Combinatorics, Probability and Computing (2011) 20, 53–71. © Cambridge University Press 2010
doi:10.1017/S0963548310000209

The 3-Colour Ramsey Number of a 3-Uniform Berge Cycle

ANDRÁS GYÁRFÁS[1†] and GÁBOR N. SÁRKÖZY[1,2‡]

[1]Computer and Automation Research Institute, Hungarian Academy of Sciences, Budapest, PO Box 63, Budapest, Hungary, H-1518
(e-mail: gyarfas@sztaki.hu)

[2]Computer Science Department, Worcester Polytechnic Institute, Worcester, MA 01609, USA
(e-mail: gsarkozy@cs.wpi.edu)

Received 31 July 2007; revised 20 May 2010; first published online 2 July 2010

The asymptotics of 2-colour Ramsey numbers of loose and tight cycles in 3-uniform hypergraphs were recently determined [16, 17]. We address the same problem for Berge cycles and for 3 colours. Our main result is that the 3-colour Ramsey number of a 3-uniform Berge cycle of length n is asymptotic to $\frac{5n}{4}$. The result is proved with the Regularity Lemma via the existence of a monochromatic connected matching covering asymptotically $4n/5$ vertices in the multicoloured 2-shadow graph induced by the colouring 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 for two colours and for paths, were looked at some years later in [8], and for three colours and for paths and cycles in [5], [14] and [19].

There are several possible ways to define paths and cycles in hypergraphs. In this paper we address the case of the *Berge cycle*; the earliest definition of a cycle in hypergraphs is probably the one in the book by Berge [1]. Turán-type problems for Berge paths and Berge cycles of hypergraphs perhaps made their first appearance in [2]. Other types of hypergraph cycles, *loose* and *tight*, were studied in [18] and [27]. Investigations of the corresponding Ramsey problems began quite recently with [16] and [17], where Ramsey numbers of loose and tight cycles were determined asymptotically for two colours and for 3-uniform hypergraphs.

† Research supported in part by OTKA grant K68322.
‡ Research supported in part by the National Science Foundation under grant DMS-0456401, by OTKA grant K68322 and by a János Bolyai Research Scholarship.

Let $\mathcal{H}$ be a 3-uniform hypergraph (3-element subsets of a set). For vertices $x, y \in V(\mathcal{H})$ we say that x is adjacent to y if there exists an edge $e \in E(\mathcal{H})$ such that $x, y \in e$. Let $K_n^{(3)}$ denote the complete 3-uniform hypergraph on n vertices. A 3-uniform ℓ-cycle, or *Berge cycle* of length ℓ, denoted by $C_\ell^{(3)}$, 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 $i, 1 \leqslant i \leqslant \ell$, where $v_{\ell+1} := v_1$. When 3-uniformity is clearly understood we may simply write C_ℓ for $C_\ell^{(3)}$. It is important to keep in mind that a 3-uniform Berge cycle C_ℓ is not determined uniquely: it is considered to be 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 colours. 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-colouring of $K_n^{(3)}$ (see [11]). In this paper we explore the 3-colour Ramsey number of a Berge cycle in 3-uniform hypergraphs. Our main result is that $R_3(C_n) = (1 + o(1))\frac{5n}{4}$; as far as we know, this is the first 3-colour Ramsey-type result for cycles in hypergraphs. It seems purely incidental that our result has the same asymptotics as the 2-colour Ramsey number of the loose n-cycle in 3-uniform hypergraphs (see [16]).

Theorem 1.1. *For all $\eta > 0$ there exists n_0 such that for every $n > n_0$, every colouring of the edges of $K_n^{(3)}$ with 3 colours contains a monochromatic Berge cycle of length at least $(\frac{4}{5} - \eta)n$.*

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

Theorem 1.2. *For all $\eta > 0$ there exists n_0 such that for every $n > n_0$, we have the following:*

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

Perhaps Theorem 1.1 can be extended as follows.

Conjecture 1.3. *For all $\eta > 0$ and positive integer r, there exists $n_0 = n_0(\eta, r)$ such that for every $n > n_0$, every colouring of the edges of $K_n^{(r)}$ with r colours contains a monochromatic Berge cycle of length at least $(\frac{2r-2}{2r-1} - \eta)n$.*

Conjecture 1.3 (and thus Theorem 1.1) is asymptotically best possible, 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 coloured with colour 1. The r-edges that are not yet coloured and do not contain a vertex from A_2 are coloured with colour 2. We continue in this fashion. Finally, the r-edges that are not yet coloured with colours $1, \ldots, r-2$ and do not contain a vertex from A_{r-1} are coloured with colour $r-1$. The r-edges that contain a vertex from all $r-1$ sets $A_1, \ldots, A_{r-1}$ get colour r. We claim that in this r-colouring of the edges of

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

The proofs of Theorems 1.1 and 1.2 use the following approach. For a given 3-uniform hypergraph $\mathcal{H}$, consider the 2-*shadow* (or simply shadow) graph $\Gamma(\mathcal{H})$ on the same vertex set, with edge $(x, y) \in E(\Gamma(\mathcal{H}))$ if and only if x, y is covered by some hyperedge. To a given 3-colouring of the edges of the 3-uniform hypergraph associate an edge multicolouring of the shadow graph by colouring each edge with all colours appearing on hyperedges containing that pair. Edge (multi-)colourings of $\Gamma(\mathcal{H})$ defined in this way will be called 3-*uniform colourings* of $\Gamma(\mathcal{H})$.

Then, following the method established in [24] and refined later in several papers (see [5], [14], [15], [12], [13], [16], [17] and [19]), Theorems 1.1 and 1.2 can be reduced to finding a large (of size at least $\frac{2n}{5}$ asymptotically) monochromatic connected matching in any 3-uniform 3-colouring of $\Gamma(\mathcal{H})$ obtained from an almost complete hypergraph $\mathcal{H}$ with n vertices. Almost complete (or $(1-\varepsilon)$-dense) means that $\mathcal{H}$ 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 like Lemma 1.4, and will be proved in Section 2.

