Available online at www.sciencedirect.com

SCIENCE DIRECT®

Journal of Combinatorial Theory, Series A 113 (2006) 67–83

Journal of Combinatorial Theory Series A

www.elsevier.com/locate/jcta

The Ramsey number for hypergraph cycles I

P.E. Haxell[a], T. Łuczak[b], Y. Peng[c], V. Rödl[d], A. Ruciński[b], M. Simonovits[e], J. Skokan[f,g]

[a] *Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

[b] *Department of Discrete Mathematics, Adam Mickiewicz University, 61-614 Poznań, Poland*

[c] *Department of Mathematics and Computer Science, Indiana State University, Terre Haute, IN 47809, USA*

[d] *Department of Mathematics and Computer Science, Emory University, Atlanta, GA 30032, USA*

[e] *Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, H-1053 Budapest, Reáltanoda u. 13-15.*

[f] *Department of Mathematics, University of Illinois at Urbana-Champaign, 1409 W. Green Street, Urbana, IL 61801, USA*

[g] *Instituto de Matemática e Estatística, Universidade de São Paulo, Rua do Matão 1010, 05508–090 São Paulo, Brazil*

Received 19 March 2004

Available online 29 March 2005

Abstract

Let $\mathcal{C}_n$ denote the 3-uniform hypergraph *loose cycle*, that is the hypergraph with vertices $v_1, \ldots, v_n$ and edges $v_1v_2v_3$, $v_3v_4v_5$, $v_5v_6v_7, \ldots, v_{n-1}v_nv_1$. We prove that every red-blue colouring of the edges of the complete 3-uniform hypergraph with N vertices contains a monochromatic copy of $\mathcal{C}_n$, where N is asymptotically equal to $5n/4$. Moreover this result is (asymptotically) best possible.

Keywords: Ramsey number; Hypergraph; Colouring; Regularity lemma; Cycle

1. Introduction

A classical result in graph theory states that for $n \geqslant 5$, the 2-colour Ramsey number $r(C_n, C_n)$ of the cycle C_n with n vertices is $2n - 1$ if n is odd and $3n/2 - 1$ if n is even [1,3,10]. Bondy and Erdős conjectured in 1973 that the 3-colour Ramsey number

E-mail addresses: pehaxell@math.uwaterloo.ca (P.E. Haxell), tomasz@amu.edu.pl (T. Łuczak), mapeng@isugw.indstate.edu (Y. Peng), rodl@mathcs.emory.edu (V. Rödl), andrzej@mathcs.emory.edu (A. Ruciński), miki@renyi.hu (M. Simonovits), jozef@member.ams.org (J. Skokan).

doi:10.1016/j.jcta.2005.02.005

$r(C_n, C_n, C_n) \leqslant 4n - 3$, and this appears to be quite a difficult problem. An asymptotic version was proved by Łuczak in 1999 [8], who showed that $r(C_n, C_n, C_n) = (4 + o(1))n$ if n is odd. Recently, Figaj and Łuczak [4] found the asymptotic value of the Ramsey number for all triples of even long cycles, proving, in particular, that $r(C_n, C_n, C_n) = (2 + o(1))n$ if n is even. Thus, in this case as well, the Ramsey number depends in a significant way on the parity of n.

In this paper we consider the problem of finding the 2-colour Ramsey number for a 3-uniform hypergraph cycle with n vertices. There are several natural definitions for a cycle in a 3-uniform hypergraph. The one we focus on here is what we call the *loose* cycle $\mathcal{C}_n$, which has vertices $v_1, \dots, v_n$ and edge set $\{v_1v_2v_3, v_3v_4v_5, v_5v_6v_7, \dots, v_{n-1}v_nv_1\}$. Note that the number of vertices n of $\mathcal{C}_n$ is necessarily even, and $\mathcal{C}_n$ contains exactly $n/2$ vertices of degree one and $n/2$ vertices of degree two. The number of edges of $\mathcal{C}_n$, called the *length* of $\mathcal{C}_n$, is $n/2$. We remark that in another paper [7] we consider the analogous problem for the *tight* cycle, which has vertex set $v_1, \dots, v_n$ and edge set $\{v_1v_2v_3, v_2v_3v_4, v_3v_4v_5, \dots, v_nv_1v_2\}$. The Ramsey number for the tight cycle is larger than that of the loose cycle, and the techniques required in the argument are substantially different.

The Ramsey number $r(\mathcal{C}_n, \mathcal{C}_n)$ of $\mathcal{C}_n$ is defined to be the smallest integer N for which every colouring of the edges of the complete 3-uniform hypergraph $K_N^{(3)}$ contains a monochromatic copy of $\mathcal{C}_n$, that is, a copy of $\mathcal{C}_n$ whose edges are all coloured the same colour. We first note the following lower bound on $r(\mathcal{C}_n, \mathcal{C}_n)$.

Lemma 1.1. *We have $r(\mathcal{C}_{4k}, \mathcal{C}_{4k}) > 5k - 2$ and $r(\mathcal{C}_{4k+2}, \mathcal{C}_{4k+2}) > 5k + 1$.*

Proof. To prove the first assertion we exhibit a 2-colouring of the edges of $K_{5k-2}^{(3)}$ that contains no monochromatic copy of $\mathcal{C}_{4k}$. We partition the vertex set into parts A and B, where $|A| = k - 1$ and $|B| = 4k - 1$. We colour all edges that contain a vertex of A blue, and the rest red. Now this colouring cannot contain a red copy of $\mathcal{C}_{4k}$, since any such copy must have all vertices in B. Suppose that a blue copy of $\mathcal{C}_{4k}$ exists, then each of its edges must contain a vertex of A, that is, a vertex-cover of $\mathcal{C}_{4k}$ would be contained in A. But $\mathcal{C}_{4k}$ has $2k$ edges and maximum degree two, and so cannot have a vertex cover of size smaller than k. Therefore, since $|A| = k - 1$, no blue copy of $\mathcal{C}_{4k}$ can exist in this colouring, hence the lower bound is proved.

The argument for the second bound is exactly the same, except that we take $|A| = k$ and $|B| = 4k + 1$. □

Our main aim in this paper is to prove that this lower bound is asymptotically tight, in other words that $r(\mathcal{C}_n, \mathcal{C}_n) \sim 5n/4$.

Theorem 1.2. *For all $\eta > 0$ there exists $n_0 = n_0(\eta)$ such that for every $n > n_0$, every 2-colouring of $K_{5(1+\eta)n/4}^{(3)}$ contains a monochromatic copy of $\mathcal{C}_n$.*

As in [4,8], our proof is based on the Regularity Lemma, here however we will need a hypergraph variant of this result. Note that in contrast to the graph case, the parity of the length of our cycle does not have a significant effect on its Ramsey number.

In our arguments we will often need to find loose *paths* between specified vertices. The *loose path* $\mathcal{P}_n$ has vertices $v_1, \dots, v_n$ and edge set $\{v_1v_2v_3, v_3v_4v_5, v_5v_6v_7, \dots, v_{n-2}v_{n-1}v_n\}$. By the *length* of a loose path we mean its number of edges, so $\mathcal{P}_n$ has length $(n-1)/2$. Often we will abbreviate the terms "3-uniform hypergraph", "loose cycle", and "loose path" to "hypergraph", "cycle", and "path" whenever there is no danger of confusion.

2. Proof of Theorem 1.2

In this section we state the principal lemmas required to prove Theorem 1.2, and show how they lead to the proof.

We begin by introducing a regularity lemma for hypergraphs. There are several generalizations of the Regularity Lemma of Szemerédi [11] to hypergraphs, due to various authors, e.g. [2,5,6,9]. Here, we will use the most straightforward one (see Chung [2]).

Let $\varepsilon > 0$, let V_1, V_2, V_3 be disjoint vertex sets of size m, and let $\mathcal{H}$ be a hypergraph such that each edge of $\mathcal{H}$ contains exactly one element of each V_i for $i = 1, 2, 3$. Let $d = |\mathcal{H}|/m^3$. Then $\mathcal{H}$ is said to be *ε-regular of density d* if for every choice of $X_i \subseteq V_i$ with $|X_1||X_2||X_3| > \varepsilon m^3$ we have

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

Here by $\mathcal{H}[X_1, X_2, X_3]$ we mean the subhypergraph of $\mathcal{H}$ induced by the vertex set $X_1 \cup X_2 \cup X_3$. We often refer to the hypergraph $\mathcal{H} = \mathcal{H}[V_1, V_2, V_3]$ as an ε-regular *triple*. In this setting the (weak) regularity lemma for hypergraphs from [2] can be stated as follows:

Theorem 2.1. *Let $\varepsilon > 0$ and $t_0 \in \mathbb{N}$ be given. Then there exist $t(\varepsilon, t_0)$ and $N(\varepsilon, t_0)$ such that every 3-uniform hypergraph $\mathcal{M}$ with $N = |V(\mathcal{M})| > N(\varepsilon, t_0)$ vertices has a partition $V_0 \cup V_1 \cup \dots \cup V_t$ of its vertex set $V(\mathcal{M})$, where $t_0 \leqslant t \leqslant t(\varepsilon, t_0)$, such that*

(1) $|V_0| < \varepsilon N$,
(2) $|V_i| = m$ *for* $i = 1, \dots, t$, *where* $(1-\varepsilon)N/t \leqslant m \leqslant N/t$,
(3) *all but at most* εN^3 *edges of* $\mathcal{M}$ *lie in some ε-regular triple* $\mathcal{M}[V_i, V_j, V_k]$ *with* $1 \leqslant i < j < k \leqslant t$.

