_B \cup S_C \) and \( W = W_A \cup W_B \cup W_C \), now we have only \(|W| \leq |S|\). Denote by \( S(W) \) the set of strong endpoints of \( M_2 \) that are matched to \( W \). We have the following claim.

**Claim 5.** All edges of \( H \) in \( S(W) \) are \( p \)-green in \( H \) (they cannot be blue).

In fact, if we had a blue edge in \( S(W) \), then we could increase the size of the blue matching along an alternating path with five edges, a contradiction.

**Case 3:** \(|S_C| > 0 \) (then \(|S_C| > \delta n\)).

Here we define \( H_1 \) as the green edges in the union of the subgraphs

\[
\begin{align*}
H[R_A, W_B], H[R_A, S_B], H[R_A, S_C], H[R_A, W_C], \\
H[R_B, S_C], H[R_B, S_C], H[S_B, W_C], H[S_C, W_C].
\end{align*}
\]

Similarly as above we can show that \( H_1 \) contains almost all edges of \( H \) in these subgraphs, so these subgraphs are almost totally green. We can prove again with a Tutte-Berge argument that \( H_1 \) contains an almost perfect matching \( M_3 \). As above we have to show that for each nonempty independent set \( S \) of \( G(H_1) \) here we have \( cr^*(S) \leq -\eta n \). We have to check this inequality for all the subsets of the maximal independent sets \( \{W_B, W_C\}, \{W_B, S_B\}, \{R_A\} \) and \( \{S_C\} \). For example for the independent sets \( S_1 = \{W_B, W_C\} \) and \( S_2 = \{S_C\} \) we get

\[
\begin{align*}
&cr^*(S_1) = (|W_B| + |W_C|) - (|S_B| + |S_C| + |R_A|) \leq -\eta n, \\
&cr^*(S_2) = |S_C| - (|B| + |R_A| + |W_C|) \leq -\eta n
\end{align*}
\]

from the following inequalities

\[
|W_B| + |W_C| \leq |W| \leq |S| = |S_A| + |S_B| + |S_C| \leq |S_B| + |S_C| + |R_A| - \eta n,
\]

and

\[
|S_C| \leq |C| \leq |R| = |R_A| \leq |R_A| + |B| + |W_C| - \eta n
\]

(using \(|C| \leq |R|\), i.e. \(|M_1| \geq |M_2|\)). Similarly, the other independent sets can be checked, and we get that here for each nonempty independent set \( S \) of \( G(H_1) \) we have \( cr^*(S) \leq -\eta n \). Thus we have an almost perfect matching \( M_3 \) in \( H_1 \).

This matching leaves out only at most

\[
|W_A| + |S_A| + 10\delta n \leq \frac{1}{5} n
\]

vertices (using (1), (2) and (5)), as we wanted.

**Case 4:** Finally we may assume that \( S_C = \emptyset \).

Here we define \( H_1 \) as the green edges in the union of the subgraphs

\[
\begin{align*}
H[R_A, S_B], H[R_A, C], H[S_B, C].
\end{align*}
\]

Again \( H_1 \) contains almost all edges in these subgraphs, however, this \( H_1 \) leaves out possibly too many vertices (\(|W_A| + |W_B| + |S_A|\) which could be close to \( \frac{2}{3} n \)). Thus we have to extend \( H_1 \).
Let us consider those vertices in $W_B \cup W_C$ for which the corresponding strong endpoints in $S(W_B \cup W_C)$ are in $S_A$. Denote the set of these vertices by $W'$ and their strong endpoints by $S' = S(W')$. Thus $S' \subset S_A$. Write $S_B^1 = S((W_B \cup W_C) \setminus W') \subset S_B$ and $S_B^2 = S_B \setminus S_B^1$. We have the following estimate on the size of $S'$.

$$|S'| \geq |W_B| - (\rho_2 + \delta)n.$$  \hspace{1cm} (7)

Indeed, as the strong endpoints corresponding to vertices in

$$\left((W_A \cup S_A) \setminus S'\right) \cup \left((W_B \cup W_C) \setminus W'\right)$$

should all go to $S_B$, we have

$$|W_A| + |S_A| - |S'| + |W_C| + |W_B| - |S'| \leq |S_B| = |B| - |W_B|.$$  

From this, (3) and (5) we get the estimate

$$2|S'| \geq |W_A| + |S_A| + 2|W_B| + |C| - |B| = |A| - |R_A| + 2|W_B| + |C| - |B| =$$

$$= |A| - |R| + 2|W_B| + |C| - |B| \geq 2|W_B| + 2(\rho_1 - \rho_2 - \delta)n \geq 2(|W_B| - (\rho_2 + \delta)n),$$

and thus we get (7).