Lemma 1.4. *For all $\eta > 0$ there exist $\varepsilon > 0$ and n_0 such that for every $n > n_0$, the following is true. In every 3-uniform 3-colouring of $\Gamma(\mathcal{H})$ obtained from a $(1-\epsilon)$-dense 3-uniform hypergraph $\mathcal{H}$, there is a connected monochromatic matching of size at least $(\frac{2}{5} - \eta)n$.*

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 1.1 and 1.2. Although the approach outlined above is now becoming 'standard', there are several technical solutions to handling '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 $L \subset V(\mathcal{H})$ of k distinct vertices was obtained by a δ-*bounded selection* if its elements are chosen in k consecutive steps so that, at each step, there is a set of at most δn forbidden vertices that cannot be included as the next element. These sets of forbidden vertices may depend on the choices of the vertices at the previous steps. For simplicity, we sometimes call the sequence itself a δ-bounded selection. A basic property of almost complete hypergraphs is expressed in the following lemma.

Lemma 1.5. *Assume that $\mathcal{H}$ is a $(1-\epsilon)$-dense r-uniform hypergraph ($r \geqslant 2$) and set $\delta = \epsilon^{2^{-r}}$. There are forbidden sets such that, with respect to them, every δ-bounded selection $L \subset V(\mathcal{H})$ of length at most r is contained in at least $(1-\delta)\frac{n^{r-|L|}}{(r-|L|)!}$ edges of $\mathcal{H}$.*

The case $|L| = r$ in Lemma 1.5 is very important, because we get that every δ-bounded selection L is an edge of $\mathcal{H}$. The case $|L| = 0$ states that $\mathcal{H}$ has at least $\frac{(1-\delta)n^r}{r!}$ edges.

To illustrate how to use Lemma 1.5, we generalize a result in [10] (a more general form is given in [7]) from complete hypergraphs to almost complete ones. We start with a proposition (from [11]) about the connected components of a hypergraph.

Proposition 1.6. *Assume $\mathcal{H}$ is an arbitrary hypergraph and $0 < s < 1/3$. Then either there is a connected component $\mathcal{H}'$ of $\mathcal{H}$ with at least $(1-s)n$ vertices or the connected components of $\mathcal{H}$ can be partitioned into two groups so that each group contains more than sn vertices.* □

Proof. Mark the connected components of $\mathcal{H}$ until the union of them has at most sn vertices. If one unmarked component remains, it can be $\mathcal{H}'$. 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 1.7. *Assume that $\mathcal{H}$ is a $(1-\epsilon)$-dense r-uniform hypergraph with n vertices and $\delta = \epsilon^{2^{-r}} < \frac{1}{4}$. Then in every r-colouring of the edges of $\mathcal{H}$ there exists a monochromatic connected component covering all but at most δn vertices of $\mathcal{H}$.*

Proof. If the first possibility of Proposition 1.6 holds with $s = \delta$ in any of the hypergraphs determined by the edges of the different colour classes, we have nothing to prove. Otherwise the components of each colour class can be partitioned into X_i, Y_i so that both have more than δn vertices. We will reach a contradiction by applying Lemma 1.5 and defining a δ-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; in fact we will have $x_1 \in X_i$ for $i = 1, 2, \ldots, r$. This can be done as follows. We may assume that $|X_2| \leqslant |Y_2|$. Pick an arbitrary $y \in X_2$ (apart from the δn forbidden vertices); we may assume without loss of generality that $y \in X_i$ for all i, $1 \leqslant i \leqslant r$. Try a u 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 $z \in Y_2$ so that z is not in the exceptional set of y or u (there is a choice since $|Y_2| > 2\delta n$) and observe that either $L = (u, z)$ or $L = (y, z)$ satisfies the requirement for (x_1, x_2).

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 δ-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 $\mathcal{H}$, so it has a colour k, $1 \leqslant k \leqslant r$. 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-colourings

In this section we will prove Lemma 1.4. 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, [23, Theorem 3.1.14]). We will use $c_o(G)$ for the number of odd components of a graph G, and $\mathrm{def}(G)$, the deficiency of G, is defined as $|V(G)| - 2\nu(G)$.

Lemma 2.1. *For any graph G, $\mathrm{def}(G) = \max\{c_o(G \setminus T) - |T|\}$, where the maximum is taken over all $T \subseteq V(G)$.*

We also need the following obvious property of maximum matchings.

Lemma 2.2. *Suppose $M = \{e_1, \dots, e_k\}$ is a maximum matching in a graph G. Then $V(G) \setminus V(M)$ spans an independent set and one can select one endpoint x_i of each e_i – we call it the* strong point *– so that for each i, $1 \leqslant i \leqslant k$, there is at most one edge in G from x_i to $V(G) \setminus V(M)$.*

We assume that n is sufficiently large,

$$0 < \epsilon \ll \delta \ll \eta, \tag{2.1}$$

where δ is a technical parameter tending to zero if ϵ tends to zero; its role is to handle $(1-\epsilon)$-dense hypergraphs more easily (by Lemma 1.5). We may also assume that η is sufficiently small, since the statement (existence of a monochromatic connected matching of size $\left(\frac{2}{5} - \eta\right)n$) of Lemma 1.4 from any fixed η follows automatically for any larger η.

To prove Lemma 1.4, consider an arbitrary 3-uniform 3-colouring of $K = \Gamma(\mathcal{H})$, where $\mathcal{H}$ is a $(1-\epsilon)$-complete 3-uniform hypergraph with n vertices. Applying Lemma 1.5 with $\epsilon, r = 3$, we get that with $\delta = \epsilon^{1/8}$ every δ-bounded selection of at most three vertices is contained in the required number of edges ('many' edges for fewer than three vertices and one edge for exactly three vertices). It is more convenient to work with a 'truncated' hypergraph obtained from $\mathcal{H}$ as follows. Delete from $\mathcal{H}$ the (at most δn) vertices that are excluded as a first vertex of any δ-bounded selection (with respect to the forbidden sets ensured by Lemma 1.5) together with all edges incident to the deleted vertices. Moreover, also delete those (at most δn^3) edges $\{x_1, x_2, x_3\}$ of $\mathcal{H}$ for which x_2 is in the forbidden set of x_1. It is easy to see that in the truncated hypergraph, with a slightly larger δ, for example $\delta = 4\epsilon^{1/8}$, every δ-bounded selection of at most three vertices still contains the required number of edges (and, for example, the first vertex can be arbitrary). It is also clear that it is enough to find the required matching in the truncated hypergraph. Therefore, in the rest of the proof we assume that $\mathcal{H}$ is the truncated hypergraph (it is $(1 - O(\delta))$-dense so both δ and ϵ will change).