We will prove Theorem 1.2 in three steps. In the first step, given a 2-colouring of the edges of $K_N^{(3)}$ for suitably chosen N, we apply Theorem 2.1 to the subhypergraph consisting of the red edges, to obtain a new 2-coloured structure called the *cluster* hypergraph. In the second step we show that the cluster hypergraph contains a monochromatic subhypergraph $\mathcal{L}$ with certain special properties. Then in the third step, we prove that the monochromatic subhypergraph of $K_N^{(3)}$ corresponding to $\mathcal{L}$ contains a copy of $\mathcal{C}_n$.

Step 1: Let the parameter η from Theorem 1.2 be given. We may assume $\eta < 1/5$. Let $\varepsilon_0 < 500^{-24}$ be small enough such that $0 < g(\varepsilon_0) < 1+\eta$, where $g(\varepsilon_0) = (1-500\varepsilon_0^{1/24})^{-3}$, and $t_0 = \lceil \varepsilon_0^{-6} \rceil$. We set $n_0 = 100t(\varepsilon_0, t_0)^2 M(\varepsilon_0, t_0)\varepsilon_0^{-1}$, where $t(\varepsilon_0, t_0)$ and $M(\varepsilon_0, t_0)$ are chosen in such a way that the assertion of Theorem 2.1 holds. For $n > n_0$, consider an arbitrary colouring of the edges of $K_N^{(3)}$ with red and blue, where $N = 5g(\varepsilon_0)n/4$. Let $\mathcal{J}$

denote the subhypergraph consisting of the edges coloured red, and let $\bar{\mathcal{J}}$ denote the blue subhypergraph. We apply Theorem 2.1 to $\mathcal{J}$ with ε_0 to obtain a partition $V_0 \cup V_1 \cup \cdots \cup V_t$ of $V(K_N^{(3)})$ with the properties stated in Theorem 2.1. Then we have the following lemma, whose very standard proof appears in Section 3.

Lemma 2.2. *Let $\varepsilon = \varepsilon_0^{1/4}$. Then*

(1) *all but at most εt^3 triples $\mathcal{J}[V_i, V_j, V_k]$ are ε-regular,*
(2) *if $\mathcal{J}[V_i, V_j, V_k]$ is ε-regular with density d then $\bar{\mathcal{J}}[V_i, V_j, V_k]$ is ε-regular with density $1 - d$,*
(3) $20t < \varepsilon m$.

We define the *cluster hypergraph* $\mathcal{J}_0$ as follows. The vertex set of $\mathcal{J}_0$ is $\{1, \ldots, t\}$, and *ijk* forms an edge of $\mathcal{J}_0$ precisely when $\mathcal{J}[V_i, V_j, V_k]$ is ε-regular (of some density). Then by Lemma 2.2, all but at most εt^3 of the triples *ijk* are edges of $\mathcal{J}_0$.

We consider the edges of $\mathcal{J}_0$ to be coloured red or blue, as follows. We colour the edge $ijk \in \mathcal{J}_0$ red if $\mathcal{J}[V_i, V_j, V_k]$ has density at least $1/2$, otherwise we colour it blue. Then by Lemma 2.2(2), if *ijk* is coloured blue, then $\bar{\mathcal{J}}[V_i, V_j, V_k]$ is ε-regular of density more than $1/2$. This completes Step 1.

Step 2: To describe Step 2 we need to introduce a few more definitions. Let $\mathcal{M}$ be a 3-uniform hypergraph. The *shadow* graph $\Gamma(\mathcal{M})$ of $\mathcal{M}$ is defined on the vertex set $V(\mathcal{M})$ by joining vertices x and y by an edge if and only if there exists an edge $xyz \in \mathcal{M}$. We call $\mathcal{M}$ a *connected* hypergraph if $\Gamma(\mathcal{M})$ is connected in the ordinary graph sense. A subhypergraph of $\mathcal{M}$ that is maximal with respect to being connected is called a *component* of $\mathcal{M}$. Thus a component of $\mathcal{M}$ is determined by its set of vertices. We will use this concept when $\mathcal{M}$ is the subhypergraph of a 2-coloured hypergraph $\mathcal{K}$ consisting of all the red edges (or all the blue edges). A component of $\mathcal{M}$ will be called a *monochromatic component* of $\mathcal{K}$, so each monochromatic component is either red or blue. Thus each vertex in a 2-coloured hypergraph $\mathcal{K}$ is in one red monochromatic component and one blue monochromatic component.

We also define a special small hypergraph called a *diamond*. A diamond $\mathcal{D}$ has vertex set $\{x_1, x_2, x_3, x_4\}$ and edge set $\{x_1x_2x_3, x_2x_3x_4\}$. The two vertices x_2 and x_3 of degree 2 in $\mathcal{D}$ are called the *central points* of $\mathcal{D}$.

Step 2 of the proof of Theorem 1.2 is accomplished by the following lemma.

Lemma 2.3. *The cluster hypergraph $\mathcal{J}_0$ has a monochromatic component $\mathcal{L}$ that contains $s = \lceil f(\varepsilon)\frac{n}{N}\frac{t}{4}\rceil$ vertex-disjoint diamonds $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_s$, where $f(\varepsilon) = (1 - 500\varepsilon^{1/6})^{-2}$.*

The proof of Lemma 2.3 appears in Section 4. Note that since $N = 5g(\varepsilon_0)n/4$, and $f(\varepsilon)$ and $g(\varepsilon_0)$ are very close to 1, the parameter s is approximately equal to $t/5$.

Step 3: To complete Step 3 we show that the monochromatic component $\mathcal{L}$ guaranteed by Lemma 2.3 corresponds to a monochromatic subhypergraph $\mathcal{H}$ of $\mathcal{J}^* = \mathcal{J}$ or $\bar{\mathcal{J}}$ that contains a copy of $\mathcal{C}_n$. If $\mathcal{L}$ is red then we let $\mathcal{J}^* = \mathcal{J}$, if $\mathcal{L}$ is blue then we choose $\mathcal{J}^* = \bar{\mathcal{J}}$. Then $\mathcal{H}$ is defined in the natural way, that is, $V(\mathcal{H}) = \bigcup\{V_i : i \in V(\mathcal{L})\}$ and $\mathcal{H} = \bigcup\{\mathcal{J}^*[V_i, V_j, V_k] : ijk \in \mathcal{L}\}$.

Lemma 2.4. *$\mathcal{H}$ contains a copy of $\mathcal{C}_n$.*

To prove Lemma 2.4, we first trace a "route" in the monochromatic component $\mathcal{L}$ of the cluster hypergraph $\mathcal{J}_0$, that visits all of the s disjoint diamonds $\mathcal{D}_1, \ldots, \mathcal{D}_s$ in $\mathcal{L}$ (see Lemma 5.1). Then (using Lemma 5.3), we choose a collection of short loose paths (of length three or six) in the hypergraph $\mathcal{H}$, that link together to form a cycle, following the chosen route. Finally, to obtain the cycle $\mathcal{C}_n$ in $\mathcal{H}$ we "blow-up" $s \sim t/5$ short paths (of length 3) corresponding to diamonds by long paths (each of length $\sim 4m \sim 4N/t \sim 5n/t$). More precisely, for each diamond $\mathcal{D}_i = \{hjk, hjp\}$, we replace the short path that starts in V_h and ends in V_j by a long path with the same end-vertices, that uses almost all the vertices in $V_h \cup V_j \cup V_k \cup V_p$. (This step uses Lemma 5.5.) Note that these long paths are mutually vertex disjoint since all diamonds $\mathcal{D}_i$ are vertex disjoint. Therefore, to obtain a cycle, we just need to make sure that the short paths do not intersect and they do not interfere with the long paths. The full proof of Lemma 2.4 appears in Section 5.

