Available online at www.sciencedirect.com
ScienceDirect
Journal of Combinatorial Theory, Series B 98 (2008) 342–358
Journal of Combinatorial Theory
Series B
www.elsevier.com/locate/jctb

Monochromatic Hamiltonian Berge-cycles in colored complete uniform hypergraphs

András Gyárfás [a], Jenő Lehel [a,b,1], Gábor N. Sárközy [a,c,2], R.H. Schelp [b]

[a] *Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest, PO Box 63, Budapest, Hungary H-1518*

[b] *University of Memphis, Department of Mathematical Sciences, Memphis, TN 38152, USA*

[c] *Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609, USA*

Received 30 January 2007

Available online 22 October 2007

Abstract

We conjecture that for any fixed r and sufficiently large n, there is a monochromatic Hamiltonian Berge-cycle in every $(r-1)$-coloring of the edges of $K_n^{(r)}$, the complete r-uniform hypergraph on n vertices. We prove the conjecture for $r=3, n \geqslant 5$ and its asymptotic version for $r=4$. For general r we prove weaker forms of the conjecture: there is a Hamiltonian Berge-cycle in $\lfloor (r-1)/2 \rfloor$-colorings of $K_n^{(r)}$ for large n; and a Berge-cycle of order $(1-o(1))n$ in $(r-\lfloor \log_2 r \rfloor)$-colorings of $K_n^{(r)}$. The asymptotic results are obtained with the Regularity Lemma via the existence of monochromatic connected almost perfect matchings in the multicolored shadow graph induced by the coloring of $K_n^{(r)}$.

Keywords: Colored complete uniform hypergraphs; Monochromatic Hamiltonian Berge-cycles; Almost perfect connected matchings; Regularity Lemma

1. Introduction

Let $\mathcal{H}$ be an r-uniform hypergraph (a family of some r-element subsets of a set). For vertices $x, y \in V(\mathcal{H})$ we say x is adjacent to y, if there exists an edge $e \in E(\mathcal{H})$ such that $x, y \in e$. Let

E-mail addresses: gyarfas@sztaki.hu (A. Gyárfás), jlehel@memphis.edu (J. Lehel), gsarkozy@cs.wpi.edu (G.N. Sárközy), schelpr@msci.memphis.edu (R.H. Schelp).

[1] Research funded by the 2006–2007 Professional Development Assignment program of The University of Memphis.

[2] Research supported in part by the National Science Foundation under Grant No. DMS-0456401.

0095-8956/$ – see front matter
doi:10.1016/j.jctb.2007.07.002

$K_n^{(r)}$ denote the complete r-uniform hypergraph on n vertices. An r-uniform ℓ-cycle, or *Berge-cycle* of length ℓ, denoted $C_\ell^{(r)}$, is a sequence of distinct vertices $v_1, v_2, \dots, v_\ell$, the *core of the cycle*, together with a set of distinct edges $e_1, \dots, e_\ell$ such that e_i contains v_i, v_{i+1} ($v_{\ell+1} \equiv v_1$). When the uniformity is clearly understood we may simply write C_ℓ for $C_\ell^{(r)}$. A Berge-cycle of length n in a hypergraph of n vertices is called a Hamiltonian Berge-cycle.

Let $R_k(C_\ell^{(r)})$ denote the Ramsey number of the r-uniform ℓ cycle using k colors, that is the smallest n such that every k-coloring of the edges of $K_n^{(r)}$ admits a monochromatic Berge-cycle of length ℓ. It is important to keep in mind that—in contrast to the case $r = 2$— a Berge-cycle $C_\ell^{(r)}$ is not determined uniquely for $r > 2$; it is considered as an arbitrary choice from many possible cycles with the same pair of parameters. It is worth mentioning two special Berge-cycles, the *loose* and the *tight* cycles. In a loose cycle the edges of the cycle intersect the core sequence in consecutive pairs and are pairwise disjoint outside the core, while in a tight cycle the edges are the consecutive r-element subsets of the core sequence. The asymptotic values of 2-color Ramsey numbers for loose and tight cycles have been determined recently, see [10,11].

In this paper we give a conjecture about the number of colors in edge colorings of $K_n^{(r)}$ that ensures a monochromatic Hamiltonian Berge-cycle. Thinking in terms of graphs, such an attempt seems hopeless, since in many 2-colorings of K_n there are no monochromatic Hamiltonian cycles. For example, if each edge incident to a fixed vertex is red and the other edges are blue, there is no monochromatic Hamiltonian cycle. However, from the nature of Berge-cycles, this example does not extend to hypergraphs. If $K_n^{(3)}$ is colored in this way, there is a red Hamiltonian Berge-cycle (for $n \geqslant 5$).

Conjecture 1.1. *Assume that $r \geqslant 2$ is fixed and n is sufficiently large. Then every $(r-1)$-coloring of $K_n^{(r)}$ contains a monochromatic Hamiltonian Berge-cycle.*

It is worth noting that the number of colors, $r - 1$, cannot be increased in the conjecture. A construction in [9] shows that for r colors the size of the largest monochromatic Berge-cycle can be at most $(2r - 2)n/(2r - 1)$.

In Section 2 we shall prove Conjecture 1.1 for $r = 3$ and a general but weaker form for $r \geqslant 4$ when the number of colors is $\lfloor (r - 1)/2 \rfloor$ (the latter is trivial for $r = 4$).

Theorem 1.2. *If $K_n^{(3)}$, $n \geqslant 5$, is colored with two colors, then there exists a monochromatic Hamiltonian Berge-cycle.*

Theorem 1.3. *If $K_n^{(r)}$, $r \geqslant 4$, is colored with $\lfloor (r - 1)/2 \rfloor$ colors, and n is sufficiently large, then there exists a monochromatic Hamiltonian Berge-cycle.*

In Sections 3 and 4 we prove our main results, Conjecture 1.1 for $r = 4$ in asymptotic form, and a weaker version for general r.

Theorem 1.4. *For all $\eta > 0$ there exists $n_0 = n_0(\eta)$ such that every coloring of the edges of $K_n^{(4)}$, $n > n_0$, with 3 colors contains a monochromatic Berge-cycle of length at least $(1 - \eta)n$.*

Theorem 1.5. *For all $\eta > 0$ and all integers $r, k \geqslant 2$ with $r \geqslant k + \lfloor \log_2(k + 1) \rfloor$, there exists $n_0 = n_0(\eta, r, k)$ such that for every $n > n_0$, every coloring of the edges of $K_n^{(r)}$ with k colors contains a monochromatic Berge-cycle of length at least $(1 - \eta)n$.*

The proofs of Theorems 1.4 and 1.5 use the approach of [9]. Assume that $\mathcal{H}$ is an r-uniform hypergraph. The *shadow graph* of $\mathcal{H}$ is defined as the graph $\Gamma(\mathcal{H})$ on the same vertex set, where two vertices are adjacent if they are covered by at least one edge of $\mathcal{H}$. A hypergraph is called connected if its shadow graph is connected (and its components are defined similarly). A coloring of the edges of an r-uniform hypergraph $\mathcal{H}$, $r \geqslant 2$, induces a multicoloring on the edges of the shadow graph $\Gamma(\mathcal{H})$ in a natural way; every edge e of $\Gamma(\mathcal{H})$ receives the color of all hyperedges containing e. A multicoloring obtained in this way will be called an *r-uniform coloring* of $\Gamma(\mathcal{H})$. A subgraph of $\Gamma(\mathcal{H})$ is *monochromatic* if the color sets of its edges have a nonempty intersection.

We shall assume that $n = |V(\mathcal{H})|$ tends to infinity and define an almost Hamiltonian Berge-cycle of $\mathcal{H}$ as a Berge-cycle of length $(1 - o(1))n$. Similarly, a set of pairwise disjoint edges of the shadow graph saturating $n - o(n)$ vertices is called an almost perfect matching of $\Gamma(\mathcal{H})$. An r-uniform hypergraph is almost complete, if it has at least $(1 - o(1))\binom{n}{r}$ edges. A matching in a graph is *connected* if its edges are in the same connected component of the graph.

Following the method established in [16] and refined later in various papers, see [2,5,6,9–12], the asymptotic version of Conjecture 1.1 can be reduced to the following conjecture, which seems to have independent interest.