The edges of $\overline{K}$, the complement of K, will sometimes be referred to as the 'missing edges'. These are the edges uncovered by the hyperedges of $\mathcal{H}$. For convenience, we shall also consider the exceptional edges (from x_i to $V(G) \setminus V(M)$) of Lemma 2.2 as missing edges.

We call colour i good if there is a $V' \subseteq V$ such that $|V'| \geqslant (1-\delta)n$ and the edges of colour i in V' form only one non-trivial component C, where a trivial component is a single vertex. (We shall use the fact that in a good colour no edge in $V' \setminus V(C)$ has colour i.) There exists at least one good colour since, by Lemma 1.7, there is a colour with a connected component of at least $(1-\delta)n$ vertices in K. We select M_1 as a largest monochromatic matching among matchings in good colours, say M_1 is red in V'. We may assume that $|M_1| = k_1 = (\frac{2}{5} - \eta - \rho_1)n$ with some $0 < \rho_1 \leqslant (\frac{2}{5} - \eta)$, otherwise we are done. Furthermore, by a result of [14], we may also assume that $|M_1| = k_1 \geqslant (\frac{1}{4} - \eta)n$ (indeed this is true for *any* 3-colouring of an almost complete graph). Apply Lemma 2.2 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 $C = V' \setminus V(M_1)$. Thus we have

$$|A| = |B| = k_1 = \left(\frac{2}{5} - \eta - \rho_1\right)n \geqslant \left(\frac{1}{4} - \eta\right)n, \tag{2.2}$$

$$\left(\frac{1}{5} + 2\eta + 2\rho_1 - \delta\right)n \leqslant |C| \leqslant \left(\frac{1}{2} + 2\eta\right)n. \tag{2.3}$$

Call an edge of K *purely*-{blue, green} (or simply p-{blue, green}) if this edge cannot be red, and can therefore only be blue and/or green. (We have a multicolouring!) Similarly, a p-green edge can only be green. Note that – using the convention that the exceptional red edges from each vertex of B are considered to be missing edges – every edge of K inside C and in the bipartite graph $[B, C]$ is p-{blue, green}. We will frequently use the following fact.

Fact. Consider an edge $e \in \mathcal{H}$ such that the triangle defined by e in K contains a p-{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.

We have the following structural information about the 3-uniform 3-colouring on $A \cup B \cup C$ induced by K. Every edge of K inside C and in the bipartite graph $[B, C]$ is p-{blue, green}. Using this and the fact that, by δ-bounded selection, every $x_1, x_2 \in A \cup B \cup C$, such that not both x_1 and x_2 are in A, is contained in an edge $e \in \mathcal{H}$ having a p-{blue, green} edge, it follows that every edge of K within $A \cup B \cup C$ is blue and/or green except possibly the edges inside A. 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 the bipartite graph $[A, B \cup C]$. Now all edges of H have blue and/or green colours. Some of the edges of H might have a red colour as well, but we ignore that, *i.e.*, we consider H to be a 2-multicoloured graph. The pairs in $B \cup C$ and in $[A, B \cup C]$ that are not in K will be referred to as the 'missing edges' of H.

Proposition 2.3. *All vertices of H have missing degree at most δn.* □

Proof. By Lemma 1.5 every δ-bounded selection x_1, x_2 is covered by at least one (in fact by at least $(1 - \delta)n$) edge of $\mathcal{H}$. By assumption, every $x_1 \in V(H)$ can start such a selection and at most δn choices are forbidden for x_2. Thus, in the case of $x_1 \in A$ there are at most δn choices of $x_2 \in B \cup C$ such that no edge of $\mathcal{H}$ covers $\{x_1, x_2\}$. Similarly, for $x_1 \in B \cup C$ there are at most δn choices of $x_2 \in V(H)$ such that no edge of $\mathcal{H}$ covers $\{x_1, x_2\}$. □

Next we establish some facts about the monochromatic components of H. From (2.2) we have $|A|, |B \cup C| > 2\delta n$. It is easy to see that there exists $U_1 \subseteq B \cup C$ such that $|(B \cup C) \setminus U_1| \leqslant 2\delta n$ and $H[U_1]$ is connected in blue or green, say, without loss of generality, in blue. Indeed, if U is a green component of $H[B \cup C]$ such that $|V(H[B \cup C] \setminus U|, |U| > 2\delta n$, then Proposition 2.3 implies that the bipartite subgraph $[U, V(H[B \cup C] \setminus U] \subset H$ is connected in blue. (A worse bound could be obtained from Lemma 1.7.) In fact, we may assume that $B \cup C = U_1$, since deleting at most $2\delta n$ vertices does not influence the proof.

Let U_2 be the set of vertices in A with at least one blue neighbour 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 non-trivial blue component of H, so blue is a good colour.

Case I: $A \setminus U_2$ is non-empty. 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 non-trivial green component of H, so green is a good colour. In this case we define M_2 to be 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 non-trivial green component (*i.e.*, if green is a good colour) then we have the symmetry of Case I, and M_2 is defined to be the larger of two maximal matchings; without loss of generality, M_2 is blue again. Otherwise, if there is more than one non-trivial green component, M_2 is defined to be a maximum matching in K_1. (However, as will become clear later, we can find the required monochromatic matching without dealing with this possibility.)