Note that the above strategy should work with any chosen simple hypergraph $\mathcal{U}$ in place of the diamond $\mathcal{D}$, provided

- enough disjoint copies of $\mathcal{U}$ can be found to cover $\sim 4t/5$ vertices of $\mathcal{J}_0$, and
- each $\mathcal{U}$ corresponds to a subhypergraph of $\mathcal{J}^*$ that contains a path covering almost all its vertices.

For example, taking $\mathcal{U}$ to be just a single edge may seem more natural than choosing $\mathcal{U} = \mathcal{D}$, the diamond. Then we would want $\mathcal{L}$ to be a component containing a matching (a set of disjoint edges) of size $\sim 4t/15$. However, unfortunately, such a monochromatic component is not guaranteed to exist in every 2-coloured $\mathcal{J}_0$. A construction similar to that given in Lemma 1.1, but with $|A| = t/4$ and $|B| = 3t/4$, gives a colouring of $\mathcal{J}_0$ in which the largest monochromatic matching has size $t/4$. Thus we need to choose $\mathcal{U}$ to be something more than just a single edge, and the diamond turns out to be a suitable choice. Note that the single blue component in the above example in fact contains $t/4$ disjoint diamonds. For the colouring with proportions essentially as given in Lemma 1.1, with $|A| = t/5$ and $|B| = 4t/5$, each of the red and the blue components contains exactly $t/5$ disjoint diamonds, showing that Lemma 2.3 is asymptotically best possible.

3. Proof of Lemma 2.2

With the definitions given in Step 1, we consider the first assertion of Lemma 2.2. We begin by claiming that if $|\mathcal{J}[V_i, V_j, V_k]| < \varepsilon_0^{1/2} m^3$, then it is ε-regular. To see this, note that $0 \leqslant d < \varepsilon_0^{1/2} < \varepsilon$ for this triple, so if $|X_i||X_j||X_k| > \varepsilon m^3$, then

$$0 \leqslant \frac{|\mathcal{J}[X_i, X_j, X_k]|}{|X_i||X_j||X_k|} < \frac{\varepsilon_0^{1/2} m^3}{\varepsilon_0^{1/4} m^3} = \varepsilon.$$

Therefore the density of $\mathcal{J}[X_i, X_j, X_k]$ differs from d by less than ε, as required.

Suppose on the contrary that at least εt^3 triples are not ε-regular in $\mathcal{J}$. Then they are certainly not ε_0-regular either. By the above claim, the total number of edges of $\mathcal{J}$ contained in these triples is at least $\varepsilon\varepsilon_0^{1/2}m^3t^3 > \varepsilon_0^{3/4}(1-\varepsilon_0)^3N^3$ by Theorem 2.1(2). But since $\varepsilon_0 < 1/5000$ this number is more than $\varepsilon_0 N^3$, contradicting Theorem 2.1(3). Therefore the first assertion of Lemma 2.2 holds.

To check the second assertion, note that

$$|\bar{\mathcal{J}}[X_i, X_j, X_k]| = |X_i||X_j||X_k| - |\mathcal{J}[X_i, X_j, X_k]|$$

holds for arbitrary sets $X_i \subseteq V_i$, $X_j \subseteq V_j$ and $X_k \subseteq V_k$. In particular, the densities $d_{\bar{\mathcal{J}}}$ of $\bar{\mathcal{J}}[V_i, V_j, V_k]$ and $d_{\mathcal{J}}$ of $\mathcal{J}[V_i, V_j, V_k]$ are related by $d_{\bar{\mathcal{J}}} = 1 - d_{\mathcal{J}}$. Thus if $\mathcal{J}[V_i, V_j, V_k]$ is ε-regular then for any sets $X_i \subseteq V_i$, $X_j \subseteq V_j$ and $X_k \subseteq V_k$ with $|X_i||X_j||X_k| > \varepsilon m^3$ we have

$$\frac{|\bar{\mathcal{J}}[X_i, X_j, X_k]|}{|X_i||X_j||X_k|} = 1 - \frac{|\mathcal{J}[X_i, X_j, X_k]|}{|X_i||X_j||X_k|},$$

and so this number differs from $d_{\bar{\mathcal{J}}} = 1 - d_{\mathcal{J}}$ by less than ε as required.

The third assertion is simply a calculation from the definitions given in Step 1. We have $N > n > \frac{100t(\varepsilon_0)^2}{\varepsilon_0} > \frac{100t^2}{\varepsilon}$ so $\varepsilon(1-\varepsilon)N > 20t^2$. This together with Theorem 2.1(2) implies $\varepsilon m > \varepsilon(1-\varepsilon_0)\frac{N}{t} > \varepsilon(1-\varepsilon)\frac{N}{t} > 20t$ as required.

4. Proof of Lemma 2.3

We begin this section by focusing on a subhypergraph of $\mathcal{J}_0$ with convenient properties.

Lemma 4.1. *The hypergraph $\mathcal{J}_0$ contains a subhypergraph $\mathcal{J}_1$ with the following properties. Here $\varepsilon_1 = 10\varepsilon^{1/6}$.*

(1) *$\mathcal{J}_1$ has $t_1 > (1-\varepsilon_1)t$ vertices,*
(2) *if some edge of $\mathcal{J}_1$ contains the vertices x and y then more than $(1-\varepsilon_1)t$ edges contain both x and y,*
(3) *for every vertex x, more than $(1-\varepsilon_1)t$ pairs xy are such that at least $(1-\varepsilon_1)t$ edges contain both x and y.*

Proof. Recall from Step 1 (Lemma 2.2) that $\mathcal{J}_0$ has t vertices and at least $\binom{t}{3} - \varepsilon t^3$ edges. We call a vertex x of $\mathcal{J}_0$ *bad* if it is in fewer than $\binom{t-1}{2} - \varepsilon^{1/2}t^2$ edges. We first estimate the number αt of bad vertices by counting the edges of $\mathcal{J}_0$ over the vertices as follows:

$$3|\mathcal{J}_0| \leqslant (1-\alpha)t\binom{t-1}{2} + \alpha t\left(\binom{t-1}{2} - \varepsilon^{1/2}t^2\right).$$

Therefore

$$\begin{aligned}\binom{t}{3} - \varepsilon t^3 &\leqslant (1-\alpha)\binom{t}{3} + \alpha\binom{t}{3} - \alpha\varepsilon^{1/2}t^3/3\\ &= \binom{t}{3} - \alpha\varepsilon^{1/2}t^3/3.\end{aligned}$$

Thus the number αt of bad vertices must be at most $3\varepsilon^{1/2}t$. Remove these, and denote the resulting hypergraph by $\mathcal{J}'$. Then every vertex of $\mathcal{J}'$ is in at least $\binom{t-1}{2} - \varepsilon^{1/2}t^2 - 3\varepsilon^{1/2}t^2 = \binom{t-1}{2} - 4\varepsilon^{1/2}t^2$ edges.

We call a pair xy of vertices of $\mathcal{J}'$ *bad* if it is contained in fewer than $(1 - 2\varepsilon^{1/4})t$ edges. Then the maximum degree of the graph on $V(\mathcal{J}')$ formed by bad pairs is less than $4\varepsilon^{1/4}t$, by a similar double counting argument using the vertex property of $\mathcal{J}'$.

We let $\mathcal{J}_1$ be the hypergraph formed by removing all edges that contain bad pairs. To check (1), note that $|V(\mathcal{J}_1)| = |V(\mathcal{J}')| > t - 3\varepsilon^{1/2}t > (1 - \varepsilon_1)t$.

Property (2) is true because if an edge remains, no pair in it was bad. Therefore, if x and y are contained in this edge then they were contained in at least $(1-2\varepsilon^{1/4})t$ edges of $\mathcal{J}'$. Thus, the number of edges lost for xy is at most $4\varepsilon^{1/4}t$ for the bad pairs incident to x, plus $4\varepsilon^{1/4}t$ for the bad pairs incident to y. Therefore xy is contained in at least $(1 - 10\varepsilon^{1/4})t > (1 - \varepsilon_1)t$ edges of $\mathcal{J}_1$.

Finally to verify (3), note that the above argument tells us that each pair xy that was not bad satisfies the property in (2). Thus x is incident to more than $|V(\mathcal{J}_1)| - 4\varepsilon^{1/4}t > (1 - 3\varepsilon^{1/2} - 4\varepsilon^{1/4})t > (1 - \varepsilon_1)t$ such pairs. □

The following lemma is a simple calculation that relates s to the parameters of $\mathcal{J}_1$.

Lemma 4.2. *We have* $s \leqslant \frac{1}{5}|V(\mathcal{J}_1)| - 6\varepsilon_1 t$.