Conjecture 1.6. *Assume that $r \geqslant 2$ is fixed, $\mathcal{H}$ is an almost complete r-uniform hypergraph with n vertices, and its edges are colored with $r - 1$ colors. Then the r-uniform coloring induced on its shadow graph $\Gamma(\mathcal{H})$ contains a monochromatic almost perfect connected matching.*

Here is the outline how to derive the asymptotic version of Conjecture 1.1. If Conjecture 1.6 is true then for any given $\eta > 0$ there is a small enough ϵ such that for every $(r - 1)$-coloring of an r-uniform $(1 - \epsilon)$-complete hypergraph $\mathcal{H}$ (missing at most $\epsilon\binom{n}{r}$ edges), the induced coloring on $\Gamma(\mathcal{H})$ has a monochromatic connected matching covering at least $(1 - \eta)|V(\Gamma(\mathcal{H}))|$ vertices. Consider an $(r - 1)$-coloring of $\mathcal{H} = K_n^{(r)}$ and take an ϵ-regular partition on its vertices into clusters. (Here we apply a "colored" hypergraph-version of the Regularity Lemma.) Then Conjecture 1.6 is applicable to $\mathcal{H}^R$, the "reduced hypergraph" associated to the clusters and equipped with the standard majority coloring. This gives a large monochromatic connected matching in the shadow graph of $\mathcal{H}^R$. The final step is completed by Lemma 4.2 to appropriately connect the vertices of the clusters associated with the matching, to form a monochromatic Berge-cycle of length at least $(1 - 3\epsilon)(1 - \eta)n$ in $\mathcal{H} = K_n^{(r)}$.

We give an inductive argument in (Proposition 3.5) showing that it is enough to prove Conjecture 1.6 in a weaker form, dropping the connectivity condition of the required large monochromatic matching. This allows us to prove our results inductively, starting from the case $k = 1, r = 2$. We prove Conjecture 1.6 in Section 3 for $r = 2, 3, 4$. In general we can prove only Theorem 3.9, a weaker version of the conjecture, where the number of colors is the largest integer k such that $k + \lfloor \log_2(k + 1) \rfloor \leqslant r$ (k is at least $r - \lfloor \log_2 r \rfloor$). In Section 4 we show how to use the Regularity Lemma to convert connected matchings into Berge-cycles, i.e. how to finish the proof of Theorems 1.4 and 1.5.

2. Monochromatic Hamiltonian Berge-cycles

Proof of Theorem 1.2. If $n = 5$, let $\mathcal{H}_r$ and $\mathcal{H}_b$ be the hypergraphs formed by the red and blue edges, respectively, in a 2-coloring of $K_5^{(3)}$. Easy inspection shows that a 3-uniform hypergraph $\mathcal{H}$ with five vertices and at least five edges has a Berge-cycle $C_5^{(3)}$ unless $\mathcal{H}$ is isomorphic to $K_4^{(3)}$

extended with an edge that intersects $V(K_4^{(3)})$ in two vertices. Because one of $\mathcal{H}_r$ and $\mathcal{H}_b$ must be different from this exceptional hypergraph, the theorem follows for $n=5$. For $n>5$ we proceed by induction. Assume that the vertex set of a 2-colored $K_n^{(3)}$ is $[n]$ and let $P=(1,2,\dots,n-1)$ be a cyclic permutation of $[n-1]$ representing the core sequence of a red Berge-cycle $C_{n-1}^{(3)}$ that exists by induction. Color the consecutive pairs $(i,i+1)$ of P with the color of the edge $\{n,i,i+1\}\in E(K_n^{(3)})$. If the pairs $(i,i+1)$ and $(i+1,i+2)$ are both red then we have a red $C_n^{(3)}$ with the core sequence obtained by inserting n between $i+1$ and $i+2$ and using the red edges $\{i,i+1,n\}$ and $\{n,i+1,i+2\}$ to cover $(i+1,n)$ and $(n,i+2)$. Similarly, if $(i,i+1)$ is red and at least one of the edges $\{i,n,j\}$ $(j\neq i+1)$ and $\{i+1,n,j\}$ $(j\neq i)$ is red, we obtain a red Berge-cycle $C_n^{(3)}$.

Therefore, if (a_i,b_i), $i=1,\dots,k$ are the red consecutive pairs of P following the orientation on P, then we may assume that every $\{n,a_i,x\}$ $(x\neq b_i)$ and every $\{n,b_i,y\}$ $(y\neq a_i)$ is blue. This allows us to easily find a blue Berge-cycle $C_n^{(3)}$ as follows. If $k=1$ then the blue core sequence Q is obtained from P by including n in P between a_1 and b_1. To obtain the blue $C_n^{(3)}$, use the blue edges $\{a_1,n,b_1+1\}$ and $\{n,b_1,a_1-1\}$ to cover the pairs (a_1,n) and (n,b_1), then use the blue edges $\{n,i,i+1\}$ to cover all other consecutive pairs. Otherwise, Q is defined by the cyclic order $Q=(a_1,a_2,-,b_1,a_3,-,b_2,a_4,-,\dots,-,b_{k-1},n,b_k,+)$ where the minuses indicate blue subpaths following P backwards and the plus means a subpath following P forward. By the assumption, every consecutive pair on Q which does not contain n can be extended to a blue triple by adding n to it. The pairs (b_{k-1},n) and (n,b_k) can be extended to a blue edge by b_k and a_{k-1}, respectively, thus defining a blue Berge-cycle $C_n^{(3)}$. □

Proof of Theorem 1.3. It is enough to prove the theorem for odd r, $r=2t+1$, $t\geqslant 1$. Indeed, since for $r=2t+2$ the same number of colors are used, one can have a color transfer by any injection of the $(2t+1)$-element subsets of $[n]$ into their $2t+2$-element supersets ($n\geqslant 4t+2$ will be ensured). Then the theorem follows from the next proposition.

Proposition 2.1. *If $t\geqslant 1$, $n\geqslant 2t^2-2t+7$, then $R_t(C_n^{(2t+1)})=n$.*

We first prove the following lemma.

Lemma 2.2. *Let c be a fixed positive integer and let $n\geqslant 3c+4$. Then a 3-uniform hypergraph $\mathcal{H}$ of order n with at least $\binom{n}{3}-cn$ edges has a Hamiltonian Berge-cycle.*

Proof. By averaging, there exists a vertex $x\in V(\mathcal{H})$ contained in at least $\binom{n-1}{2}-3c$ triples of $\mathcal{H}$. Each such triple $\{x,y,z\}$ defines an edge yz in a graph G with vertex set $V(\mathcal{H})\setminus\{x\}$. The condition $\binom{n-1}{2}-3c\geqslant\binom{n-2}{2}+2$ (which is equivalent to $n\geqslant 3c+4$) implies that G contains a Hamiltonian cycle $(x_1,\dots,x_{n-1})$ with a chord, say $x_1x_j\in E(G)$ with $j\notin\{2,n-1\}$. This corresponds to the core sequence of a Berge-cycle in $\mathcal{H}$ with edges $\{x_i,x_{i+1},x\}$, $i=1,\dots,n-1$, where $x_n\equiv x_1$. Then the vertex x can be inserted between x_1 and x_2 using the edges $\{x_1,x,x_j\}$ and $\{x,x_2,x_1\}$, thus yielding a Hamiltonian Berge-cycle in $\mathcal{H}$. □

For $S\subseteq V(K_n^{(r)})$, $|S|\leqslant r$, let $E_S=\{e\mid e\in E(K_n^{(r)})\text{ with }S\subseteq e\}$. We shall prove Proposition 2.1 in a stronger form as follows.