Since M_2 is defined in a good colour, $|M_1| \geqslant |M_2|$ (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 $0 < \rho_1 \leqslant \rho_2 \leqslant (\frac{2}{5} - \eta)$, otherwise we are done. In the remainder of the proof of Lemma 1.4, we will show in all cases that either we can find a green connected matching of size at least $\frac{2}{5}n$, or there is only one non-trivial 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 2.2 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 $18\delta n$ vertices – we may assume that every non-empty atom has order larger than $2\delta n$. We have

$$|S| = |W| = k_2 = \left(\frac{2}{5} - \eta - \rho_2\right)n, \tag{2.4}$$

$$\left(\frac{1}{5} + 2\eta + 2\rho_2 - 10\delta\right)n \leqslant |R| \leqslant \left(\frac{1}{5} + 2\eta + 2\rho_2\right)n. \tag{2.5}$$

Note that, considering at most one blue edge from each vertex of S as a missing edge, every edge of H in R and in $[S, R]$ is p-green.

Case 1: $R_C \neq \emptyset$ (then $|R_C| > 2\delta n$). Consider a δ-bounded selection starting with $v \in S$. Here we have the following claim.

Claim 1. *All but at most δn 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 δ-bounded selection (this is possible since $|R_C| > \delta n$), and let (v, u) be p-green in H . From Lemma 1.5, for all but at most δn choices of $w \in B \cup C$, the triple $\{u, v, w\}$ is an edge of $\mathcal{H}$. The colour of this edge cannot be red because of the edge (u, w) that is p-{blue, green} (since $u \in C$), and 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, w) is indeed green, proving the claim (since the edge (v, u) is also green).

Since every vertex $v \in S$ can start a δ-bounded selection, the green colour 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_C, a contradiction. Thus, from now on we may assume that $R_C = \emptyset$.

Case 2: $R_C = \emptyset$ and $R_B \neq \emptyset$ (then $|R_B| > 2\delta n$). We define an auxiliary green subgraph H_1 of H. The vertex set of H_1 is $S_A \cup S_B \cup R_A \cup R_B \cup C$ and its edge set is the set of green edges in the union of the following subgraphs:

$$\begin{aligned} &H[S_A, R_B],\ H[S_B, R_A],\ H[S_B, R_B],\ H[R_A, R_B],\ H[R_B], \\ &\quad H[S_A, C],\ H[R_A, C],\ H[S_B, C],\ H[R_B, C]. \end{aligned} \tag{2.6}$$

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 any vertex of H_1, all but at most δn 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 (2.6), since edges of R and of $[S, R]$ are p-green in H. For the subgraphs involving C we proceed as in Claim 1. For $H[S_A, C]$ and $H[R_B, C]$, start a δ-bounded selection with $v \in S_A$. Continue with $u \in R_B$ 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). Let w be the third vertex of the selection from C. Consider the colour of the edge $\{u, v, w\}$ in $\mathcal{H}$. It cannot be red because of the edge (u, w) that is p-{blue, green} (since $w \in C$), and 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, 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_1 has a matching M_3, leaving out at most constant times δn 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 $k_2 - |C| + 10\delta n \leqslant \frac{1}{5}n$ vertices using (2.1), (2.3) and (2.4), and thus

$$|M_3| \geqslant \frac{2}{5}n.$$

From Claim 2 and from the assumption on the sizes of the atoms, H_1 is connected, and thus the matching M_3 is connected.

To show that $\nu(H_1)$ is large, we bound $\mathrm{def}(H_1) = \max\{c_o(H_1 \setminus T) - |T|\}$ in the Tutte–Berge formula with Lemma 2.4. To prepare this, we assign a base graph $G = G(H_1)$ to H_1 whose vertices are the atoms of H_1; with the present H_1, G has five vertices. To each vertex $v \in V(G)$ the number of vertices of the atom corresponding to v is assigned as the weight, $w(v)$. Two (not necessarily distinct) vertices of G are adjacent if there are edges of H_1 between the corresponding atoms; see (2.6). Using the same notation for the vertices of G as for the atoms of $V(H_1)$, our base graph has one loop at vertex R_B, and all but two pairs of distinct vertices are adjacent, the exceptions being $(S_A, S_B), (S_A, R_A)$. The weight of $Z \subseteq V(G)$, $w(Z)$, is the sum of its vertex weights. Define $cr^*(Z) = w(Z) - w(N_G(Z))$, where $N_G(Z)$ is the open neighbourhood of S in G, *i.e.*, the set of vertices in $V(G) \setminus Z$ adjacent to at least one vertex of Z. Set $c^*(G) = \max_S cr^*(S)$, where the maximum is taken over all independent sets S of the base graph G (a vertex with a loop cannot be in any independent set). We state the following lemma for a general base graph with t vertices, although $t \leqslant 9$ in all applications.

Lemma 2.4. *Let G be the base graph of H_1, $|V(G)| = t$. Then* $\mathrm{def}(H_1) \leqslant c^*(G) + 4t\delta n + t$

Proof. Let F_i denote the atoms of H_1 and let i denote the corresponding vertex in the base. Suppose (see Lemma 2.1) that $\mathrm{def}(H_1) = c_0(H_1 \setminus T) - |T|$ for some $T \subseteq V(H_1)$. Set $X_i = F_i \cap T, Y_i = F_i \setminus X_i$. If $|Y_i| \leqslant 2\delta n$ then Y_i is called small, otherwise it is large. Let W denote the set of vertices in G (set of indices) for which Y_i is small.