Proof. Recall that $s = \lceil f(\varepsilon)\frac{n}{N}\frac{t}{4}\rceil$, where $f(\varepsilon)$ is as defined in Lemma 2.3, and $N = 5g(\varepsilon_0)n/4$ (see Step 1). Then since $\varepsilon = \varepsilon_0^{1/4}$ (see Lemma 2.2) and $\varepsilon_1 = 10\varepsilon^{1/6}$ (see Lemma 4.1) we have $N = 5(1 - 50\varepsilon_1)^{-1} f(\varepsilon)n/4$. Therefore

$$f(\varepsilon)\frac{n}{N}\frac{5t}{4} = (1 - 50\varepsilon_1)t = (1 - \varepsilon_1)t - 49\varepsilon_1 t.$$

Hence $f(\varepsilon)\frac{n}{N}\frac{t}{4} < \frac{1}{5}(1 - \varepsilon_1)t - 9\varepsilon_1 t$, which implies the desired result because $|V(\mathcal{J}_1)| > (1 - \varepsilon_1)t$ by Lemma 4.1(1). □

We are interested in the monochromatic components of the 2-coloured hypergraph $\mathcal{J}_1$ (as defined in Step 2). Note that each vertex of $\mathcal{J}_1$ is in one red monochromatic component and one blue monochromatic component.

Lemma 4.3. *The hypergraph $\mathcal{J}_1$ has the following properties.*

(1) *$\mathcal{J}_1$ contains a monochromatic component with more than $(1 - 3\varepsilon_1)t$ vertices. Let us say without loss of generality that this component $\mathcal{R}$ is red.*
(2) *if $\mathcal{R}$ does not contain s disjoint diamonds, then the largest blue component $\mathcal{B}$ has at least $4|V(\mathcal{J}_1)|/5$ vertices.*

Proof. For the first assertion, consider the shadow graph $G = \Gamma(\mathcal{J}_1)$ as defined in Step 2. Then the minimum degree of G is more than $(1 - \varepsilon_1)t$ by Lemma 4.1(3). We call an edge xy of G *red* if it is contained in a red edge of $\mathcal{J}_1$, and blue if it is in a blue edge of

$\mathcal{J}_1$. Note that some edges may be both red and blue. Then a monochromatic component of $\mathcal{J}_1$ is by definition the same as a monochromatic component of G. Suppose that the largest monochromatic (say blue) component C in G does not cover all vertices of G. Certainly $|C| > (1-\varepsilon_1)t/2$ because there is a monochromatic star of this size by the minimum degree bound. Let x be a vertex outside C. Then all the edges of G joining x to a vertex of C are red, and there are more than $|C| - \varepsilon_1 t$ such edges. Let $y \in C$ be a neighbour of x, then all the edges of G joining y to $V(G) \setminus C$ are red as well, and there are more than $|V(G) \setminus C| - \varepsilon_1 t > (1-2\varepsilon_1)t - |C|$ such edges. Then this gives a red component in G of size more than $(1-3\varepsilon_1)t$ as required.

To prove (2), suppose the largest blue component $\mathcal{B}$ has at most $4|V(G)|/5$ vertices. We first claim that there exists a set $A \subset V(G)$ with $|V(G)|/5 \leqslant |A| \leqslant |V(G)|/2$ such that all edges of G that contain a vertex of A and a vertex of $\bar{A} = V(G) \setminus A$ are red but not blue. This is clearly true if $|V(\mathcal{B})| \geqslant |V(G)|/5$, since either $V(\mathcal{B})$ or $V(G) \setminus V(\mathcal{B})$ would do. If $|V(\mathcal{B})| < |V(G)|/5$ then there is a subset S of the blue components such that $A' = \bigcup_{\mathcal{B}_i \in S} V(\mathcal{B}_i)$ satisfies $2|V(G)|/5 \leqslant |A'| \leqslant 3|V(G)|/5$. (Note that some blue components may just be isolated vertices.) Then again either A' or $V(\mathcal{J}_1) \setminus A'$ is a suitable choice for A.

Note then that the red component of $\mathcal{J}_1$ given by these red edges of G joining A and $\bar{A}$ must be $\mathcal{R}$. Next we show that this structure for $\mathcal{R}$ guarantees that it contains s disjoint diamonds. We construct greedily a set of at least $\frac{1}{6}(4|A| - |\bar{A}|) - \varepsilon_1 t$ disjoint diamonds in $\mathcal{R}$ as follows. Suppose $i < \frac{1}{6}(4|A| - |\bar{A}|) - \varepsilon_1 t$, and that disjoint diamonds $\mathcal{D}_1, \ldots, \mathcal{D}_{i-1}$ have already been found in $\mathcal{R}$, each of which has two vertices in each of A and $\bar{A}$. Then there remain at least $|A| - 2(i-1) \geqslant 2\varepsilon_1 t$ unused vertices in A (and hence also in $\bar{A}$). Then by Lemma 4.1(3), there is an edge xy of G with $x \in A$ and $y \in \bar{A}$ both unused, which is contained in more than $(1-\varepsilon_1)t$ edges of $\mathcal{J}_1$, which are all red by our choice of A. Therefore there exists another unused vertex $w \in A$, and another unused $z \in \bar{A}$, such that xyw and xyz are both red edges of $\mathcal{J}_1$. Then we set $\mathcal{D}_i$ to be this red diamond.

When we have found $q = \max\{0, \lceil \frac{1}{6}(4|A| - |\bar{A}|) - \varepsilon_1 t \rceil\}$ diamonds in $\mathcal{R}$ as above, we add $\frac{1}{3}(|\bar{A}| - |A|) - 2\varepsilon_1 t$ more diamonds that have one vertex in A and three in $\bar{A}$. Suppose $i < \frac{1}{3}(|\bar{A}| - |A|) - 2\varepsilon_1 t$, and that disjoint diamonds $\mathcal{D}_1, \ldots, \mathcal{D}_{i-1}$ have already been found in $\mathcal{R}$, each of which has one vertex in A and three in $\bar{A}$. Then there remain at least $|A| - 2q - i - 1 \geqslant 2\varepsilon_1 t$ unused vertices in A and $|\bar{A}| - 2q - 3(i-1) \geqslant 8\varepsilon_1 t$ in $\bar{A}$. Then by Lemma 4.1(3), there is an edge xy of G with $x \in A$ and $y \in \bar{A}$ both unused, which is contained in more than $(1-\varepsilon_1)t$ edges of $\mathcal{J}_1$, which are all red. Therefore there exist two more unused vertices $w, z \in \bar{A}$, such that xyw and xyz are both red edges of $\mathcal{J}_1$. Then we set $\mathcal{D}_i$ to be this red diamond.

The total number of disjoint diamonds we find by this greedy construction is at least $\frac{1}{6}(4|A| - |\bar{A}|) - \varepsilon_1 t + \frac{1}{3}(|\bar{A}| - |A|) - 2\varepsilon_1 t \geqslant \frac{|V(G)|}{5} - 3\varepsilon_1 t$, which is more than s by Lemma 4.2. $\square$

In addition to Lemma 4.3, the following technical lemma will be important for the proof of Lemma 2.3. By a *diadem* we mean any coloured hypergraph with 5 vertices that contains both a red diamond and a blue diamond.

Lemma 4.4. *Let $X = \{x_1, x_2, x_3, x_4\}$ and $Y = \{y_1, y_2, y_3, y_4, y_5\}$ be disjoint vertex sets such that $x_i x_j y_k \in \mathcal{J}_1$ for all i, j, k except possibly when $\{i, j\} = \{1, 4\}$. Let $S \subset X$ be* *a subset of size at least two. Suppose further that $x_1x_2x_3, x_2x_3x_4 \in \mathcal{J}_1$ and are both red. Then one of the following holds.*

(1) *$X \cup Y$ contains two disjoint red diamonds, both of which intersect X.*
(2) *$X \cup Y$ contains a diadem $\mathcal{E}$ such that $3 \leqslant |V(\mathcal{E}) \cap X| \leqslant 4$ whose blue diamond intersects both S and Y.*

Proof. Since $|S| \geqslant 2$, either we have both $\{x_1, x_2\} \cap S \neq \emptyset$ and $\{x_3, x_4\} \cap S \neq \emptyset$, or $\{x_1, x_3\} \cap S \neq \emptyset$ and $\{x_2, x_4\} \cap S \neq \emptyset$. Since x_2 and x_3 are symmetric in the assumptions of the lemma, we may assume without loss of generality that the former possibility holds. We look first at the edges $x_1x_2y_i$, $i = 1, \dots, 5$. We argue that if more than one of them is blue, then $X \cup Y$ contains a diadem.

Indeed, suppose that, say, the edges $x_1x_2y_1$ and $x_1x_2y_2$ are blue. Then, if the edge $x_1x_3y_1$ is blue, we take the vertex set W of $\mathcal{E}$ to be $W = \{x_1, x_2, x_3, x_4, y_1\}$, and, if $x_1x_3y_1$ is red, then we put $W = \{x_1, x_2, x_3, y_1, y_2\}$. Consequently, we may assume that, say, the edges $x_1x_2y_i$, $i = 1, \dots, 4$, are coloured red.