Lemma 2.3. *For $t \geqslant 1$, $n \geqslant 2t^2 - 2t + 7$, let $S \subseteq V(K_n^{(2t+1)})$ be a set of even cardinality with $0 \leqslant |S| \leqslant 2t-2$, and color a subset of m edges in E_S with $u = t - |S|/2$ colors. If $m \geqslant \binom{n-|S|}{2t+1-|S|} - (t-u)n > 0$, then E_S contains a monochromatic Hamiltonian Berge-cycle.*

Proof. Let $F_S \subseteq E_S$, $|F_S| = m$, be the set of colored edges in E_S. Fix $t \geqslant 1$. The proof is by induction on u. If $u = 1$, then $|S| = 2t-2$ and since $n \geqslant 2t^2 - 2t$, we obtain that

$$m \geqslant \binom{n-|S|}{2t+1-|S|} - (t-1)n = \binom{n-2t+2}{3} - (t-1)n$$
$$\geqslant \binom{n-2t+2}{3} - t(n-2t+2).$$

Define the 3-uniform hypergraph $\mathcal{H}_S$ with $V(\mathcal{H}_S) = V(K_n^{(2t+1)}) \setminus S$ and $E(\mathcal{H}_S) = \{e \setminus S \mid e \in F_S\}$. Clearly $n - 2t + 2 \geqslant 3t + 4$, thus Lemma 2.2 implies that $\mathcal{H}_S$ contains a Hamiltonian Berge-cycle $C^{(3)}_{n-2t+2}$. Because the corresponding Berge-cycle $C^{(2t+1)}_{n-2t+2}$ in E_S uses only $n - 2t + 2$ edges of F_S, it is easy to extend it by including all vertices of S into a Hamiltonian Berge-cycle $C_n^{(2t+1)}$.

Let $u \geqslant 2$, $|S| = 2t - 2u$, $|F_S| = m \geqslant \binom{n-|S|}{2t+1-|S|} - (t-u)n > 0$, and assume that the theorem holds for $(u-1)$-colorings. Let ℓ be the maximum length of a monochromatic Berge-cycle of F_S. Suppose $\ell < n$, and $C_\ell^{(2t+1)}$ is a maximum Berge-cycle in color 1 with core sequence $(z_1, z_2, \ldots, z_\ell)$. Let $z \in V(K_n^{(2t+1)}) \setminus V(C_\ell^{(2t+1)})$. If there is a j, $1 \leqslant j \leqslant \ell$, such that some $e \in E_{S\cup\{z,z_j\}}$ is in color 1, then by the maximality of ℓ, no edge in $E_{S\cup\{z,z_{j-1}\}} \setminus E(C_\ell^{(2t+1)} + e)$ is colored with 1. Therefore, all but at most $(t-u)n + \ell + 1 \leqslant (t-u+1)n$ edges of $E_{S\cup\{z,z_{j-1}\}}$ are colored with $u-1$ colors. In this case let S' be any set of $|S|+2$ vertices containing $S \cup \{z, z_{j-1}\}$.

If the condition above fails, then for each j, $1 \leqslant j \leqslant \ell$, all but possibly $(t-u)n$ uncolored edges in $E_{S\cup\{z,z_j\}} \setminus E(C_\ell^{(2t+1)})$ are in one of the colors $2, 3, \ldots, u$. In this case let S' be any set of $|S|+2$ vertices containing $S \cup \{z, z_j\}$.

In either case we have $|S'| = 2t - 2(u-1)$, furthermore,

$$m' \geqslant |E_{S'}| - (t-u+1)n = \binom{n-|S'|}{2t+1-|S'|} - \bigl(t-(u-1)\bigr)n > 0$$

edges of $E_{S'}$ are colored with at most $u-1$ colors. By induction, $E_{S'}$ contains a monochromatic Hamiltonian Berge-cycle $C_n^{(2t+1)}$, contradicting the maximality of ℓ. □

The proof of Proposition 2.1 and thus Theorem 1.3 follow applying Lemma 2.3 with $S = \emptyset$. □

3. Almost perfect connected matchings in almost complete hypergraphs

Throughout this section $r \geqslant 2$ is a fixed integer, $0 < \epsilon < 1$ is a fixed and arbitrary small real, and n approaches infinity (thus is arbitrarily large). Hypergraph $\mathcal{H}$ is a $(1-\epsilon)$-complete r-uniform hypergraph on n vertices, i.e. is obtained from $K_n^{(r)}$ by deleting at most $\epsilon\binom{n}{r}$ edges. For easier computation we shall assume that $|E(\mathcal{H})| \geqslant (1-\epsilon)n^r/r!$. A coloring of the edges of $\mathcal{H}$ induces a multicoloring on the shadow graph $\Gamma(\mathcal{H})$ that will be called an *almost complete r-uniform coloring* of $\Gamma(\mathcal{H})$.

Different technical lemmas have been used earlier to handle almost complete graphs and 3-uniform hypergraphs (see [5,10]). We introduce here a tool, the concept of sequential selection, that proves to be convenient for almost complete hypergraphs in general when one needs to show that there exists at least one edge at a prescribed spot or there are many edges where they need to be.

For $0 < \delta < 1$ fixed, we say that a sequence $L \subset V(\mathcal{H})$ of k distinct vertices was obtained by a *δ-bounded selection* (or sometimes we just say shortly that L *is* a δ-bounded selection) if its elements are chosen in k consecutive steps so that in each step there are at most δn forbidden vertices that cannot be included as the next element. These sets of δn forbidden vertices may depend on the choices of the vertices chosen in the previous steps. Observe that a δ-bounded selection L is also a δ'-bounded selection for any $\delta' > \delta$.

In the subsequent applications when specifying a sequential selection of length k, $0 \leqslant k \leqslant r$, we would like to guarantee that at least $(1-\delta)n^{r-k}/(r-k)!$ edges of $\mathcal{H}$ contain the selection. We shall see that this is always possible in a $(1-\epsilon)$-complete hypergraph, because at each step there are at most δn forbidden vertices, where δ depends only on ϵ. For $k = 0$ we need that $\mathcal{H}$ has at least $(1-\delta)n^r/r!$ edges, which is obvious with $\epsilon = \delta$. For larger k our argument will be based on the following recurrence lemma.

Lemma 3.1. *Let $L_0 \subset V(\mathcal{H})$ be contained in at least $(1-\delta_0)\frac{n^{r-|L_0|}}{(r-|L_0|)!}$ edges of $\mathcal{H}$. If $|L_0| < r$ and $\delta = \sqrt{\delta_0}$, then there exists $F_0 \subset V(\mathcal{H})$, $|F_0| \leqslant \delta n$, such that for every $x \in V(\mathcal{H}) \setminus (L_0 \cup F_0)$ at least $(1-\delta)\frac{n^{r-|L|}}{(r-|L|)!}$ edges of $\mathcal{H}$ contain $L = L_0 \cup \{x\}$.*

Proof. Let $|L_0| = i$. By the assumption, there are $\beta \leqslant \delta_0 n^{r-i}/(r-i)!$ distinct $(r-i)$-element "bad" subsets $B \subseteq V(\mathcal{H}) \setminus L_0$ with $L_0 \cup B \notin E(\mathcal{H})$. Let $F_0 \subseteq V(\mathcal{H}) \setminus L_0$ be the set of all vertices contained in more than $\delta n^{r-i-1}/(r-i-1)!$ distinct $(r-i)$-element bad sets. We clearly have $\beta \geqslant |F_0|\delta n^{r-i-1}/(r-i)!$.

By comparing these two bounds on β, we obtain that $|F_0| \leqslant \frac{\delta_0}{\delta} n = \delta n$ and the lemma follows. $\square$

Lemma 3.1 immediately gives the following.

Lemma 3.2. *Assume that $\mathcal{H}$ is a $(1-\epsilon)$-complete r-uniform hypergraph $(r \geqslant 2)$ and set $\delta = \epsilon^{2^{-r}}$. There are forbidden sets such that for every $L \subset V(\mathcal{H})$ of length at most r that was obtained by a δ-bounded selection (with respect to these forbidden sets), at least $(1-\delta)\frac{n^{r-|L|}}{(r-|L|)!}$ edges of $\mathcal{H}$ contain L.*

Proof. Let $\delta_0 = \epsilon$ and let $\delta_{i+1} = \delta_i^{1/2}$, for $i = 0, 1, \ldots, r-1$. By applying successively Lemma 3.1 as an arbitrary sequential selection process, we obtain the forbidden sets $F_i \subset V(\mathcal{H})$, $i = 0, \ldots, r$, such that $|F_i| \leqslant \delta_i n$. Because $\delta_0 < \delta_1 < \cdots < \delta_r$, every δ_k-bounded selection of length k $(0 \leqslant k \leqslant r)$ is contained in at least $(1-\delta_k)n^{r-k}/(n-k)!$ edges of $\mathcal{H}$. Hence every δ_r-bounded selection $L \subset V(\mathcal{H})$ is contained in at least $(1-\delta_r)n^{r-|L|}/(r-|L|)!$ edges of $\mathcal{H}$ and the lemma follows by choosing $\delta = \delta_r = \epsilon^{2^{-r}}$. $\square$

When applying Lemma 3.2, a δ-bounded selection L with $|L| = r$ will be used most of the time, in which case L becomes an edge of $\mathcal{H}$. Throughout this paper we shall use $\delta = \epsilon^{2^{-r}}$ as defined in Lemma 3.2.