Clearly at most $2t\delta n$ odd components of $H_1 \setminus T$ intersect the union of the small Y_is. The other odd components are in the union of the large Y_is; however, only a few, at most t, can be non-trivial. Indeed, suppose a component C has an edge $e \in E(H_i)$ between two large Y_i, Y_j for $i \neq j$. Since the missing degree of any vertex of H_1 is at most δn, C spans $Y_i \cup Y_j$. Similarly, if an edge $e \in E(H_1)$ of a component is within a large Y_i, then i has a loop and the component must span Y_i in this case. Therefore at most t non-trivial components (in particular odd non-trivial components) are in the union of the large Y_is (this can occur if all edges of G are loops and each weight is odd). The set B of other odd components has to be trivial, *i.e.*, one-vertex components in the union of some large Y_is, $B \subseteq \cup_{i \in S} Y_i$. Note that S is an independent set of G (and cannot contain loops either) and $|B| \leqslant \sum_{i \in S} |Y_i| \leqslant w(S)$. Thus

$$c_o(H_1 \setminus T) \leqslant 2t\delta n + t + \sum_{i \in S} |Y_i| \leqslant 2t\delta n + t + w(S). \tag{2.7}$$

Note that $N_G(S) \subseteq W$, for otherwise there is an edge $uv \in E(H_1)$ such that $u \in B, v \in Y_j$ with a large Y_j, which contradicts the fact that u is an isolated vertex in $V(H_1) \setminus T$. Using this and the definition of W, we have

$$\sum_{i \in W} |X_i| > \sum_{i \in W} w(i) - 2t\delta n = w(W) - 2t\delta n \geqslant w(N_G(S)) - 2t\delta n. \tag{2.8}$$

Combining (2.7) and (2.8),

$$\begin{aligned}\mathrm{def}(H_1) = c_o(H_1 \setminus T) - |T| &\leqslant 2t\delta n + t + w(S) - \sum_{i=1}^{t} |X_i| \\ &\leqslant 2t\delta n + t + w(S) - \sum_{i\in W} |X_i| < 4t\delta n + t + w(S) - w(N_G(S)) \\ &= cr^*(S) + 4t\delta n + t \leqslant c^*(G) + 4t\delta n + t, \qquad (2.9)\end{aligned}$$

giving the claimed estimate of the lemma. □

Using Lemma 2.4, in the next claim we estimate $cr^*(S)$ for the independent sets S of $G(H_1)$. As before, with a slight abuse of notation, vertices of $G(H_1)$ are denoted the same way as the corresponding atoms in $V(H_1)$. Lemma 2.4 will be used in the proof of other subcases, at which points we will omit some of the details.

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

$$\begin{aligned}cr^*(S_1) &= (|S_A| + |R_A|) - (|S_B| + |R_B| + |C|) \leqslant -\eta n,\\ cr^*(S_2) &= (|S_A| + |S_B|) - (|R_A| + |R_B| + |C|) \leqslant -\eta n,\\ cr^*(S_3) &= |C| - (|R_A| + |R_B| + |S_A| + |S_B|) \leqslant -\eta n.\\ cr^*(S_4) &= |S_A| - (|R_B| + |C|) \leqslant -\eta n.\end{aligned}$$

Indeed, we get these estimates from the inequalities

$$|S_A| + |R_A| \leqslant |A| = |B| = |S_B| + |R_B| + |W_B| < |S_B| + |R_B| + |C| - \eta n,$$

where we used

$$|W_B| < |C| - \eta n. \qquad (2.10)$$

To get this inequality, note that here all vertices in C are weak endpoints, so $W_C = C$, but then using (2.3) we have $|W_C| = |C| > \frac{n}{5} + \eta n$. However, from (2.4) we get $|W_B \cup W_C| \leqslant |W| < \frac{2}{5}n$, which implies $|W_B| < \frac{n}{5}$, and thus $|W_B| < \frac{n}{5} < |C| - \eta n$, as desired. Then

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

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

and

$$|S_A| + |R_A| \leqslant |A| \leqslant \left(\frac{2}{5} - \eta\right)n < |R| + |C| - \eta n|R_A| + |R_B| + |C| - \eta n.$$

Furthermore, note that for $S_5 = \{R_A\}$ and $S_6 = \{S_B\}$ we have $cr^*(S_5) \leqslant cr^*(S_1) \leqslant -\eta n$ and $cr^*(S_6) \leqslant cr^*(S_2) \leqslant -\eta n$, since $N_G(S_5) = N_G(S_1)$ and $N_G(S_6) = N_G(S_2)$. Thus, for each independent set S of $G(H_1)$ we have $cr^*(S) \leqslant -\eta n$, and thus $c^*(G(H_1)) \leqslant -\eta n$. Thus, from Lemma 2.4 with $t = 5$ (and using that $\delta \ll \eta$), we get the following.

Claim 4. $\operatorname{def}(H_1) \leqslant 20\delta n + 5$.

This shows that H_1 has an almost perfect matching, thus finishing the subcase.

Subcase 2.2: $|S_C| > 0$ (then $|S_C| > 2\delta n$). 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 now contains 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 δ-bounded selection starting with $v \in S_C$. Continue with $u \in R_A$ 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, and all other edges are good). Finish the selection with $w \in W_B$. Consider the colour 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 $v \in C$), and 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 will show again that H_1 has a matching of size at least $\frac{2}{5}n$. For this purpose we apply Lemma 2.4 again, as above, so we will not present all the details. As above, we have to show that for each maximal independent set S of $G(H_1)$ we have $cr^*(S) \leqslant (\frac{1}{5} - \eta)n$. 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

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

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

from the following inequalities (using (2.2), (2.3) and (2.5)):

$$\begin{aligned} |S_A| + |W_A| + |R_A| + |S_B| + |W_B| + |R_B| &= |A| + |B| \leqslant \left(\frac{4}{5} - \eta\right)n \\ &\leqslant 2|R| + |C| + \left(\frac{1}{5} - \eta\right)n = 2(|R_A| + |R_B|) + |C| + \left(\frac{1}{5} - \eta\right)n, \end{aligned}$$

and

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

Similarly, the other independent sets also satisfy $cr^*(S) \leqslant (\frac{1}{5} - \eta)n$. Then, from Lemma 2.4 (and using that $\delta \ll \eta$), we again get a matching M_3 in H_1 of size almost $\frac{2}{5}n$. 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 2.2 every edge of M_2 has at least one (now perhaps 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 $|W| \leqslant |S|$. 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: $R_B = R_C = \emptyset$ and $|S_C| > 0$ (then $|S_C| > 2\delta n$). Here we define H_1 to be the green edges in the union of the subgraphs