An analogous argument shows that, to avoid a diadem, at most one of the edges $x_3x_4y_i$, $i = 1, \dots, 4$, is blue, so suppose that $x_3x_4y_i$, $i = 1, 2, 3$, are red. But then $X \cup Y$ contains two disjoint red diamonds with edges $x_1x_2y_4$, $x_1x_2y_3$, $x_3x_4y_2$ and $x_3x_4y_1$. $\square$

We are now ready to prove Lemma 2.3.

Proof of Lemma 2.3. Let $\{\mathcal{D}_1, \dots, \mathcal{D}_q\}$ be a set of disjoint diamonds in $\mathcal{R}$ of maximum size q. If $q \geqslant s$ we have proved the lemma, so let us assume $q < s$. Our aim is to show that the largest blue component $\mathcal{B}$ contains s disjoint diamonds.

Our first step is to replace the red diamonds in $\mathcal{R}$ one-by-one by disjoint diadems. Suppose $0 \leqslant i < q$, let $Q_i = \bigcup_{j=i+1}^{q} V(\mathcal{D}_j)$, and suppose that disjoint diadems $\mathcal{E}_1, \dots, \mathcal{E}_i$ have been found such that

(a) $V(\mathcal{E}_1 \cup \dots \cup \mathcal{E}_i) \cap Q_i = \emptyset$,
(b) each $\mathcal{E}_j$ has its red diamond in $\mathcal{R}$ and its blue diamond in $\mathcal{B}$,
(c) the set $V_i = V(\mathcal{J}_1) \setminus V(\mathcal{E}_1 \cup \dots \cup \mathcal{E}_i)$ contains at least $\frac{4}{5}|V_i|$ vertices of $\mathcal{B}$.

Lemma 4.3(2) guarantees that this process can start.

We now consider two cases, and in each case our aim is to find a diadem $\mathcal{E}_{i+1}$ such that (a), (b), and (c) hold for $i + 1$ in place of i. The next lemma lays out two useful facts that will be used in the case analysis to follow.

Lemma 4.5. *If* (a), (b), *and* (c) *hold then*

(1) $|V_i \setminus Q_i| = |V(\mathcal{J}_1)| - 4q - i$.
(2) *Let $X = V(\mathcal{D}_j)$ for some $j \in \{i + 1, \dots, q\}$ and let $Z \subset V_i \setminus Q_i$ be a set of size at least $5\varepsilon_1 t + 5$. Then there exists a set $Y \subset Z$, $|Y| = 5$, such that $X \cup Y$ satisfies the assumptions of Lemma* 4.4.

Proof. The proof of (1) is immediate since by the definitions we have $|V_i \setminus Q_i| = |V(\mathcal{J}_1)| - 5i - 4(q - i) = |V(\mathcal{J}_1)| - 4q - i$.

To prove (2), let x_1, x_2, x_3, x_4 be the vertices of X such that $x_1x_2x_3$ and $x_2x_3x_4$ are the edges of $\mathcal{D}_j$. Then each of the five pairs $\{i, j\} \subset \{1, 2, 3, 4\}$ with $\{i, j\} \neq \{1, 4\}$ is such that

$x_i x_j$ is in an edge of $\mathcal{J}_1$. Thus by Lemma 4.1(2), at most $\varepsilon_1 t$ vertices $z \in Z$ are such that $x_i x_j z$ is not an edge of $\mathcal{J}_1$. Therefore since $|Z| \geqslant 5\varepsilon_1 t + 5$, there are at least five vertices z such that $x_i y_j z$ is an edge of $\mathcal{J}_1$ for each of the five pairs $\{i, j\} \neq \{1, 4\}$, as required. $\square$

We consider two cases according to the size of $V_i \setminus (V(\mathcal{B}) \cup Q_i)$.

Case 1: $|V_i \setminus (V(\mathcal{B}) \cup Q_i)| < 5\varepsilon_1 t + 5$.

Using Lemma 4.5(1) we see that $|V_i \setminus Q_i| \geqslant |V(\mathcal{J}_1)| - 5q > |V(\mathcal{J}_1)| - 5s$, so $|V_i \setminus Q_i| > 10\varepsilon_1 t + 10$ by Lemma 4.2. Thus $|(V_i \setminus Q_i) \cap V(\mathcal{B})| \geqslant 5\varepsilon_1 t + 5$.

We choose $j \in \{i+1, \ldots, q\}$ such that $\mathcal{D}_j$ has the smallest possible number p of vertices of $\mathcal{B}$. Then with $X = V(\mathcal{D}_j)$ and $Z = (V_i \setminus Q_i) \cap V(\mathcal{B})$, by Lemma 4.5(2) there exists $Y \subset (V_i \setminus Q_i) \cap V(\mathcal{B})$, $|Y| = 5$, such that $X \cup Y$ satisfies the conditions of Lemma 4.4. We apply Lemma 4.4 with an arbitrary set $S \subset X$ of size two. Note that outcome (1) is not possible, since otherwise we would have a set of disjoint red diamonds of size $q + 1$ in $\mathcal{R}$, contradicting the definition of q. Therefore we have a diadem $\mathcal{E}$ satisfying (2), so its blue diamond intersects $Y \subset V(\mathcal{B})$, and hence is in $\mathcal{B}$. Since $|V(\mathcal{E}) \cap X| \geqslant 3$ the red diamond of $\mathcal{E}$ is in $\mathcal{R}$. Thus (a) and (b) hold for $\mathcal{E}_{i+1} = \mathcal{E}$.

Now if $p \leqslant 2$ then $X = V(\mathcal{D}_j)$ has at least two vertices that are not in $V(\mathcal{B})$, so since $|V(\mathcal{E}) \cap X| \geqslant 3$ we know $|V(\mathcal{E}) \cap V(\mathcal{B})| \leqslant 4$. Therefore $V_{i+1} = V_i \setminus V(\mathcal{E})$ contains at least $\frac{4}{5}|V_{i+1}|$ vertices of $\mathcal{B}$, so (c) holds. But if $p \geqslant 3$ then by definition of p we know

$$\begin{aligned}
|V_i \cap V(\mathcal{B})| &\geqslant \frac{3}{4}|V_i \cap Q_i| + |(V_i \setminus Q_i) \cap V(\mathcal{B})| \\
&\geqslant 3q - 3i + |V_i \setminus Q_i| - 5\varepsilon_1 t - 5 \\
&\geqslant 3q - 3i + |V(\mathcal{J}_1)| - 4q - i - 5\varepsilon_1 t - 5 \qquad \text{by Lemma 4.5(1)} \\
&= |V(\mathcal{J}_1)| - q - 4i - 5\varepsilon_1 t - 5 \\
&= \left(\frac{4}{5}|V(\mathcal{J}_1)| - 4i\right) + \frac{1}{5}|V(\mathcal{J}_1)| - q - 5\varepsilon_1 t - 5 \\
&\geqslant \frac{4}{5}|V_i| + 1,
\end{aligned}$$

where the last line follows since $|V_i| = |V(\mathcal{J}_1)| - 5i$ and $q < s \leqslant \frac{1}{5}|V(\mathcal{J}_1)| - 5\varepsilon_1 t - 6$ by Lemma 4.2. Therefore $|V_{i+1} \cap V(\mathcal{B})| \geqslant |V_i \cap V(\mathcal{B})| - 5 \geqslant \frac{4}{5}|V_i| - 4 = \frac{4}{5}(|V_i| - 5) = \frac{4}{5}|V_{i+1}|$, verifying (c).

Case 2: $|V_i \setminus (V(\mathcal{B}) \cup Q_i)| \geqslant 5\varepsilon_1 t + 5$.

First suppose that some $j \in \{i+1, \ldots, q\}$, say without loss of generality $j = i + 1$, is such that $\mathcal{D}_j$ has at least 2 vertices in $\mathcal{B}$. Set $X = V(\mathcal{D}_{i+1})$. By Lemma 4.5(2) with $Z = V_i \setminus (V(\mathcal{B}) \cup Q_i)$, there exists $Y \subset V_i \setminus (V(\mathcal{B}) \cup Q_i)$, $|Y| = 5$, such that $X \cup Y$ satisfies the conditions of Lemma 4.4. We let $S = X \cap V(\mathcal{B})$ and apply Lemma 4.4. Again note that outcome (1) is not possible, since otherwise we would have a set of disjoint red diamonds of size $q + 1$ in $\mathcal{R}$. Therefore we have a diadem $\mathcal{E}$ satisfying (2). But then the blue diamond in $\mathcal{E}$ intersects S and hence is in $\mathcal{B}$, the red diamond of $\mathcal{E}$ intersects X and hence is in $\mathcal{R}$, and $|V(\mathcal{E}) \cap V(\mathcal{B})| \leqslant 4$ because $Y \cap V(\mathcal{B}) = \emptyset$. Therefore $V_{i+1} = V_i \setminus V(\mathcal{E})$ contains at least $\frac{4}{5}|V_{i+1}|$ vertices of $\mathcal{B}$. Thus (a), (b), and (c) hold for $i + 1$ if we set $\mathcal{E}_{i+1} = \mathcal{E}$.