A matching in a graph of order n saturating $(1-\delta)n$ vertices is called a $(1-\delta)$-*perfect matching*. In an r-uniform multicoloring of the edges of the shadow graph $\Gamma(\mathcal{H}) = (V, E)$ the set of colors present on $xy \in E$ is denoted by $c(xy)$. If $\chi \in c(xy)$ we say that xy is a χ-edge, or x and y are χ-neighbors. Let E_χ be the set of all χ-edges, and let $G_\chi = (V, E_\chi)$. A matching of G_χ is a *connected matching*, if every matching edge belongs to the same connected component of G_χ. In particular, if G_χ is a connected graph on V, then every matching is automatically a connected one.

Proposition 3.3. *Assume $\mathcal{H}$ is an arbitrary hypergraph and $0 < \delta < 1/3$. It is either possible to delete at most δn vertices from $\mathcal{H}$ so that the remaining hypergraph $\mathcal{H}'$ is connected or the connected components of $\mathcal{H}$ can be partitioned into two groups so that each group contains more than δn vertices.*

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

Proposition 3.4. *An almost complete r-uniform hypergraph has a connected component that admits an almost perfect connected matching in its shadow graph.*

Proof. Choose an ϵ such that $\delta = \epsilon^{2^{-r}} < 1/3$. Let $\mathcal{H}$ be a $(1-\epsilon)$-complete r-uniform hypergraph ($r \geqslant 2$).

Proposition 3.3 is applied to $\mathcal{H}$, it gives two possibilities, we show that the second cannot hold. Indeed, suppose that there is a partition $X \cup Y = V(\mathcal{H})$ of the components of $\mathcal{H}$ such that $|X|, |Y| > \delta n$. Apply Lemma 3.2 and let us consider a δ-bounded selection $L = (x, y)$ such that $x \in X$, $y \in Y$. Since there is an edge $e \in E(\mathcal{H})$ containing L, we obtain $e \cap X \neq \emptyset$, $e \cap Y \neq \emptyset$, a contradiction.

Thus the first possibility holds, so we can delete at most δn vertices from $\mathcal{H}$ so that the remaining hypergraph $\mathcal{H}'$ is connected, so a maximum matching M in $\Gamma(\mathcal{H}')$ is connected. Moreover, M saturates all but at most δn vertices in $\Gamma(\mathcal{H}')$. Indeed, otherwise let $U \subset V(\mathcal{H}')$ be the set of vertices unsaturated by M. Apply Lemma 3.2 with a δ-bounded selection $L = (x_1, x_2) \subset U$. Then there is an edge $e \in E(\mathcal{H})$ with $x_1, x_2 \in e$, thus x_1x_2 is an edge of $\Gamma(\mathcal{H})$, contradicting the maximality of M. Hence $\Gamma(\mathcal{H})$ has a $(1-2\delta)$-perfect connected matching. □

In Section 4 we shall discuss how connected matchings of the shadow graph can be converted into Hamiltonian Berge-cycles. Here we show how to remove the connectivity requirement imposed on the matchings in Conjecture 1.6 by an inductive argument. For a compact formulation, let $S(k, r)$ denote the statement of Conjecture 1.6 with parameters k, r: any r-uniform k-coloring induced on the shadow graph of an almost complete r-uniform hypergraph contains a monochromatic almost perfect connected matching. The statement $S^-(k, r)$ is the weakening of $S(k, r)$ by dropping the connectivity requirement from its conclusion.

Proposition 3.5. *Assume that $1 \leqslant k < r$ and $S(k, r)$, $S^-(k+1, r+1)$ are both true. Then $S(k+1, r+1)$ is also true.*

Proof. To prove $S(k+1, r+1)$, let $\mathcal{H}$ be an almost complete $(r+1)$-uniform hypergraph colored with colors $1, 2, \ldots, k+1$. By the assumption that $S^-(k+1, r+1)$ is true, this coloring admits a monochromatic almost perfect matching M, say in color $k+1$. Let $\mathcal{H}(k+1)$ be the hypergraph determined by the edges of color $k+1$ and apply Proposition 3.3 to it. If the first possibility holds, i.e. $\mathcal{H}(k+1)$ has a connected component C containing all but at most δn vertices, we are done by deleting from M the edges outside C.

If the second possibility holds, the vertex set of $\mathcal{H}(k+1)$ has a partition $X \cup Y$ such that $|X|, |Y| > \delta n$ and every edge $e \in E(\mathcal{H})$ with $e \cap X \neq \emptyset$ and $e \cap Y \neq \emptyset$ has a color different from $k+1$.

Apply Lemma 3.2 to $\mathcal{H}$, and consider the r-uniform hypergraph $\mathcal{H}^*$ defined by the vertex sets of the sequences obtained by δ-bounded selections of r vertices. By Lemma 3.2, each edge f of $\mathcal{H}^*$ is contained in at least $(1-\delta)n^{r+1-r}/(r+1-r)! = (1-\delta)n$ edges of $\mathcal{H}$. Therefore (using also that $|X|, |Y|$ are both larger than δn), there exists $e \in \mathcal{H}$, $f \subset e$ such that e intersect both X and Y. Use the color $\chi \neq k+1$ of e to color f. Since $\mathcal{H}^*$ has at least $(1-\delta)^r n^r/r!$ edges, k-colored and r-uniform, $S(k, r)$ applies to it, giving a monochromatic almost perfect connected matching M in $\Gamma(\mathcal{H}^*)$. To conclude the proof, observe that M is a connected matching in $\Gamma(\mathcal{H})$ as well. $\square$

Due to Proposition 3.5, when looking for almost perfect connected matchings of $\Gamma(\mathcal{H})$ in some color, it is enough to find arbitrary (not necessarily connected) matchings. Next we introduce the concept of a strong transversal that proves to be helpful in the forthcoming investigation.

Assume that $\Gamma(\mathcal{H})$ is colored with $1, 2, \ldots, k$, let $M_i \subseteq V$ be the vertex set saturated by a maximum monochromatic matching $\mathcal{M}_i$ in color i, and set $C_i = V \setminus M_i$, for $i = 1, \ldots, k$. A vertex $x \in M_i$ with at least two i-neighbors in C_i is called *exposed*, otherwise it is *unexposed*. Observe that every edge of $\mathcal{M}_i$ has at least one unexposed vertex otherwise there is an augmenting path of three edges, contradicting the maximality of $\mathcal{M}_i$.

Let $W_i \subseteq M_i$ be the set of all exposed vertices in color i. In the set $S_i = M_i \setminus W_i$ of the unexposed vertices every vertex u has at most one i-neighbor in C_i. If such an i-neighbor $v \in C_i$ of $u \in S_i$ exists, we say that the ordered pair (u, v) is *exceptional* in color i. From the nature of r-uniform colorings, the same ordered pair can be exceptional in many colors. Also, it is quite conceivable that both (u, v) and (v, u) are exceptional (in different colors). Let D be the digraph whose vertex set is V and whose arc set is the set of exceptional ordered pairs. Notice that every vertex of D has outdegree at most k.

The k partitions $V = C_i \cup S_i \cup W_i$, $i = 1, \ldots, k$, decompose V into 3^k atoms, the *atom* $A(x)$ of a vertex $x \in V$ is obtained by specifying for each i which element of $\{C_i, S_i, W_i\}$ contains x. We shall assume that each of these atoms is either empty or has cardinality proportional to n. Otherwise, removing all vertices of a 'small' atom from the shadow graph would reduce its order by $o(n)$, an immaterial change in size when seeking almost perfect matchings.

A set $T \subset V$ is a *strong transversal* (with respect to a k-coloring and a fixed selection of maximum matchings in each color) if for every $i = 1, \ldots, k$, either $|T \cap C_i| \geqslant 2$ or both $T \cap C_i$ and $T \cap S_i$ are non-empty sets. Observe that for $n > 1$ strong transversals exist if (and only if) each C_i is non-empty. Indeed, to define a strong transversal with at most $2k$ elements, simply pick one element from each C_i and from each S_i with $|S_i| > 0$. If $|S_i| = 0$, i.e. color i is not present at all, pick two elements from C_i. The reason for interest in strong transversals is the following lemma.