We extend $H_1$ with the green edges in the union of the subgraphs

$$H[S', C], H[S', S_B^1].$$

Again we can show that almost all edges in these subgraphs are in $H_1$, so they are green. Indeed, let us take a $\delta$-bounded selection with $u \in S', v \in S_B^1$ and $w \in C$. Consider the color of the triple $\{u, v, w\}$ in the original $\mathcal{H}$. This cannot be red because of the edge $(v, w)$ that is $p$-blue, green (since $w \in C$), it cannot be blue because of the edge $(v, u)$ that cannot be blue by Claim 5, so it must be green. Thus the edges $(u, v)$ and $(u, w)$ are indeed green.

In this case again we can prove with a Tutte-Berge argument that $H_1$ contains a matching $M_3$ covering all but at most $|S'| - |W_B| + (\rho_2 + 2\delta)n \geq \delta n$ using (7)) vertices of $H_1$. This will be enough as by (7) this matching leaves out only at most

$$|W_B| + |W_A| + |S_A| - |S'| + |S'| - |W_B| + (\rho_2 + 2\delta)n + 10\delta n \leq$$

$$\leq |W_A| + |S_A| + \rho_2 n + 12\delta n \leq \frac{1}{5} n$$

vertices, as we wanted. Here the last inequality follows from (1), (2), (5) and

$$|W_A| + |S_A| + \frac{1}{5} n + \rho_2 n + 12\delta n \leq |W_A| + |S_A| + |R_A| = |A| \leq \frac{2}{5} n.$$

For the existence of the matching $M_3$ again we have to show that for each nonempty independent set $S$ of $G(H_1)$ here we have

$$cr^*(S) \leq |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n.$$ \hspace{1cm} (8)
We have to check this inequality for all the subsets of the maximal independent sets \( \{S', R_A\}, \{S', S_B^2\}, \{S_B\} \) and \( \{C\} \). For example for the independent sets \( S_1 = \{S', R_A\} \) and \( S_2 = \{S', S_B^2\} \) we get
\[
cr^*(S_1) = (|S'| + |R_A|) - (|S_B| + |C|) \leq |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n,
\]
\[
cr^*(S_2) = (|S'| + |S_B^2|) - (|S_B| + |R_A| + |C|) \leq |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n
\]
from the following inequalities
\[
|S'| + |R_A| \leq |A| = |B| = |S_B| + |W_B| \leq |S_B| + |C| - \eta n \leq |S_B| + |C| + |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n
\]
and
\[
|S'| + |S_B^2| = |S'| - |W_B| + |W_B| + |S_B^2| \leq |S'| - |W_B| + (\rho_2 + 2\delta)n + |B| \leq |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n + \frac{2}{\eta}n \leq |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n + |R_A| + |C| + |S_B^2|.
\]
Similarly, the other independent sets can be checked, and we get that for each nonempty independent set \( S \) of \( G(H_1) \) (8) is satisfied, as we wanted. This completes the proof of Lemma 4.

3 From connected matchings to Berge-cycles

We shall assume throughout the rest of the paper that \( n \) is sufficiently large. First we will need a generalization of the Regularity Lemma ([28]) for hypergraphs. There are several generalizations of the Regularity Lemma for hypergraphs due to various authors (see [3], [6], [9], [27] and [29]). Here we will use the simplest one due to Chung [3]. First we need to define the notion of \( \varepsilon \)-regularity. Let \( \varepsilon > 0 \) and let \( V_1, V_2, V_3 \) be disjoint vertex sets of size \( m \), and let \( H \) be a 3-uniform hypergraph such that every edge of \( H \) contains exactly one vertex from each \( V_i \) for \( i = 1, 2, 3 \). The density of \( H \) is \( d_{H} = \frac{|E(H)|}{m^3} \). The triple \( \{V_1, V_2, V_3\} \) is called an \((\varepsilon, H)\)-regular triple of density \( d_{H} \) if for every choice of \( X_i \subset V_i, |X_i| > \varepsilon |V_i|, i = 1, 2, 3 \) we have
\[
|\mathcal{H}[X_1, X_2, X_3] - d_{H}| \leq \varepsilon.
\]
Here by \( \mathcal{H}[X_1, X_2, X_3] \) we denote the subhypergraph of \( H \) induced by the vertex set \( X_1 \cup X_2 \cup X_3 \). In this setting the 3-color version of the (weak) Hypergraph Regularity Lemma from [3] can be stated as follows.