Now suppose that each $\mathcal{D}_j$ with $j \in \{i+1, \ldots, q\}$ has at most one vertex of $\mathcal{B}$. We claim that in this case we must have $|(V_i \setminus Q_i) \cap V(\mathcal{B})| \geqslant 5\varepsilon_1 t + 5$. To see this, suppose the contrary. Then

$$\begin{aligned}\frac{4}{5}|V(\mathcal{J}_1)| - 4i &= \frac{4}{5}|V_i| \leqslant |V_i \cap V(\mathcal{B})| \\ &= |V_i \cap Q_i \cap V(\mathcal{B})| + |(V_i \setminus Q_i) \cap V(\mathcal{B})| \\ &\leqslant q - i + 5\varepsilon_1 t + 5.\end{aligned}$$

Here the first inequality comes from (c) and the last follows because each $\mathcal{D}_j$ has at most one point in $\mathcal{B}$. Therefore $\frac{4}{5}|V(\mathcal{J}_1)| \leqslant q + 3i + 5\varepsilon_1 t + 5 < 4s + 5\varepsilon_1 t + 5$, implying $\frac{1}{5}|V(\mathcal{J}_1)| < s + \frac{5}{4}\varepsilon_1 t + \frac{5}{4}$, which contradicts Lemma 4.2. Thus this cannot occur, and so we may assume $|(V_i \setminus Q_i) \cap V(\mathcal{B})| \geqslant 5\varepsilon_1 t + 5$.

Set $X = V(\mathcal{D}_{i+1})$ and $Z = (V_i \setminus Q_i) \cap V(\mathcal{B})$, and apply Lemma 4.5(2) to get a set $Y \subset (V_i \setminus Q_i) \cap V(\mathcal{B})$, $|Y| = 5$, such that $X \cup Y$ satisfies the conditions of Lemma 4.4. As in Case 1, we apply Lemma 4.4 with an arbitrary set $S \subset X$ of size two. Again we cannot get two disjoint red diamonds, so we get a diadem $\mathcal{E}$ satisfying (2). Its blue diamond intersects $Y \subset V(\mathcal{B})$, and hence is in $\mathcal{B}$. Since $|V(\mathcal{E}) \cap X| \geqslant 3$ the red diamond of $\mathcal{E}$ is in $\mathcal{R}$. Thus (a) and (b) hold for $\mathcal{E}_{i+1} = \mathcal{E}$. Note that $|V(\mathcal{E}) \cap X| \geqslant 3$ also implies that $|V(\mathcal{E}) \cap V(\mathcal{B})| \leqslant 3$, and so $V_{i+1} = V_i \setminus V(\mathcal{E})$ contains at least $\frac{4}{5}|V_{i+1}|$ vertices of $\mathcal{B}$, so (c) holds.

Therefore in all cases we can complete the above construction to find disjoint diadems $\mathcal{E}_1, \ldots, \mathcal{E}_q$, whose red diamonds are all in $\mathcal{R}$ and whose blue diamonds are all in $\mathcal{B}$, such that $V_q = V(\mathcal{J}_1) \setminus V(\mathcal{E}_1 \cup \cdots \cup \mathcal{E}_q)$ contains at least $\frac{4}{5}|V_q|$ vertices of $\mathcal{B}$. Note that $|V_q| = |V(\mathcal{J}_1)| - 5q \geqslant 30\varepsilon_1 t$ by Lemma 4.2, and so by Lemma 4.3(1) we have $|V_q \cap V(\mathcal{R})| \geqslant |V_q| - 3\varepsilon_1 t \geqslant 27\varepsilon_1 t$. Note also that by definition of q, no red diamond can be present in $V_q \cap V(\mathcal{R})$.

We now complete the proof by showing that $V_q \cap V(\mathcal{R}) \cap V(\mathcal{B})$ contains a set U of $s - q$ disjoint blue diamonds. Then U together with the q blue diamonds in $\mathcal{E}_1, \ldots, \mathcal{E}_q$ will give a set of s disjoint blue diamonds in $\mathcal{B}$ as required.

First we observe that for any $x \in V_q \cap V(\mathcal{R})$, since $|V_q \cap V(\mathcal{R})| > \varepsilon_1 t$, by Lemma 4.1(3) there exists $y \in V_q \cap V(\mathcal{R})$ such that xy is in at least $(1 - \varepsilon_1)t$ edges of $\mathcal{J}_1$. Therefore, since $V_q \cap V(\mathcal{R})$ contains no red diamonds, all but at most $\varepsilon_1 t + 1$ of the vertices $z \in V_q \cap V(\mathcal{R})$ are such that $xyz \in \mathcal{J}_1$ and xyz is blue. Therefore in particular

$$\begin{aligned}|V_q \cap V(\mathcal{R}) \cap V(\mathcal{B})| &\geqslant |V_q \cap V(\mathcal{R})| - \varepsilon_1 t - 1 \\ &\geqslant |V(\mathcal{J}_1)| - 5q - 4\varepsilon_1 t - 1 > 4(s - q) + 2\varepsilon_1 t,\end{aligned}$$

where the last line follows since $|V(\mathcal{J}_1)| > 5s + 30\varepsilon_1 t > 4s + q + 6\varepsilon_1 t + 1$ by Lemma 4.2.

We can now find the required set U of $s - q$ blue diamonds in $V_q \cap V(\mathcal{R}) \cap V(\mathcal{B})$ by a greedy construction. Suppose we have put into U some disjoint blue diamonds, and there remains a set W of at least $2\varepsilon_1 t$ unused vertices. Then as above, we can choose $x \in W$ and $y \in W$ such that there exist at least $|W \setminus \{x, y\}| - \varepsilon_1 t + 1 \geqslant 2$ vertices $z \in W \setminus \{x, y\}$ such that xyz is a blue edge of $\mathcal{J}_1$. Thus we can add another blue diamond to U. Thus since $|V_q \cap V(\mathcal{R}) \cap V(\mathcal{B})| > 4(s - q) + 2\varepsilon_1 t$ this construction yields a set U of at least $s - q$ disjoint blue diamonds as required. $\square$

5. Proof of Lemma 2.4

We follow the outline presented in Step 3. Let $G = \Gamma(\mathcal{L})$ be the shadow graph of $\mathcal{L}$, as defined in Step 2. Then for each of the s disjoint diamonds $\mathcal{D}_i = \{hjk, hjp\}$ in $\mathcal{L}$, there is an edge hj of G joining the two central points of $\mathcal{D}_i$. Let M denote the matching in G consisting of these s disjoint edges. By a *closed directed trail* in G we mean a sequence of (not necessarily distinct) vertices $x_1, x_2, \dots, x_r$, where $x_r = x_1$, of G, together with the edges (x_i, x_{i+1}) for $i = 1, \dots, r-1$ which we consider as being directed from x_i to x_{i+1} and call *arcs*.

The following easy lemma provides the route that our cycle will eventually follow.

Lemma 5.1. *Let $G = \Gamma(\mathcal{L})$, and let M be defined as above. Then G contains a closed directed trail with $r < 2t$ vertices that uses every edge of M as an arc, and no edge of G is used as an arc more than once in each direction.*

Proof. Recall that $|V(G)| \leqslant t$. Let T be any spanning tree of G that contains the edges of M. We can easily obtain a closed directed trail with the desired properties as follows. Fix a planar embedding of T, and let x_1 be the root of T. Construct the trail by taking a walk starting at x_1 around the boundary of the (single) face of the embedding, ending back at x_1. Then each edge of T is traversed exactly once in each direction, so the number of vertices of this trail is at most $2t - 1$. Moreover, every edge of M is used as an arc. $\square$

For each diamond $\mathcal{D}_i$, $1 \leqslant i \leqslant s$, we choose one arc that uses the edge of M corresponding to $\mathcal{D}_i$ and call it the *diamond arc* for $\mathcal{D}_i$.