Lemma 3.6. *Let T be a strong transversal in a k-coloring of an almost complete r-uniform hypergraph $\mathcal{H}$ with n vertices. If each non-empty atom of the coloring has more than $u = r(2k+\delta n)$ vertices, then $|T| > r$.*

Proof. Assume to the contrary that there is a strong transversal T with $t = |T| \leqslant r$. Apply Lemma 3.2 considering a δ-bounded selections to define another strong transversal $L = (x_1, x_2, \dots, x_t)$ containing no arcs from D. The vertices of L will be selected from $Z_1 \supseteq Z_2 \cdots \supseteq Z_t$ defined as follows.

Set $T_1 = T$, $Z_1 = \bigcup_{x\in T_1} A(x) \setminus U_1$ where U_1 is the set of forbidden vertices for x_1 ($|U_1| \leqslant \delta n$). Let x_1 be a vertex of minimum indegree (at most k) in $D[Z_1]$. Assume that T_i, Z_i, x_i are already defined for some $1 \leqslant i < t$. Let $T_{i+1} = T_i \setminus \{y\}$, where $y \in T_i \cap A(x_i)$, and

$$Z_{i+1} = Z_i \cap \left(\bigcup_{x\in T_{i+1}} A(x)\right) \Big\backslash \big(N^+_{Z_i}(x_i) \cup N^-_{Z_i}(x_i) \cup U_{i+1}\big)$$

where U_{i+1} is the set of forbidden vertices for x_{i+1} ($|U_{i+1}| \leqslant \delta n$) and $N^-_{Z_i}, N^+_{Z_i}$ are the set of endpoints of incoming and outgoing arcs of x_i in the subgraph $D[Z_i]$. Let x_{i+1} be a vertex of minimum indegree (at most k) in $D[Z_{i+1}]$.

For each i, $i = 0, 1, \dots, i$, the outdegree of vertex x_{i+1} is at most k in $D[Z_{i+1}]$ (at most one in each color) and its indegree is also at most k in $D[Z_{i+1}]$ since a vertex of minimum indegree was selected. Hence in each step, before selecting the next element into L, at most $2k + \delta n$ vertices are removed from any atom. The number of steps is $|L| = t \leqslant r$ and initially each atom has more than $u = r(2k + \delta n)$ vertices. Thus Z_t is non-empty so L is well defined.

Observe that no arc of D is contained in L. Indeed, consider $x_j, x_i \in L$, $1 \leqslant j < i \leqslant t$ and notice that no vertex of $N^+_{Z_j}(x_j) \cup N^-_{Z_j}(x_j)$ is in Z_{j+1} so not in $Z_i \subseteq Z_{j+1}$.

The set L is a strong transversal because its vertices are selected from the same set of atoms (with the same multiplicity) as the set of atoms containing T.

Notice that $L \subseteq e$ for some $e \in E(\mathcal{H})$. If χ is the color of e, then each pair $x, y \in L$ is a χ-edge of $\Gamma(\mathcal{H})$. Since L is a strong transversal, either $|L \cap C_\chi| \geqslant 2$ or both $L \cap C_\chi$ and $L \cap S_\chi$ are non-empty. In the first case if $x, y \in L \cap C_\chi$, then the χ-edge xy extends M_χ contradicting the maximality of M_χ. In the second case let $x \in S_\chi$ and $y \in C_\chi$. Because there is at most one χ-edge from the unexposed vertex $x \in S_\chi$ to C_χ, the arc xy must belong to D. However, no ordered pair of L appears in D, a contradiction. $\square$

We prove Conjecture 1.6 next for $r = 3$. Although a direct simpler proof is possible, we prove it with the method that will be applied for the case $r = 4$.

Theorem 3.7. *Every 2-coloring of an almost complete 3-uniform hypergraph $\mathcal{H}$ admits a monochromatic almost perfect connected matching in $\Gamma(\mathcal{H})$.*

Proof. Using Proposition 3.4 with $r = 2$ and Proposition 3.5, it is enough to find an almost perfect monochromatic matching. To be able to apply Lemma 3.6, we first delete all (non-empty) atoms of size at most $3(4 + \delta n)$ from $V(\Gamma(\mathcal{H}))$—for convenience, we keep all notation for the remaining set of vertices. We may assume that the sets C_i, $i = 1, 2$, are non-empty—consequently large, say $|C_i| > \delta n$—since otherwise, there is an almost perfect matching in some color. Notice that the existence of a strong transversal T with at most three vertices contradicts Lemma 3.6 finishing the proof.

We may also assume that no S_i, $i=1,2$, is empty (consequently each is large). Assume not, say S_1 is empty. This means that no edge of $\Gamma(\mathcal{H})$ is colored by color 1, so C_1 contains all vertices (but at most one). Therefore picking two vertices from C_2 we have a strong transversal—contradiction.

Observe that $|C_i \cup S_i| \geqslant |C_i| + |M_i|/2 = |C_i| + |V \setminus C_i|/2 > n/2$. Thus there exists a vertex $x \in (C_1 \cup S_1) \cap (C_2 \cup S_2)$. Then for distinct vertices $y \in C_1$, $z \in C_2$ different from x, $T = \{x, y, z\}$ is a strong transversal, contradiction. □

Theorem 3.8. *Every* 3*-coloring of an almost complete* 4*-uniform hypergraph* $\mathcal{H}$ *admits a monochromatic almost perfect connected matching in* $\Gamma(\mathcal{H})$.

Proof. We use notation already introduced above. We shall follow the argument used in the proof of Theorem 3.7. By Theorem 3.7 and Proposition 3.5, it is enough to find an almost perfect (not necessarily connected) monochromatic matching. To be able to apply Lemma 3.6, we first delete all (non-empty) atoms of size at most $4(6+\delta n)$ from $V(\Gamma(\mathcal{H}))$—for convenience, we keep all notation for the remaining set of vertices. We may assume that the sets C_i, $i=1,2,3$ are non-empty—consequently large, say $|C_i| > \delta n$—since otherwise, there is an almost perfect matching in some color. Notice that the existence of a strong transversal T with at most four vertices contradicts Lemma 3.6 finishing the proof.

We may also assume that no S_i, $i=1,2,3$ is empty (consequently each is large). Assume not, say S_1 is empty. This means that no edge of $\Gamma(\mathcal{H})$ is colored by color 1, so C_1 contains all vertices (but at most one). Therefore picking two vertices from C_2 and two vertices from C_3 we have a strong transversal, a contradiction.

Observe that $|C_i \cup S_i| \geqslant |C_i| + |M_i|/2 = |C_i| + |V \setminus C_i|/2 > n/2$. As a corollary we obtain that any two sets among $C_i \cup S_i$, $i=1,2,3$, have a common vertex.

Case 1. $C_i \cap C_j \neq \emptyset$, for some $1 \leqslant i < j \leqslant 3$. Let k be the third index different from i and j. If $x, y \in C_i \cap C_j$, and $v, w \in C_k$, then $T = \{x, y, v, w\}$ is a strong transversal, a contradiction.

Case 2. $C_i \cap S_j \neq \emptyset$ and $C_j \cap S_k \neq \emptyset$ for $\{i, j, k\} = \{1, 2, 3\}$. If $y, z \in C_i \cap S_j$, $x \in C_j \cap S_k$, and $w \in C_k$, then $T = \{x, y, z, w\}$ is a strong transversal, a contradiction.

Case 3. $C_i \cap S_j \neq \emptyset$, for some $1 \leqslant i < j \leqslant 3$, and none of the previous cases applies. Let k be the third index different from i and j. Then there is a vertex $y \in S_i \cap S_k$ or a vertex $y \in C_i \cap S_k$. Then let $x \in C_i \cap S_j$, $v \in C_j$ and $w \in C_k$. In both cases $T = \{x, y, v, w\}$ is a strong transversal, a contradiction.

Case 4. $S_i \cap S_j \neq \emptyset$, for all $1 \leqslant i, j \leqslant 3$. Assume in addition that none of the previous cases applies, in particular, $C_i \cap C_j = \emptyset$ and $C_i \cap S_j = \emptyset$, for all $1 \leqslant i, j \leqslant 3$.