$$H[R_A, W_B],\ H[R_A, S_B],\ H[R_A, S_C],\ H[R_A, W_C],$$
$$H[W_B, S_C],\ H[S_B, S_C],\ H[S_B, W_C],\ H[S_C, W_C].$$

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 by Lemma 2.4 that H_1 contains an almost perfect matching M_3. As above, we have to show that for each independent set S of $G(H_1)$ we have $cr^*(S) \leqslant -\eta n$. 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

$$cr^*(S_1) = (|W_B| + |W_C|) - (|S_B| + |S_C| + |R_A|) \leqslant -\eta n,$$
$$cr^*(S_2) = |S_C| - (|B| + |R_A| + |W_C|) \leqslant -\eta n$$

from the inequalities

$$|W_B| + |W_C| \leqslant |W| \leqslant |S| = |S_A| + |S_B| + |S_C| \leqslant |S_B| + |S_C| + |R_A| - \eta n$$

and

$$|S_C| \leqslant |C| \leqslant |R| = |R_A| \leqslant |R_A| + |B| + |W_C| - \eta n$$

(using $|C| \leqslant |R|$, *i.e.*, $|M_1| \geqslant |M_2|$). Similarly, the other independent sets can be checked, and we find that $cr^*(S) \leqslant -\eta n$ holds for all of them. Thus, again we have an almost perfect matching M_3 in H_1 from Lemma 2.4. This matching leaves out only at most

$$|W_A| + |S_A| + 10\delta n \leqslant \frac{1}{5} n$$

vertices (using (2.1), (2.2) and (2.5)), as we wanted.

Case 4: Finally we may assume that $R_B = R_C = \emptyset$ and $S_C = \emptyset$. Hence $C = W_C$. Here we define H_1 to be the green edges in the union of the subgraphs

$$H[R_A, S_B],\ H[R_A, C],\ H[S_B, C].$$

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 $\frac{2}{5}n$). 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 of the size of S':

$$|S'| \geqslant |W_B| - (\rho_2 + \delta)n. \tag{2.11}$$

Indeed, as the strong endpoints corresponding to vertices in

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

should all go to S_B (using the fact that there are no edges of H in A), we have

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

From this, (2.3) and (2.5), we get the estimate

$$\begin{aligned}2|S'| &\geqslant |W_A| + |S_A| + 2|W_B| + |C| - |B| = |A| - |R_A| + 2|W_B| + |C| - |B| \\ &= |A| - |R| + 2|W_B| + |C| - |B| \geqslant 2|W_B| + 2(\rho_1 - \rho_2 - \delta)n \geqslant 2(|W_B| - (\rho_2 + \delta)n),\end{aligned}$$

and thus we get (2.11).

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 δ-bounded selection with $u \in S', v \in S_B^1$ and $w \in C$. Consider the colour of the triple $\{u, v, w\}$ in the original $\mathcal{H}$. This cannot be red because of the edge (v, w), which is p-{blue, green} (since $w \in C$), and it cannot be blue because of the edge (v, u), which 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 too we can prove with Lemma 2.4 that H_1 contains a matching M_3 covering all but at most $|S'| - |W_B| + (\rho_2 + 2\delta)n$ ($\geqslant \delta n$ using (2.11)) vertices of H_1. This will be enough, as by (2.11) this matching leaves out only at most

$$\begin{aligned}&|W_B| + |W_A| + |S_A| - |S'| + |S'| - |W_B| + (\rho_2 + 2\delta)n + 10\delta n \\ &\qquad \leqslant |W_A| + |S_A| + \rho_2 n + 12\delta n \leqslant \frac{1}{5}n\end{aligned}$$

vertices, as we wanted. Here the last inequality follows from (2.1), (2.2), (2.5), and

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

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

$$cr^*(S) \leqslant |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n. \tag{2.12}$$

We have to check this inequality for all the subsets of the maximal independent sets $\{S', R_A\}$, $\{S', S_B^2\}$, $\{S_B^1, S_B^2\}$ and $\{C\}$. For example, for the independent sets $S_1 = \{S', R_A\}$

and $S_2 = \{S', S_B^2\}$, we get

$$cr^*(S_1) = (|S'| + |R_A|) - (|S_B| + |C|) \leqslant |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n,$$
$$cr^*(S_2) = (|S'| + |S_B^2|) - (|S_B^1| + |R_A| + |C|) \leqslant |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n$$

from the following inequalities (again using (2.10)):

$$\begin{aligned}|S'| + |R_A| \leqslant |A| = |B| = |S_B| + |W_B| &\leqslant |S_B| + |C| - \eta n \\ &\leqslant |S_B| + |C| + |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n\end{aligned}$$

and

$$\begin{aligned}|S'| + |S_B^2| = |S'| - |W_B| + |W_B| + |S_B^2| &\leqslant |S'| - |W_B| + (\rho_2 + 2\delta)n + |B| \\ &\leqslant |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n + \frac{2}{5}n \\ &\leqslant |S'| - |W_B| + (\rho_2 + 2\delta - \eta)n + |R_A| + |C| + |S_B^1|.\end{aligned}$$

Similarly, the other independent sets can be checked, and we find that, for each one, (2.12) is satisfied, as we wanted. This completes the proof of Lemma 1.4. □

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 [29] for hypergraphs. There are several generalizations of the Regularity Lemma for hypergraphs due to various authors (see [3], [6], [9], [28] and [30]). Here we will use the simplest one due to Chung [3]. First we need to define the notion of ε-regularity. Let $\varepsilon > 0$ and let V_1, V_2, V_3 be disjoint vertex sets of size m, and let $\mathcal{H}$ be a 3-uniform hypergraph such that every edge of $\mathcal{H}$ contains exactly one vertex from each V_i for $i = 1, 2, 3$. The density of $\mathcal{H}$ is $d_{\mathcal{H}} = \frac{|E(\mathcal{H})|}{m^3}$. The triple $\{V_1, V_2, V_3\}$ is called an $(\varepsilon, \mathcal{H})$-regular triple of density $d_{\mathcal{H}}$ if, for every choice of $X_i \subset V_i$, $|X_i| > \varepsilon|V_i|$, $i = 1, 2, 3$, we have

$$\left| \frac{\mathcal{H}[X_1, X_2, X_3]}{|X_1||X_2||X_3|} - d_{\mathcal{H}} \right| < \varepsilon.$$

Here, by $\mathcal{H}[X_1, X_2, X_3]$ we denote the subhypergraph of $\mathcal{H}$ induced by the vertex set $X_1 \cup X_2 \cup X_3$. In this setting the 3-colour version of the (weak) Hypergraph Regularity Lemma from [3] can be stated as follows.