When choosing short paths that will link up to form a cycle, we need to make sure the end-vertices of these paths can indeed coincide. Let V_h, V_j, and V_k be any vertex classes such that $\mathcal{H}[V_h, V_j, V_k]$ is ε-regular with density $d > 2\varepsilon^{1/2}$ (recall $|V_h| = |V_j| = |V_k| = m$), and for $i \in \{h, j, k\}$ let $U_i \subseteq V_i$ be arbitrary subsets. We say that a vertex $x \in V_h$ is *good* for the triple $\mathcal{H}[U_h, U_j, U_k]$ if

(i) for at least $d|U_j|/2$ vertices $y \in U_j$, there are at least $d|U_k|/2$ vertices $z \in U_k$ such that $xyz \in \mathcal{H}$, and
(ii) for at least $d|U_k|/2$ vertices $z \in U_k$, there are at least $d|U_j|/2$ vertices $y \in U_j$ such that $xyz \in \mathcal{H}$.

Note that for $x \in V_h$, the property of being good for $\mathcal{H}[U_h, U_j, U_k]$ is independent of the choice of U_h.

We define vertices in V_j and V_k to be good for $\mathcal{H}[U_h, U_j, U_k]$ in a similar way. The next lemma implies that, as long as U_h, U_j and U_k are reasonably large, most vertices in $V_h \cup V_j \cup V_k$ are good.

Lemma 5.2. *Suppose $d > 2\varepsilon^{1/2}$. With the above definitions, the number of vertices in V_h that are good for $\mathcal{H}[U_h, U_j, U_k]$ is at least $m - \frac{\varepsilon m^3}{|U_j||U_k|}$. The analogous bounds hold for V_j and V_k. In particular, for each $i \in \{h, j, k\}$ the number of vertices in V_i that are good for $\mathcal{H}[V_h, V_j, V_k]$ is at least $(1-\varepsilon)m$.*

Proof. Suppose $x \in V_h$ fails to satisfy (i). Then the total number of edges of $\mathcal{H}[\{x\}, U_j, U_k]$ is smaller than

$$(d|U_j|/2)|U_k| + (1 - d/2)|U_j|(d|U_k|/2) = (d - d^2/4)|U_j||U_k|.$$

If x fails to satisfy (ii) then we find the same bound. Let $X \subset V_h$ denote the set of vertices that are not good. Then $|\mathcal{H}[X, U_j, U_k]| \leqslant |X|(d - d^2/4)|U_j||U_k|$. Thus if $|X||U_j||U_k| > \varepsilon m^3$ we find

$$\frac{|\mathcal{H}[X, U_j, U_k]|}{|X||U_j||U_k|} \leqslant d - d^2/4 < d - \varepsilon,$$

which contradicts the fact that $\mathcal{H}[V_h, V_j, V_k]$ is ε-regular. Therefore the number of vertices in V_h that are not good for $\mathcal{H}[U_h, U_j, U_k]$ is at most $\frac{\varepsilon m^3}{|U_j||U_k|}$. Similar bounds can be found for V_j and V_k. □

The set of vertices in $V_h \cup V_j \cup V_k$ that are good for $\mathcal{H}[V_h, V_j, V_k]$ will simply be called *good*. To make sure the end-vertices of our short paths coincide where necessary, we will require them to be good. This idea will be made precise in the upcoming proof of Lemma 2.4.

We will also need to choose our short paths in ε-regular triples so that they avoid some "bad" set B of vertices. The next lemma allows us to find these short paths (of length three).

Lemma 5.3. *Let $\mathcal{H}[V_h, V_j, V_k]$ be an ε-regular triple with density $d > 2\varepsilon^{1/3}$. Then for every pair of good vertices $x \in V_h$ and $y \in V_j$, and for every set $B \subset V_h \cup V_j \cup V_k \setminus \{x, y\}$ that contains all non-good vertices and satisfies $|B \cap V_i| < (d/2 - \varepsilon^{1/3})m$ for $i \in \{h, j, k\}$, there is a path of length three in $\mathcal{H}[V_h, V_j, V_k]$ joining x to y that is disjoint from B (and hence contains only good vertices). Moreover the path can be chosen so that one vertex of degree two in the path is in V_h, and the other is in V_j.*

Proof. Since x is good, there exists a set $U_x \subset V_j$, $|U_x| \geqslant dm/2$, such that for each $z \in U_x$, there are at least $dm/2$ vertices $w \in V_k$ such that $xzw \in \mathcal{H}$. Since y is good, there exists a set $U_y \subset V_h$ with similar properties. Then, writing $b = (d/2 - \varepsilon^{1/3})$, we have

$$|U_x \setminus B||U_y \setminus B||V_k \setminus B| > (d/2 - b)^2(1 - b)m^3 > \varepsilon m^3.$$

Therefore, since $\mathcal{H}$ is ε-regular, we know that $|\mathcal{H}[U_x \setminus B, U_y \setminus B, V_k \setminus B]| \geqslant (d - \varepsilon)|U_x \setminus B||U_y \setminus B||V_k \setminus B|$. We may therefore choose vertices $z \in U_x \setminus \{y\}$ and $v \in U_y \setminus \{x\}$, and distinct good vertices $u_h, u_j, u_k \in V_k$ such that $xzu_h, yvu_j, zvu_k \in \mathcal{H}$. This gives the required path. □

The following immediate consequence of the above lemma will be used to adjust the parity of the length of a path.

Corollary 5.4. *Let $\mathcal{H} = \mathcal{H}[V_h, V_j, V_k]$ be an ε-regular triple with density $d > 2\varepsilon^{1/3}$. Then for every pair of good vertices $x \in V_h$ and $y \in V_j$, and for every set $B \subset V_h \cup V_j \cup V_k \setminus \{x, y\}$ that contains all non-good vertices and satisfies $|B \cap V_i| < (d/2 - \varepsilon^{1/3})m - 6$ for each $i \in \{h, j, k\}$, there is a path of length six in $\mathcal{H}$ joining x to y that is disjoint from B (and hence contains only good vertices).*

Proof. Let w be any vertex in $V_k \setminus B$. Then by Lemma 5.3 there is a path $\mathcal{P}$ of length three joining x and w that avoids $B \cup \{y\}$. Let $B' = B \cup V(\mathcal{P}) \setminus \{w\}$. Then applying Lemma 5.3 again with w, y and B' we may extend $\mathcal{P}$ to a path of length six with the required properties. □

Our final preparatory lemma tells us that in a "blown-up diamond" we can find long paths between specified vertices, that avoid any given small set B. We remark that the condition on the parity of the path is not essential, it is present simply to make the proof a bit easier.

Lemma 5.5. *Let V_h, V_j, V_k, V_p be classes such that $\mathcal{H}_1 = \mathcal{H}[V_h, V_j, V_k]$ and $\mathcal{H}_2 = \mathcal{H}[V_j, V_k, V_p]$ are both ε-regular of density $d \geqslant 32\varepsilon^{1/6}$. Let $x \in V_j$ and $y \in V_k$ be good vertices for both $\mathcal{H}_1$ and $\mathcal{H}_2$. Let $B \subset V_h \cup V_j \cup V_k \cup V_p \setminus \{x, y\}$ be such that B contains all vertices that are not good for $\mathcal{H}_1$ or not good for $\mathcal{H}_2$, and $|B \cap V_i| < 2\varepsilon m$ for each $i \in \{h, j, k, p\}$. Then there is a path joining x and y in $\mathcal{H}_1 \cup \mathcal{H}_2$ of length ℓ for all odd integers ℓ satisfying $2dm/7 \leqslant \ell \leqslant (1 - 2\varepsilon^{1/6})2m$ that avoids B.*

Proof. We prove the lemma by means of the following two claims. The first is an easy greedy argument using Lemma 5.3.

Claim 5.6. *Let $0 \leqslant q < \lfloor (d/2 - \varepsilon^{1/3})m/6 \rfloor$, and let $B' \subset V_h \cup V_j \cup V_k \setminus \{x, y\}$ be such that B' contains all vertices that are not good for $\mathcal{H}_1$, and $|B' \cap V_i| < (d/2 - \varepsilon^{1/3})m - 6q$ for each $i \in \{h, j, k\}$. Then there is a path joining x and y in $\mathcal{H}_1$ of length $6q + 3$ that avoids B', whose degree-two vertices all lie in V_j or V_k.*

Proof. We prove the claim by induction on q. For $q = 0$ it is immediately given by Lemma 5.3. Assume $q \geqslant 1$ and that the claim is true for smaller values of q. Let B' be given. Choose $z \in V_k \setminus (B' \cup \{y\})$, then z is good. Then by the induction hypothesis, there is a path $\mathcal{P}$ of length $6(q - 1) + 3$ joining x to z that avoids $B' \cup \{y\}$, and whose degree-two vertices all lie in V_j or V_k. Note then $|V(\mathcal{P}) \cap V_h| = 6(q - 1) + 3$ and $|V(\mathcal{P}) \cap V_j| = |V(\mathcal{P}) \cap V_k| = 3(q - 1) + 2$.