Lemma 11 (3-color Weak Hypergraph Regularity Lemma). For every positive \( \varepsilon \) and positive integer \( m \) there are positive integers \( M \) and \( n_0 \) such that for \( n \geq n_0 \) the following holds. For all 3-uniform hypergraphs \( H_1, H_2, H_3 \) with \( V(H_1) = V(H_2) = V(H_3) = V, |V| = n \), there is a partition of \( V \) into \( l + 1 \) classes (clusters)
\[
V = V_0 + V_1 + V_2 + \ldots + V_l
\]
such that...
\begin{itemize}
  \item $m \leq l \leq M$
  \item $|V_1| = |V_2| = \ldots = |V_l|
  \item $|V_0| < \varepsilon n$
  \item apart from at most $\varepsilon(\frac{1}{3})$ exceptional triples, the triples $\{V_1, V_2, V_3\}$ are $(\varepsilon, \mathcal{H}_s)$-regular for $s = 1, 2, 3$.
\end{itemize}

For an extensive survey on different variants of the Regularity Lemma see [22].

Consider a 3-edge coloring $(\mathcal{H}_1, \mathcal{H}_2, \mathcal{H}_3)$ of the 3-uniform complete hypergraph $K_n^{(3)}$, i.e. $\mathcal{H}_1$ is the subhypergraph induced by the first color, $\mathcal{H}_2$ is the subhypergraph induced by the second color, and $\mathcal{H}_3$ is the subhypergraph induced by the third color.

We apply the above 3-color Weak Hypergraph Regularity Lemma with a small enough $\varepsilon$ and we get a partition of $V(K_n^{(3)}) = V = \bigcup_{0 \leq i \leq l} V_i$, where $|V_i| = m, 1 \leq i \leq l$. We define the following reduced hypergraph $\mathcal{H}^R$: The vertices of $\mathcal{H}^R$ are $p_1, \ldots, p_l$, and we have a triple on vertices $p_i, p_j, p_k$ if the triple $\{V_{i}, V_{j}, V_{k}\}$ is $(\varepsilon, \mathcal{H}_s)$-regular for $s = 1, 2, 3$. Thus we have a one-to-one correspondence $f : p_i \rightarrow V_i$ between the vertices of $\mathcal{H}^R$ and the clusters of the partition. Then,

$$|E(\mathcal{H}^R)| \geq (1 - \varepsilon)\left(\frac{l}{3}\right),$$

and thus $\mathcal{H}^R$ is a $(1 - \varepsilon)$-dense 3-uniform hypergraph on $l$ vertices. Define a 3-edge coloring $(\mathcal{H}_1^R, \mathcal{H}_2^R, \mathcal{H}_3^R)$ of $\mathcal{H}^R$ with the majority color, i.e. the triple $\{V_{i_1}, V_{i_2}, V_{i_3}\} \in \mathcal{H}_s^R$ if color $s$ is the most frequent color in this triple. Note then that the density of this color is $\geq 1/3$ in this triple. Finally we consider the multi-colored shadow graph $\Gamma(\mathcal{H}^R)$. The vertices are $V(\mathcal{H}^R) = \{p_1, \ldots, p_l\}$ and we join vertices $x$ and $y$ by an edge of color $s$, $s = 1, 2, 3$ if $x$ and $y$ are contained in an edge of $\mathcal{H}^R$ that is colored with color $s$.

The main lemma that allows us to convert monochromatic connected matchings into monochromatic Berge-cycles is the following.

\textbf{Lemma 12.} Assume that for some positive constant $c$ we can find a monochromatic connected matching $M$ spanning at least $cl$ vertices in $\Gamma(\mathcal{H}^R)$. Then in the original 3-edge colored $K_n^{(3)}$ we can find a monochromatic Berge-cycle of length at least $c(1 - 3\varepsilon)n$.