Lemma 3.1 (3-colour Weak Hypergraph Regularity Lemma). *For every positive ε and positive integer M_0 there are positive integers M_1 and n_0 such that for $n \geqslant n_0$ the following holds. For all 3-uniform hypergraphs $\mathcal{H}_1$, $\mathcal{H}_2, \mathcal{H}_3$ with $V(\mathcal{H}_1) = V(\mathcal{H}_2) = V(\mathcal{H}_3) = V$, $|V| = n$, there is a partition of V into $l+1$ classes (clusters),*

$$V = V_0 \cup V_1 \cup V_2 \cup \cdots \cup V_l,$$

such that

- $M_0 \leqslant l \leqslant M_1$,

- $|V_1| = |V_2| = \cdots = |V_l|$,
- $|V_0| < \varepsilon n$,
- *apart from at most* $\varepsilon\binom{l}{3}$ *exceptional triples, the triples* $\{V_{i_1}, V_{i_2}, V_{i_3}\}$ *are* $(\varepsilon, \mathcal{H}_s)$*-regular for* $s = 1, 2, 3$.

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

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

We apply the above 3-colour Weak Hypergraph Regularity Lemma with a small enough ε, and we get a partition of $V(K_n^{(3)}) = V = \cup_{0 \leqslant i \leqslant l} V_i$, where $|V_i| = m, 1 \leqslant i \leqslant l$. 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 $p_{i_1}, p_{i_2}, p_{i_3}$ if the triple $\{V_{i_1}, V_{i_2}, V_{i_3}\}$ is $(\varepsilon, \mathcal{H}_s)$-regular for $s = 1, 2, 3$. 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,

$$|E(\mathcal{H}^R)| \geqslant (1 - \varepsilon)\binom{l}{3},$$

and thus $\mathcal{H}^R$ is a $(1 - \varepsilon)$-dense 3-uniform hypergraph on l vertices. Define a 3-edge colouring $(\mathcal{H}_1^R, \mathcal{H}_2^R, \mathcal{H}_3^R)$ of $\mathcal{H}^R$ with the majority colour, *i.e.*, the triple $\{V_{i_1}, V_{i_2}, V_{i_3}\} \in \mathcal{H}_s^R$ if colour s is the most frequent colour in this triple. Note then that the density of this colour is $\geqslant 1/3$ in this triple. Finally we consider the multicoloured 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 colour s, $s = 1, 2, 3$ if x and y are contained in an edge of $\mathcal{H}^R$ that is coloured with colour s.

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

Lemma 3.2. *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 coloured* $K_n^{(3)}$, *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 proof (first suggested by [24]) has become somewhat standard (see [5], [11], [14], [15], [12], [13] and [24]), so a proof of this lemma can be found in [11], for example. For the sake of completeness we repeat the proof.

Proof. We may assume that M is in $\Gamma(\mathcal{H}_1^R)$. Denote the edges on M by $M = \{e_1, e_2, \ldots, e_{l_1}\}$ 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 endpoints of e_i.

Next we define *good* vertices for an arbitrary edge e in $\Gamma(\mathcal{H}_1^R)$. Denote $f(e) = (V^1, V^2)$. Since e is an edge in $\Gamma(\mathcal{H}_1^R)$, the endpoints of e are contained in a triple E in $\mathcal{H}_1^R$. 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 $\geqslant 1/3$. 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 $m/6$ vertices $y \in V^{j'}$, there are at least $m/6$

triples in $\mathcal{H}_1[f(E)]$ containing both x and y (recall that m is the number of vertices in the clusters). 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

$$|X|\left(\frac{m}{6}m + \left(1 - \frac{1}{6}\right)m\frac{m}{6}\right) = \left(\frac{1}{3} - \frac{1}{36}\right)|X|m^2, \tag{3.1}$$

which contradicts the fact that $f(E)$ is an $(\varepsilon, \mathcal{H}_1)$-regular triple with density at least $1/3$ if ε 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

$$\geqslant (1/6 - \varepsilon)m > m/7 \tag{3.2}$$

vertices $y \in V^{j'}$ such that there are at least $m/6$ 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 (ε, δ, 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,

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

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 (3.2). 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 $E_G(X, Y) \leqslant |X||Y|/7$. 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

$$|X||Y|\left(\frac{m}{7} + \left(1 - \frac{1}{7}\right)\frac{m}{6}\right) = \left(\frac{1}{3} - \frac{1}{21}\right)|X||Y|m, \tag{3.3}$$

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 $f^{-1}(V_1^{i+1})$ for every $1 \leqslant i \leqslant l_1$ (for $i = l_1$ put $V_j^{i+1} = V_j^1$). 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. Denote $P_1^R = (p_1, \ldots, p_t), 2 \leqslant t \leqslant l$, 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_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 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 $m/6$ triples 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' 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 in 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 ensure that we now have the same number of vertices left in each cluster of the matching. For simplicity we keep the notation V_j^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 $1 \leqslant i \leqslant l_1$, and we now have $|V_1^i| = |V_2^i|$.

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

Lemma 3.3. *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 3.3 is used by Łuczak in [24]. Lemma 3.3 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 3.3 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 a Berge cycle. For the edges on the connecting paths P_i we have already assigned distinct $\mathcal{H}_1$ triples. 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 $m/7$ triples (we have already removed some vertices and triples) in $\mathcal{H}_1[f(E_i)]$ (here E_i denotes the triple in $\mathcal{H}_1^R$ containing the endpoints of the edge e_i). Note that the triples E_i must be distinct for each $i, 1 \leqslant i \leqslant l_1$, because the pairs (V_1^i, V_2^i), $1 \leqslant i \leqslant l_1$, are all disjoint. Furthermore, the triples containing two distinct edges from $G(V_1^i, V_2^i)$ 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 3.2. □

Putting together Lemma 3.2 with the asymptotic result of the previous section on monochromatic connected matchings (Lemma 1.4) we get the desired asymptotic result on monochromatic Berge cycles (Theorem 1.1).