For every i, j, $1 \leqslant i < j \leqslant 3$, define $U_{ij} = S_i \cap S_j$, and for the third index k, let $U_k = S_k \setminus (S_i \cup S_j)$. Observe that $S_1 \cap S_2 \cap S_3 = \emptyset$, since otherwise, if $x \in S_1 \cap S_2 \cap S_3$ and $y_i \in C_i$, then $T = \{x, y_1, y_2, y_3\}$ is a strong transversal. Then it follows that

$$S_1 = U_{12} \cup U_{13} \cup U_1,$$

$$S_2 = U_{12} \cup U_{23} \cup U_2,$$

$$S_3 = U_{13} \cup U_{23} \cup U_3.$$

Notice that distinct sets in the right-hand side are pairwise disjoint, and also disjoint from each of C_1, C_2, C_3. Let $U_0 = V \setminus \bigcup_{i=1}^{3}(C_i \cup S_i)$.

In terms of the atoms introduced above the sets of exposed vertices are partitioned as follows:

$$W_1 = M_1 \setminus S_1 = C_2 \cup C_3 \cup U_{23} \cup U_2 \cup U_3 \cup U_0,$$
$$W_2 = M_2 \setminus S_2 = C_1 \cup C_3 \cup U_{13} \cup U_1 \cup U_3 \cup U_0,$$
$$W_3 = M_3 \setminus S_3 = C_1 \cup C_2 \cup U_{12} \cup U_1 \cup U_2 \cup U_0.$$

Let wx be an edge of $\mathcal{M}_1$ with $w \in W_1$. Then $x \in S_1$, therefore $|S_1| = |U_{12}| + |U_{13}| + |U_1| \geqslant |W_1|$. We strengthen this inequality as follows. Let $A_{12} \subseteq U_{12}$, $A_{13} \subseteq U_{13}$ denote the set of vertices matched from W_1 by $\mathcal{M}_1$. We claim that at least one of the sets A_{12}, A_{13} is small, has at most δn vertices. Suppose this is not the case. Then we can apply Lemma 3.2 with a δ-bounded selection of vertices, (x_2, x_3, y_2, y_3), such that $x_2 \in A_{12}$, $x_3 \in A_{13}$, $y_2 \in C_2$, $y_3 \in C_3$ and y_j is not the exceptional j-neighbor of x_j in C_j. Since $Q = \{x_2, x_3, y_2, y_3\} \in E(\mathcal{H})$, Q has some color. But $2 \notin c(x_2y_2)$ and $3 \notin c(x_3y_3)$, so Q is colored with color 1, consequently $1 \in c(x_1x_2)$. Let z_1, z_2 be the other endpoints of the edges of $\mathcal{M}_1$ containing x_1, x_2, respectively. Now there is an augmenting path of length five in color 1; by the definition of W_1, one can select two vertices, $p, q \in C_1$ such that $1 \in c(pz_1)$, $1 \in c(qz_2)$. Thus replacing $\{z_1x_1, z_2x_2\}$ by $\{pz_1, x_1x_2, qz_2\}$ we contradict to the maximality of M_1. A similar argument holds for each color.

This implies that for color 1, either

(1) $|U_{13}| + |U_1| + \delta n \geqslant |W_1| = |C_2| + |C_3| + |U_{23}| + |U_2| + |U_3| + |U_0|$

or

(2) $|U_{12}| + |U_1| + \delta n \geqslant |C_2| + |C_3| + |U_{23}| + |U_2| + |U_3| + |U_0|$.

Similarly, for color 2, either

(3) $|U_{12}| + |U_2| + \delta n \geqslant |C_1| + |C_3| + |U_{13}| + |U_1| + |U_3| + |U_0|$

or

(4) $|U_{23}| + |U_2| + \delta n \geqslant |C_1| + |C_3| + |U_{13}| + |U_1| + |U_3| + |U_0|$.

Furthermore, for color 3 either

(5) $|U_{23}| + |U_3| + \delta n \geqslant |C_1| + |C_2| + |U_{12}| + |U_1| + |U_2| + |U_0|$

or

(6) $|U_{13}| + |U_3| + \delta n \geqslant |C_1| + |C_2| + |U_{12}| + |U_1| + |U_2| + |U_0|$.

Without loss of generality we may assume that (1) is true. This inequality contradicts inequality (4), because their combination results in

$$|U_{13}| + |U_1| + \delta n \geqslant |C_2| + |C_3| + \big(|U_{23}| + |U_2|\big) + |U_3| + |U_0|$$

$$\geqslant |C_2| + |C_3| + \big(|C_1| + |C_3| + |U_{13}| + |U_1| + |U_3| + |U_0|\big) + |U_3| + |U_0| - \delta n$$
$$\geqslant |U_{13}| + |U_1| + |C_1| + |C_2| + 2|C_3| - \delta n > |U_{13}| + |U_1| + 3\delta n.$$

Thus (3) must be true (the last inequality follows from $|C_i| > \delta n$). A similar argument excludes (6) and implies that (5) is true. Then the sum of the inequalities (1), (3), and (5) leads to an obvious contradiction. This concludes Case 4 and the proof of the theorem. □

Theorem 3.9. *Let k be the largest integer such that $k + \lfloor \log_2(k+1) \rfloor \leqslant r$. Then every k-coloring of an almost complete r-uniform hypergraph $\mathcal{H}$ admits a monochromatic almost perfect connected matching in its shadow graph.*

Proof. We use notation already introduced. Again, by Proposition 3.5, we do not have to prove the connectivity of the matching. (Notice that the inequality $k + \lfloor \log_2(k+1) \rfloor \leqslant r$ is trivially inherited from the pair $(k+1, r+1)$ to the pair (k, r) and one can start the induction from Theorem 3.8 with $(3, 5)$ or trivially with $(1, 3)$.) We may delete vertices of atoms of order at most $u = r(2k + \delta n)$ and can assume that in the remaining atoms all sets C_i are represented—otherwise we have the required almost perfect matching. We show that there is a strong transversal of at most $k + \lfloor \log_2(k+1) \rfloor \leqslant r$ vertices, contradicting Lemma 3.6.

Observe that $|C_i \cup S_i| > n/2$, $1 \leqslant i \leqslant k$ (if $|C_i \cup S_i| = n/2$, we have nothing to prove, M_i spans all vertices). Therefore $\sum_{i=1}^{k} |C_i \cup S_i| > nk/2$ showing that some vertex $v_1 \in V$ is in more than $k/2$ of the sets $V_i = C_i \cup S_i$. Repeating the argument with the sets V_i that do not contain v_1, one can eventually obtain a set $T = \{v_1, v_2, \dots, v_l\}$ such that $l \leqslant \lfloor \log_2(k+1) \rfloor$ and $T \cap V_i \neq \emptyset$, $1 \leqslant i \leqslant k$. Then T can be extended to a strong transversal T^* by adding at most k vertices to T; for every i, add a vertex of C_i or S_i to T selecting the one that has empty intersection with T (if both intersect T we do not select anything). □

One easy consequence of Theorem 3.9 is the following.

Corollary 3.10. *If $k = r - \lfloor \log_2 r \rfloor$ then every k-coloring of an almost complete r-uniform hypergraph $\mathcal{H}$ admits a monochromatic almost perfect connected matching in its shadow graph.*

4. From connected matchings to Berge-cycles

Here we show how to transform our asymptotic results on monochromatic connected matchings to asymptotic results on monochromatic Hamiltonian Berge-cycles with the use of the hypergraph version of the Regularity Lemma of Szemerédi [18]. This approach has also been used in [9,10]. We shall assume throughout the rest of the paper that n is sufficiently large and k and r are fixed.