We note here again that the use of a connected matching in this type of proofs (first suggested by [24]) has become somewhat standard by now (see [5], [11], [12], [13], [14], [15] and [24]), so a proof of this lemma can be found in [11] for example. For the sake of completeness we give the proof here again.

\textbf{Proof.} We may assume that $M$ is in $\Gamma(\mathcal{H}_1^R)$. Denote the edges on $M$ by $M = \{e_1, e_2, \ldots, e_l\}$ and thus $2l_1 \geq cl$. Furthermore, write $f(e_i) = (V_{i_1}, V_{i_2})$ for $1 \leq i \leq l_1$ where $V_{i_1}, V_{i_2}$ are the clusters assigned to the end points of $e_i$.  

14
Next we define good vertices for an arbitrary edge \( e \) in \( \mathcal{H}(\mathcal{H}^R) \). Denote \( f(e) = (V^1, V^2) \). Since \( e \) is an edge in \( \mathcal{H}(\mathcal{H}^R) \), the endpoints of \( e \) are contained in a triple \( E \) in \( \mathcal{H}^R \). By definition this triple corresponds to an \((\varepsilon, \mathcal{H}_1)\)-regular triple \( f(E) \) (containing clusters \( V^1, V^2 \) and one more cluster) that has density \( \geq 1/3 \). We say that a vertex \( x \in V^{j}, j = 1, 2 \) is good for \( V^j \), \( j' = 1, 2, j' \neq j \) if for at least \( m \) vertices \( y \in V^{j'} \), there are at least \( m \) triples in \( \mathcal{H}_1[f(E)] \) containing both \( x \) and \( y \). The next claim shows that most vertices are good in each \( V^j \).

**Claim 6.** In each \( V^j \), \( j = 1, 2 \) the number of vertices that are good for \( V^j \), \( j' = 1, 2, j' \neq j \) is at least \( (1 - \varepsilon)m \).

Indeed, let \( X \subset V^j \) denote the set of vertices in \( V^j \) that are not good for \( V^j \). Assume indirectly that \( |X| > \varepsilon m \). The total number of triples in \( \mathcal{H}_1[f(E)] \) that contain a vertex from \( X \) is smaller than

\[
|X| \left( \frac{m}{6} + (1 - \frac{1}{6})m \right) = \left( \frac{1}{3} - \frac{1}{36} \right)|X|m^2,
\]

which contradicts the fact that \( f(E) \) is an \((\varepsilon, \mathcal{H}_1)\)-regular triple with density at least \( 1/3 \) if \( \varepsilon \) is small enough. Thus the claim is true.

The good vertices determine an auxiliary bipartite graph \( G(V^1, V^2) \) in the following natural way. In \( V^j \), \( j = 1, 2 \) we keep only the vertices that are good for \( V^j \), \( j' = 1, 2, j' \neq j \). For simplicity we keep the \( V^1, V^2 \) notation. For a vertex \( x \in V^j \) that is good for \( V^j \) we connect it in \( G(V^1, V^2) \) to the \( \geq (1/6 - \varepsilon)m > m/7 \) vertices \( y \in V^{j'} \) such that there are at least \( m \) triples in \( \mathcal{H}_1[f(E)] \) containing both \( x \) and \( y \).

At this point we introduce a one-sided notion of regularity. A bipartite graph \( G(A, B) \) is \((\varepsilon, \delta, G)\)-super-regular if for every \( X \subset A \) and \( Y \subset B \) satisfying \( |X| > \varepsilon|A|, |Y| > \varepsilon|B| \) we have

\[
E_G(X, Y) > \delta|X||Y|,
\]

and furthermore,

\[
deg_G(a) > \delta|B| \quad \text{for all } a \in A, \quad \text{and} \quad deg_G(b) > \delta|A| \quad \text{for all } b \in B.
\]

Then it is not hard to see that the following is true.