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 1.2. Indeed, Lemma 3.3 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 $4 \leqslant m' \leqslant 2m$. If the parity is not right we can change the parity with the following simple trick. Consider $P_1^R = (p_1, \ldots, p_t)$. Since the edge (p_1, p_2) is in $\Gamma(\mathcal{H}_1^R)$, there is a triple E in $\mathcal{H}_1^R$ 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_1^R. In this way we have increased the length by one and thus changed the parity. Hence, in Theorem 1.2 we can find a monochromatic Berge cycle of length exactly n.

Acknowledgement

Thanks to an unknown referee whose useful remarks improved the presentation.

References

[1] Berge, C. (1973) *Graphs and Hypergraphs*, North-Holland.
[2] Bermond, J. C., Germa, A., Heydemann, M. C. and Sotteau, D. (1978) Hypergraphes hamiltoniens. In *Problèmes Combinatoires et Théorie des Graphes* (Colloq. Internat. CNRS, Orsay 1976), Vol. 260, pp. 39–43.
[3] Chung, F. (1991) Regularity lemmas for hypergraphs and quasi-randomness. *Random Struct. Alg.* **2** 241–252.
[4] Erdős, P. and Gallai, T. (1959) On maximal paths and circuits of graphs. *Acta Math. Acad. Sci. Hungar.* **10** 337–356.
[5] Figaj, A. and Łuczak, T. (2007) The Ramsey number for a triple of long even cycles. *J. Combin. Theory Ser. B* **97** 584–596.

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

[7] Füredi, Z. and Gyárfás, A. (1991) Covering t-element sets by partitions. *Europ. J. Combin.* **12** 483–489.

[8] Gerencsér, L. and Gyárfás, A. (1967) On Ramsey-type problems. *Ann. Univ. Sci. Budapest Eötvös, Sect. Math.* **10** 167–170.

[9] Gowers, W. T. (2007) Hypergraph regularity and the multidimensional Szemerédi Theorem. *Ann. of Math.* (2) **166** 897–946.

[10] Gyárfás, A. (1973) Partition coverings and blocking sets of hypergraphs (in Hungarian). Studies of Computer and Automation Research Institute, No. 71, MR0357172.

[11] Gyárfás, A., Lehel, J., Sárközy, G. N. and Schelp, R. H. (2008) Monochromatic Hamiltonian Berge cycles in colored complete hypergraphs. *J. Combin. Theory Ser. B* **98** 342–358.

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

[13] Gyárfás, A., Ruszinkó, M., Sárközy, G. N. and Szemerédi, E. (2006) One-sided coverings of colored complete bipartite graphs. In *Topics in Discrete Mathematics* (dedicated to J. Nešetřil on his 60th birthday), Vol. 26 of *Algorithms and Combinatorics* (M. Klazar *et al.*, eds), Springer, pp. 133–154.

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

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

[16] Haxell, P., Łuczak, T., Peng, Y., Rödl, V., Ruciński, A., Simonovits, M. and Skokan, J. (2006) The Ramsey number for hypergraph cycles I. *J. Combin. Theory Ser. A* **113** 67–83.

[17] Haxell, P., Łuczak, T., Peng, Y., Rödl, V., Ruciński, A. and Skokan, J. (2009) The Ramsey number for 3-uniform tight hypergraph cycles. *Combin. Probab. Comput.* **18** 165–203.

[18] Katona, G. Y. and Kierstead, H. A. (1999) Hamiltonian chains in hypergraphs. *J. Graph Theory* **30** 205–212.

[19] Kohayakawa, Y., Simonovits, M. and Skokan, J. The 3-color Ramsey number of odd cycles. Manuscript. *J. Combin. Theory Ser. B*, to appear.

[20] Komlós, J., Sárközy, G. N. and Szemerédi, E. (1997) Blow-Up Lemma. *Combinatorica* **17** 109–123.

[21] Komlós, J., Sárközy, G. N. and Szemerédi, E. (1998) An algorithmic version of the Blow-Up Lemma. *Random Struct. Alg.* **12** 297–312.

[22] Komlós, J. and Simonovits, M. (1996) Szemerédi's Regularity Lemma and its applications in graph theory. In *Combinatorics: Paul Erdős is Eighty* (D. Miklós, V. T. Sós, and T. Szőnyi, eds), Vol. 2 of *Bolyai Society Math. Studies*, pp. 295–352.

[23] Lovász, L. and Plummer, M. D. (1986) *Matching Theory*, North-Holland and Akadémiai Kiadó.

[24] Łuczak, T. (1999) $R(C_n, C_n, C_n) \leqslant (4 + o(1))n$. *J. Combin. Theory Ser. B* **75** 174–187.

[25] Rosta, V. (1973) On a Ramsey-type problem of J. A. Bondy and P. Erdős I. *J. Combin. Theory Ser. B* **15** 94–104.

[26] Rosta, V. (1973) On a Ramsey-type problem of J. A. Bondy and P. Erdős II. *J. Combin. Theory Ser. B* **15** 105–120.

[27] Rödl, V., Ruciński, A. and Szemerédi, E. (2006) A Dirac-type theorem for 3-uniform hypergraphs. *Combin. Probab. Comput.* **15** 229–251.

[28] Rödl, V. and Skokan, J. (2004) Regularity Lemma for uniform hypergraphs. *Random Struct. Alg.* **25** 1–42.

[29] Szemerédi, E. (1978) Regular partitions of graphs. In In *Problèmes Combinatoires et Théorie des Graphes* (Colloq. Internat. CNRS, Orsay 1976), Vol. 260, pp. 399–401.

[30] Tao, T. (2006) A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A* **113** 1257–1280.