There are several generalizations of the Regularity Lemma for hypergraphs due to various authors ([1,3], for an extensive survey see [15], new developments are in [4,17,19]). Here we will use the simplest one due to Chung [1]. First we need to define the notion of ε-regularity. Let $\varepsilon > 0$ and let $V_1, V_2, \dots, V_r$ be disjoint vertex sets of order m, and let $\mathcal{H}$ be an r-uniform hypergraph such that every edge of $\mathcal{H}$ contains exactly one vertex from each V_i for $i = 1, 2, \dots, r$. The density of $\mathcal{H}$ is $d_{\mathcal{H}} = \frac{|E(\mathcal{H})|}{m^r}$. The r-tuple $\{V_1, V_2, \dots, V_r\}$ is called an $(\varepsilon, \mathcal{H})$-regular r-tuple of density $d_{\mathcal{H}}$ if for every choice of $X_i \subset V_i$, $|X_i| > \varepsilon |V_i|$, $i = 1, 2, \dots, r$ we have

$$\left| \frac{|E(\mathcal{H}[X_1, \dots, X_r])|}{|X_1| \cdots |X_r|} - d_{\mathcal{H}} \right| < \varepsilon.$$

Here we denote by $\mathcal{H}[X_1, \dots, X_r]$ the subhypergraph of $\mathcal{H}$ induced by the vertex set $X_1 \cup \dots \cup X_r$. In this setting the k-color version of the Hypergraph Regularity Lemma from [1] can be stated as follows.

Lemma 4.1 *(k-color Weak Hypergraph Regularity Lemma). For every positive ε and positive integers t, r, k there are positive integers M and n_0 such that for $n \geqslant n_0$ the following holds. For all r-uniform hypergraphs $\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_k$ with $V(\mathcal{H}_1) = V(\mathcal{H}_2) = \dots = V(\mathcal{H}_k) = V$, $|V| = n$, there is a partition of V into $l+1$ classes (clusters)*

$$V = V_0 + V_1 + V_2 + \dots + V_l$$

such that

- $t \leqslant l \leqslant M$,
- $|V_1| = |V_2| = \dots = |V_l|$,
- $|V_0| < \varepsilon n$,
- *apart from at most $\varepsilon\binom{l}{r}$ exceptional r-tuples, the r-tuples $\{V_{i_1}, V_{i_2}, \dots, V_{i_r}\}$ are $(\varepsilon, \mathcal{H}_s)$-regular for $s = 1, 2, \dots, k$.*

Consider a k-edge coloring $(\mathcal{H}_1, \mathcal{H}_2, \dots, \mathcal{H}_k)$ of the r-uniform complete hypergraph $K_n^{(r)}$, i.e. $\mathcal{H}_1$ is the subhypergraph induced by the first color, $\mathcal{H}_2$ is the subhypergraph induced by the second color, etc. $\mathcal{H}_k$ is the subhypergraph induced by the kth color.

We apply the above k-color Weak Hypergraph Regularity Lemma with $t = r$ and with a small enough ε to obtain a partition of $V(K_n^{(r)}) = V = \bigcup_{0 \leqslant i \leqslant l} V_i$, where $|V_i| = \frac{n-|V_0|}{l} = m$, $1 \leqslant i \leqslant l$. We define the following *reduced hypergraph* $\mathcal{H}^R$: The vertices of $\mathcal{H}^R$ are $p_1, \dots, p_l$, and we have an r-edge on vertices $p_{i_1}, p_{i_2}, \dots, p_{i_r}$ if the r-tuple $\{V_{i_1}, V_{i_2}, \dots, V_{i_r}\}$ is $(\varepsilon, \mathcal{H}_s)$-regular for $s = 1, 2, \dots, k$. Thus we have a one-to-one correspondence $f: p_i \to V_i$ between the vertices of $\mathcal{H}^R$ and the clusters of the partition. Then,

$$\left|E(\mathcal{H}^R)\right| \geqslant (1-\varepsilon)\binom{l}{r},$$

and thus $\mathcal{H}^R$ is a $(1-\varepsilon)$-complete r-uniform hypergraph on l vertices. Define a k-edge coloring $(\mathcal{H}_1^R, \mathcal{H}_2^R, \dots, \mathcal{H}_k^R)$ of $\mathcal{H}^R$ with the majority color, i.e. the r-tuple $\{p_{i_1}, p_{i_2}, \dots, p_{i_r}\} \in E(\mathcal{H}_s^R)$ if s is the most frequent color in the r-tuple $\{V_{i_1}, V_{i_2}, \dots, V_{i_r}\} \in E(\mathcal{H}_s)$. Note then that the density of this color is $\geqslant 1/k$ in this r-tuple. Finally we consider the multicolored shadow graph $\Gamma(\mathcal{H}^R)$. The vertices are $V(\mathcal{H}^R) = \{p_1, \dots, p_l\}$ and we join vertices x and y by an edge of color s, $s = 1, 2, \dots, k$ 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 one.

Lemma 4.2. *Assume that for some positive constant c we can find a monochromatic connected matching M saturating at least cl vertices in $\Gamma(\mathcal{H}^R)$. Then in the original k-edge colored $K_n^{(r)}$ we can find a monochromatic Berge-cycle of length at least $c(1-3\varepsilon)n$.*

We here again note that the use of a connected matching in this type of proof (first suggested by [16]) has become somewhat standard by now (see [2,5–8]).

Proof. We may assume that M is in $\Gamma(\mathcal{H}_1^R)$. Denote the edges of M by $M = \{e_1, e_2, \dots, e_{l_1}\}$ and thus $2l_1 \geqslant cl$. Furthermore, write $f(e_i) = (V_1^i, V_2^i)$ for $1 \leqslant i \leqslant l_1$ where V_1^i, V_2^i are the clusters assigned to the end points of e_i.

Next we define *good* vertices for an arbitrary edge e in $\Gamma(\mathcal{H}_1^R)$. Let $f(e)$ be denoted by (V^1, V^2). Since e is an edge in $\Gamma(\mathcal{H}_1^R)$, the endpoints of e are contained in an r-edge E in $\mathcal{H}_1^R$. By definition this r-edge corresponds to an $(\varepsilon, \mathcal{H}_1)$-regular r-tuple $f(E)$ (containing clusters V^1, V^2 and $r-2$ more clusters) that has density $\geqslant 1/k$. 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 $m/2k$ vertices $y \in V^{j'}$, there are at least $m^{r-2}/2k$ r-edges in $\mathcal{H}_1[f(E)]$ containing x and y. The next claim shows that most vertices are good in each V^j.

Claim 1. *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 r-edges in $\mathcal{H}_1[f(E)]$ that contain a vertex from X is smaller than

$$|X|\left(\frac{m}{2k}m^{r-2} + \left(1 - \frac{1}{2k}\right)m\frac{m^{r-2}}{2k}\right) = \left(\frac{1}{k} - \frac{1}{4k^2}\right)|X|m^{r-1}, \tag{1}$$

which contradicts the fact that $f(E)$ is $(\varepsilon, \mathcal{H}_1)$-regular with density at least $1/k$ when ε is small. 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

$$\geqslant (1/2k - \varepsilon)m > m/4k \tag{2}$$

vertices $y \in V^{j'}$ such that there are at least $m^{r-2}/2k$ r-edges in $\mathcal{H}_1[f(E)]$ containing x and y.

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

$$\left|E_G(X, Y)\right| > \delta|X||Y|,$$

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 difficult to see that the following is true.

Claim 2. *$G(V^1, V^2)$ is a $(2\varepsilon, 1/4k, G)$-super-regular bipartite graph.*

Indeed, the second condition of super-regularity follows from (2). 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 $E_G(X, Y) \leqslant |X||Y|/4k$. The total number of r-edges in $\mathcal{H}_1[f(E)]$ that contain a vertex from X and a vertex from Y is smaller than

$$|X||Y|\left(\frac{m^{r-2}}{4k} + \left(1 - \frac{1}{4k}\right)\frac{m^{r-2}}{2k}\right) < \frac{3}{4k}|X||Y|m^{r-2}, \tag{3}$$

which again contradicts the fact that $f(E)$ is $(\varepsilon, \mathcal{H}_1)$-regular with density at least $1/k$. 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 $f^{-1}(V_1^{i+1})$ for every $1 \leqslant i \leqslant l_1$ (for $i = l_1$ set $i+1 = 1$). 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_1^{i+1} of V_1^{i+1} that is good for V_2^{i+1}. More precisely we construct P_1 with the following simple greedy strategy. Let $P_1^R = (p_1, \dots, p_t)$, $2 \leqslant t \leqslant l$, where according to the definition $f(p_1) = V_2^1$ and $f(p_t) = V_1^2$. 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_2^1 and $f(p_2)$. By Claim 1 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 1 and the fact that ε 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_t $(= v_1^2)$ of P_1 is a vertex $u_t \in (f(p_t) \cap N_{G(f(p_{t-1}), f(p_t))}(u_{t-1}))$ that is good for V_2^2.