**Claim 7.** \( G(V^1, V^2) \) is a \((2\varepsilon, 1/7, G)\)-super-regular bipartite graph.

Indeed, the second condition of super-regularity follows from (10). For the first condition let \( X \subset V^1, Y \subset V^2 \) with \( |X| > 2\varepsilon|V^1| (> \varepsilon m), |Y| > 2\varepsilon|V^2| (> \varepsilon m) \). Assume indirectly that \( E_G(X, Y) \leq |X||Y|/7 \). The total number of triples in \( \mathcal{H}_1[f(E)] \) that contain a vertex from \( X \) and a vertex from \( Y \) is smaller than

\[
|X||Y| \left( \frac{m}{7} + (1 - \frac{1}{7})m \right) = \left( \frac{1}{3} - \frac{1}{21} \right)|X||Y|m,
\]

15
which again contradicts the fact that \( f(E) \) is an \((\varepsilon, \mathcal{H}_1)\)-regular triple with density at least \( 1/3 \). Thus the claim is true.

Since \( M \) is a connected matching in \( \Gamma(\mathcal{H}_1^R) \) we can find a connecting path \( P_i^R \) in \( \Gamma(\mathcal{H}_1^R) \) from \( f^{-1}(V_2^i) \) to \( f^{-1}(V_{i+1}^1) \) for every \( 1 \leq i \leq l_1 \) (for \( i = l_1 \) put \( i+1 = 1 \)). Note that these paths in \( \Gamma(\mathcal{H}_1^R) \) may not be internally vertex disjoint. From these paths \( P_i^R \) in \( \Gamma(\mathcal{H}_1^R) \) we can construct vertex disjoint connecting paths \( P_i \) in \( \Gamma(\mathcal{H}_1) \) connecting a vertex \( v_2^i \) of \( V_2^i \) that is good for \( V_1^i \) to a vertex \( v_{i+1}^1 \) of \( V_{i+1}^1 \) that is good for \( V_1^1 \). More precisely we construct \( P_i \) with the following simple greedy strategy. Denote \( P_t^R = (p_1, \ldots, p_t), 2 \leq t \leq l \), where according to the definition \( f(p_1) = V_2^1 \) and \( f(p_t) = V_2^l \). Let the first vertex \( u_1 = v_2^1 \) of \( P_1 \) be a vertex \( u_1 \in V_2^1 \) that is good for both \( V_1^1 \) and \( f(p_2) \). By Claim 6 most of the vertices satisfy this in \( V_2^1 \). The second vertex \( u_2 \) of \( P_1 \) is a vertex \( u_2 \in (f(p_2) \cap N_G(f(p_1), f(p_2)))(u_1) \) (using the above defined bipartite graph \( G \)) that is good for \( f(p_3) \). Again using Claim 6 and the fact that \( \varepsilon \) is sufficiently small, most vertices satisfy this in \( f(p_2) \cap N_G(f(p_1), f(p_2)) \) (\( u_1 \)). The third vertex \( u_3 \) of \( P_1 \) is a vertex \( u_3 \in (f(p_3) \cap N_G(f(p_2), f(p_3))(u_2)) \) that is good for \( f(p_4) \).

We continue in this fashion, finally the last vertex \( u_l = v_2^l \) of \( P_1 \) is a vertex \( u_l \in (f(p_l) \cap N_G(f(p_{l-1}), f(p_l))(u_{l-1})) \) that is good for \( V_2^l \).

Then we move on to the next connecting path \( P_2 \). Here we follow the same greedy procedure, we pick the next vertex from the next cluster in \( P_2^R \). However, if the cluster has occurred already on the path \( P_2^R \), then we just have to make sure that we pick a vertex that has not been used on \( P_1 \).

We continue in this fashion and construct the vertex disjoint connecting paths \( P_i \) in \( \Gamma(\mathcal{H}_1) \), \( 1 \leq i \leq l_1 \). Next we have to make these connecting paths Berge-paths. By the construction, since every edge on every path \( P_i, 1 \leq i \leq l_1 \) came from an appropriate bipartite graph \( G \), the two endpoints of every edge are contained in at least \( m/6 \) triples in \( \mathcal{H}_1[f(E)] \). Since the total number of edges on the paths \( P_i \) is a constant \( (\leq l^2) \) and \( n \) (and thus \( m \)) is sufficiently large, we can clearly “assign” a triple from \( \mathcal{H}_1 \) for each edge on the paths such that the assigned triple contains the corresponding edge and the assigned \( \mathcal{H}_1 \) triples are distinct for distinct edges on the paths \( P_i \).