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 already occurred on the path P_1^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 $\Gamma(\mathcal{H}_1)$, $1 \leqslant i \leqslant l_1$. Next we have to make these connecting paths Berge-paths. By the construction, since every edge on every path P_i, $1 \leqslant i \leqslant l_1$ came from an appropriate bipartite graph G, the two endpoints of every edge are contained in at least $m^{r-2}/2k$ r-edges in $\mathcal{H}_1[f(E)]$. Since the total number of edges on the paths P_i is a constant $(\leqslant l^2)$ and n (and thus m) is sufficiently large, we can clearly "assign" an r-edge from $\mathcal{H}_1$ for each edge on the paths such that the assigned r-edge contains the corresponding edge and the assigned r-edges of $\mathcal{H}_1$ 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 r-edges from $\mathcal{H}_1$ that are assigned to the edges of the paths P_i, since these r-edges 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_1^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 $2\varepsilon m$ vertices from each cluster V_j^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_j^i. Note that by Claim 2 the remaining bipartite graph $G(V_1^i, V_2^i)$ is clearly still $(4\varepsilon, 1/8k, G)$-super-regular for every $1 \leqslant i \leqslant l_1$ and now we have $|V_1^i| = |V_2^i|$.

We will use the following property of (ε, δ, G)-super-regular pairs.

Lemma 4.3. *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 $|V_1| = |V_2| = m \geqslant m_0$, and let the pair (V_1, V_2) be (ε, δ, 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 4.3 is used by Łuczak in [16]. Lemma 4.3 is a special case of the much stronger Blow-up Lemma (see [13] and [14]). Note that an easier approximate version of this lemma would suffice as well, but for simplicity we use this lemma.

Applying Lemma 4.3 inside each $G(V_1^i, V_2^i)$, $1 \leqslant i \leqslant l_1$ together with the connecting paths P_i we get a cycle C in $\Gamma(\mathcal{H}_1)$ that has length at least

$$cl(1-2\varepsilon)m \geqslant c(1-\varepsilon)(1-2\varepsilon)n \geqslant c(1-3\varepsilon)n.$$

We only have to make this cycle the core of a Berge-cycle. For the edges on the connecting paths P_i we already have assigned distinct r-edges of $\mathcal{H}_1$. The other edges came from the bipartite graphs $G(V_1^i, V_2^i)$, $1 \leqslant i \leqslant l_1$, and thus the two endpoints of every edge are contained in at least $m^{r-2}/4k$ (we already removed some r-edges) r-edges in $\mathcal{H}_1[f(E_i)]$ (here E_i denotes the r-edge in $\mathcal{H}_1^R$ containing the endpoints of the edge e_i). For $r=2$ we are done. For $r=3$ note that the triples E_i must be distinct for each i, $1 \leqslant i \leqslant l_1$ and furthermore the triples containing two distinct edges from $G(V_1^i, V_2^i)$ are distinct. Hence for $r=3$ we can clearly assign distinct triples to each edge on C. For $r>3$ note that an r-edge E_i can be the same only for at most $\lfloor r/2 \rfloor$ values of i. At most $\lfloor r/2 \rfloor 2m \leqslant rm$ edges of C come from these values of i. Furthermore, for $r>3$ the two endpoints of every edge in $G(V_1^i, V_2^i)$ are contained in at least $m^2/4k$ r-edges in $\mathcal{H}_1[f(E_i)]$. Thus if m is sufficiently large (and thus $rm \ll m^2/4k$) we can clearly assign distinct r-edges to each edge on C and this makes the cycle C the core of a Berge-cycle, completing the proof of Lemma 4.2. □

Lemma 4.2 together with the asymptotic results of the previous section on monochromatic connected matchings give the results on monochromatic Berge-cycles stated as Theorems 1.4 and 1.5. For example, to prove Theorem 1.4, observe that by Theorem 3.8, for any given $\eta > 0$ there is a small enough ϵ such that every 3-coloring of a 4-uniform $(1-\epsilon)$-complete $\mathcal{H}$ admits a monochromatic connected matching in $\Gamma(\mathcal{H})$ covering at least $(1-\eta)|V(\Gamma(\mathcal{H}))|$ vertices. Applying this to $\mathcal{H} = \mathcal{H}^{\mathcal{R}}$, i.e. to the reduced hypergraph of $K_n^{(4)}$, we get a monochromatic connected matching covering at least $(1-\eta)l$ vertices of $\Gamma(\mathcal{H}^{\mathcal{R}})$. Then Lemma 4.2 gives a Berge-cycle of length at least $(1-3\epsilon)(1-\eta)n$ in $K_n^{(4)}$.

Acknowledgment

The authors appreciate the useful comments of the referees.

References

[1] F. Chung, Regularity lemmas for hypergraphs and quasi-randomness, Random Structures Algorithms 2 (1991) 241–252.

[2] A. Figaj, T. Łuczak, The Ramsey number for a triple of long even cycles, J. Combin. Theory Ser. B 97 (2007) 584–596.

[3] P. Frankl, V. Rödl, The uniformity lemma for hypergraphs, Graphs Combin. 8 (1992) 309–312.

[4] W.T. Gowers, Hypergraph regularity and the multidimensional Szemerédi Theorem, preprint.

[5] A. Gyárfás, M. Ruszinkó, G.N. Sárközy, E. Szemerédi, Three-color Ramsey numbers for paths, Combinatorica 27 (2007) 35–69.

[6] A. Gyárfás, M. Ruszinkó, G.N. Sárközy, E. Szemerédi, Tripartite Ramsey numbers for paths, J. Graph Theory 55 (2007) 164–174.

[7] A. Gyárfás, M. Ruszinkó, G.N. Sárközy, E. Szemerédi, An improved bound for the monochromatic cycle partition number, J. Combin. Theory Ser. B 96 (2006) 855–873.

[8] A. Gyárfás, M. Ruszinkó, G.N. Sárközy, E. Szemerédi, One-sided coverings of colored complete bipartite graphs, in: M. Klazar, et al. (Eds.), Topics in Discrete Mathematics (dedicated to J. Nesetril on his 60th birthday), in: Algorithms and Combin., vol. 26, Springer, Berlin, 2006, pp. 133–154.
[9] A. Gyárfás, G.N. Sárközy, The three-color Ramsey number of a 3-uniform Berge cycle, submitted for publication.
[10] P. Haxell, T. Łuczak, Y. Peng, V. Rödl, A. Ruciński, M. Simonovits, J. Skokan, The Ramsey number for hypergraph cycles I, J. Combin. Theory Ser. A 113 (2006) 67–83.
[11] P. Haxell, T. Łuczak, Y. Peng, V. Rödl, A. Ruciński, J. Skokan, The Ramsey number for hypergraph cycles II, manuscript.
[12] Y. Kohayakawa, M. Simonovits, Skokan, The 3-color Ramsey number of odd cycles, manuscript.
[13] J. Komlós, G.N. Sárközy, E. Szemerédi, Blow-up Lemma, Combinatorica 17 (1997) 109–123.
[14] J. Komlós, G.N. Sárközy, E. Szemerédi, An algorithmic version of the Blow-up Lemma, Random Structures Algorithms 12 (1998) 297–312.
[15] J. Komlós, M. Simonovits, Szemerédi's Regularity Lemma and its applications in graph theory, in: D. Miklós, V.T. Sós, T. Szőnyi (Eds.), Combinatorics, Paul Erdős is Eighty, Budapest, in: Bolyai Soc. Math. Stud., vol. 2, 1996, pp. 295–352.
[16] T. Łuczak, $R(C_n, C_n, C_n) \leqslant (4 + o(1))n$, J. Combin. Theory Ser. B 75 (1999) 174–187.
[17] V. Rödl, J. Skokan, Regularity Lemma for uniform hypergraphs, Random Structures Algorithms 25 (1) (2004) 1–42.
[18] E. Szemerédi, Regular partitions of graphs, in: Problèmes Combinatoires et Théorie des Graphes, Coll. Int. C.N.R.S. Orsay (1976) 399–401.
[19] T. Tao, A variant of the hypergraph removal lemma, J. Combin. Theory Ser. A 113 (2006) 1257–1280.