We remove the internal vertices of these paths \( P_i \) from \( f(M) \). We also remove the triples from \( \mathcal{H}_1 \) that are assigned to the edges of the paths \( P_i \), since these triples cannot be used again on the Berge-cycle. Note again that the number of vertices and edges that we remove this way is a constant. Furthermore, in a pair \((V_1^i, V_2^i)\) in \( V_i \) we keep only the vertices that are good for \( V_2^i \), and in \( V_2^i \) we keep only the vertices that are good for \( V_1^i \); all other vertices are removed. By these removals we may create some discrepancies in the cardinalities of the clusters of this connected matching. We remove an additional at most \( 2m \) vertices from each cluster \( V_2^i \) of the matching to assure that now we have the same number of vertices left in each cluster of the matching. For simplicity we still keep the notation \( V_2^i \). Note that by Claim 7 the remaining bipartite graph \( G(V_1^i, V_2^i) \) is clearly still \((4\varepsilon, 1/8, G)\)-super-regular for every \( 1 \leq i \leq l_1 \) and now we have \( |V_1^i| = |V_2^i| \).

We will use the following property of \((\varepsilon, \delta, G)\)-super-regular pairs.
Lemma 13. For every $\delta > 0$ there exist an $\varepsilon > 0$ and $m_0$ such that the following holds. Let $G$ be a bipartite graph with bipartition $V(G) = V_1 \cup V_2$ such that $|V_1| = |V_2| = m \geq m_0$, and let the pair $(V_1, V_2)$ be $(\varepsilon, \delta, G)$-super-regular. Then for every pair of vertices $v_1 \in V_1, v_2 \in V_2$, $G$ contains a Hamiltonian path connecting $v_1$ and $v_2$.

A lemma somewhat similar to Lemma 13 is used by Łuczak in [24]. Lemma 13 is a special case of the much stronger Blow-up Lemma (see [20] and [21]). Note that an easier approximate version of this lemma would suffice as well, but for simplicity we use this lemma.

Applying Lemma 13 inside each $G(V^i_1, V^i_2)$, $1 \leq i \leq l$, together with the connecting paths $P_i$ we get a cycle $C$ in $\Gamma(H_1)$ that has length at least
\[ cl(1 - 2\varepsilon)m \geq c(1 - \varepsilon)(1 - 2\varepsilon)n \geq c(1 - 3\varepsilon)n. \]

We only have to make this cycle a Berge-cycle. For the edges on the connecting paths $P_i$ we already have assigned distinct $H_1$ triples. The other edges came from the bipartite graphs $G(V^i_1, V^i_2)$, $1 \leq i \leq l$, and thus the two endpoints of every edge are contained in at least $m/7$ (we already removed some vertices and triples) triples in $H_1[f(E_i)]$ (here $E_i$ denotes the triple in $H^R_1$ containing the endpoints of the edge $e_i$). Note that the triples $E_i$ must be distinct for each $i, 1 \leq i \leq l$, and furthermore the triples containing two distinct edges from $G(V^i_1, V^i_2)$ are distinct. Thus if $m$ is sufficiently large we can clearly assign distinct triples to each edge on $C$ and this makes the cycle $C$ a Berge-cycle, completing the proof of Lemma 12. \hfill \Box

Putting together Lemma 12 with the asymptotic result of the previous section on monochromatic connected matchings (Lemma 4) we get the desired asymptotic result on monochromatic Berge-cycles (Theorem 1).

Here in the hypergraph case there are no parity problems, we can easily modify the proof technique of this section to yield the stronger Ramsey formulation in Theorem 2. Indeed Lemma 13 has the following stronger form. In the statement of the lemma $v_1$ and $v_2$ can actually be connected by a path of length $m'$ for every even integer $4 \leq m' \leq 2m$. If the parity is not right we can change the parity with the following simple trick. Consider $P^R = (p_1, \ldots, p_l)$. Since the edge $(p_1, p_2)$ is in $\Gamma(H^R_1)$, there is a triple $E$ in $H^R_1$ containing $p_1$ and $p_2$. Take the third vertex $p$ from $E$ that is different from $p_1$ and $p_2$ and splice in $p$ between $p_1$ and $p_2$ on the connecting path $P^R$. This way we increased the length by one and thus we changed the parity. Hence in Theorem 2 we can find a monochromatic Berge-cycle of length exactly $n$. 

17
References