Let $B'' = B' \cup V(\mathcal{P}) \cup \{y\} \setminus \{z\}$. Then $|B'' \cap V_i| < (d/2 - \varepsilon^{1/3})m - 3$. Let $w \in V_j \setminus B''$. Then w is good, so by Lemma 5.3 there is a path $\mathcal{P}'$ of length three joining z and w that avoids B''. Finally let $B''' = B'' \cup V(\mathcal{P}') \setminus \{w, y\}$. Then $|B''' \cap V_i| < (d/2 - \varepsilon^{1/3})m$. Apply Lemma 5.3 once more to obtain a path $\mathcal{P}''$ of length 3 joining w to y that avoids B'''. The concatenation $\mathcal{P}\mathcal{P}'\mathcal{P}''$ then has length $6q + 3$, avoids B', and all its degree-two vertices lie in $V_j \cup V_k$. □

Our second claim states that any reasonably long path can be lengthened by two, unless it uses almost all the vertices of $\mathcal{H}_1 \cup \mathcal{H}_2$.

Claim 5.7. *With the assumptions of Lemma 5.5, let $\mathcal{P}$ be a path in $\mathcal{H}_1 \cup \mathcal{H}_2$ joining x and y of length ℓ, where $dm/4 < \ell < (1 - 2\varepsilon^{1/6})2m$, that avoids B. Suppose all degree-two vertices of $\mathcal{P}$ lie in $V_j \cup V_k$. Then there exists a path $\mathcal{P}'$ joining x and y that avoids B, all of whose degree-two vertices lie in $V_j \cup V_k$, of length $\ell + 2$.*

Proof. The aim is to replace one edge of $\mathcal{P}$ by three edges, however we must choose the edge to be replaced appropriately.

For each $i \in \{h, j, k, p\}$ let $U_i \subset V_i$ denote the available vertices in V_i, that is $U_i = V_i \setminus (V(\mathcal{P}) \cup B)$. Then by the condition on the length of $\mathcal{P}$, we have

$|U_i| > \varepsilon^{1/6}m$ for $i \in \{j,k\}$, and $|U_i| > \varepsilon^{1/6}m$ for some $i \in \{h,p\}$, say without loss of generality h.

Now by Lemma 5.2, for each $i \in \{h,j,k\}$ at most $\varepsilon^{2/3}m$ vertices of $V_i \cap V(\mathcal{P})$ are not good for $\mathcal{H}[U_h,U_j,U_k]$. Therefore, since $|V_i \cap V(\mathcal{P})| \geqslant \ell/2 \geqslant dm/8 > 4\varepsilon^{2/3}m$ holds for $i \in \{j,k\}$, there exist $x_1 \in V_j \cap V(\mathcal{P})$ and $y_1 \in V_k \cap V(\mathcal{P})$ that are both good for $\mathcal{H}[U_h,U_j,U_k]$, such that $x_1y_1z_1$ is an edge of $\mathcal{P}$ for some $z_1 \in V_h$. Let $U_k(x_1) \subset U_k$ be the set guaranteed by the goodness of x_1, so $|U_k(x_1)| \geqslant d|U_k|/2 > \varepsilon^{1/3}m$ has property (i). Let $U_j(y_1) \subset U_j$ be defined similarly (having property (ii)). Then we have $|U_h||U_j(y_1)||U_k(x_1)| > \varepsilon m^3$, and so there exist $u \in U_k(x_1)$ and $w \in U_j(y_1)$ such that $uwz \in \mathcal{H}$ for some $z \in U_h$. Thus, a path of length three from x_1 to y_1 can easily be found, that is disjoint from the rest of $\mathcal{P}$ and from B, and whose degree-two vertices lie in V_j and V_k. □

The proof of Lemma 5.5 now follows immediately: First, we find a B-avoiding path of odd length $\ell_0 = 6(\lfloor(2d/7-\varepsilon^{1/3})m/6\rfloor - 1) + 3 > dm/4$ from x to y using Claim 5.6. Then, since $\ell_0 < 2dm/7$, we extend it step by step using Claim 5.7 until a path of the desired odd length between $2dm/7$ and $(1-2\varepsilon^{1/6})2m$ is reached. □

We are now ready to prove Lemma 2.4.

Proof of Lemma 2.4. Let $\mathcal{L}$ be as in Lemma 2.3, and let $T = x_1x_2,\dots x_r$, where $x_r = x_1$, be the closed directed trail in $\Gamma(\mathcal{L})$ given by Lemma 5.1. For each arc (x_{i-1},x_i), choose an ε-regular triple $\mathcal{H}_i = \mathcal{H}[V_{x_{i-1}},V_{x_i},V_{k_i}]$ of density at least $1/2$ for some k_i (the existence of which is guaranteed by the definition of $\Gamma(\mathcal{L})$). We also choose distinct vertices $v_i \in V_{x_i}$ for $1 \leqslant i \leqslant r-1$ such that v_i is good for both triples $\mathcal{H}_i$ and $\mathcal{H}_{i+1}$ (here indices are taken mod $r-1$). If x_i is incident to the diamond arc (h,j) for some diamond $\{hjk,hjp\} \in \{\mathcal{D}_1,\dots,\mathcal{D}_s\}$, we also require that v_i be good for the two triples $\mathcal{H}[V_h,V_j,V_k]$ and $\mathcal{H}[V_h,V_j,V_p]$ as well. (Note that these two triples may or may not have been chosen as some $\mathcal{H}_b$.) Then Lemma 5.2 together with the fact that $r < 2t < \varepsilon m$ (given by Lemma 2.2(3)) ensure that such a choice of $v_1,\dots,v_{r-1}$ can always be made. We also set $v_r = v_1$. These vertices will be the end-vertices of our short paths.

Next we join v_{i-1} to v_i by a path of length three (or six) for each i, starting with $i = 2$. We apply Lemma 5.3 to the triple $\mathcal{H}_2 = \mathcal{H}[V_{x_1},V_{x_2},V_{k_2}]$, the vertices $x = v_1$ and $y = v_2$, and the set B of vertices that are not good for $\mathcal{H}_2$ together with all vertices in $\{v_1,\dots,v_r\}$. This gives a path of length three joining v_1 and v_2, whose two vertices of degree two lie in V_{x_1} and V_{x_2}.

In the general step, for $i = 3,\dots,r$, we apply Lemma 5.3 to the triple $\mathcal{H}_i = \mathcal{H}[V_{x_{i-1}},V_{x_i},V_{k_i}]$, the vertices $x = v_{i-1}$ and $y = v_i$, and the set B of vertices that are not good for $\mathcal{H}_i$ together with all vertices of previously defined short paths and in $\{v_1,\dots,v_r\}$. Since the latter set has size at most $7r < \varepsilon m$ by Lemma 2.2(3), the assumptions on the size of B in Lemma 5.3 are satisfied. This gives a path of length three joining v_{i-1} and v_i, disjoint from all previously defined paths, whose two vertices of degree two lie in $V_{x_{i-1}}$ and V_{x_i}. These short paths link to form a loose cycle of length $3(r-1)$. If the length of our target cycle $\mathcal{C}_n$ has different parity from $3(r-1)$, we replace one path of length three that does not correspond to a diamond arc by a path of length six, using Corollary 5.4, to make the parities agree. This gives a loose cycle $\mathcal{C}$ of length c, where $3(r-1) \leqslant c \leqslant 3r$ and $c \equiv n/2 \pmod 2$.

Finally, to complete the proof we replace each short path in $\mathcal{C}$ corresponding to a diamond arc by a long path. Let (x_{i-1}, x_i) be a diamond arc, for some diamond $\{x_{i-1}x_ik, x_{i-1}x_ip\}$. Then we apply Lemma 5.5 to the two ε-regular triples $\mathcal{H}[V_{x_{i-1}}, V_{x_i}, V_k]$ and $\mathcal{H}[V_{x_{i-1}}, V_{x_i}, V_p]$ of density $d \geqslant 1/2$, corresponding to the diamond. Here we are taking $x = v_{i-1}$ and $y = v_i$, and the length $\ell = \lfloor \frac{1}{s}(\frac{n}{2} - c)\rfloor$ or $\lfloor \frac{1}{s}(\frac{n}{2} - c)\rfloor - 1$, whichever is odd. We call this number the *basic length* for our long paths. Let the set B be the set of all vertices that are not good for $\mathcal{H}[V_{x_{i-1}}, V_{x_i}, V_k]$ or for $\mathcal{H}[V_{x_{i-1}}, V_{x_i}, V_p]$ together with the set of all vertices in $\mathcal{C}$. Note that the latter set still has size at most $6r < \varepsilon m$, since the s diamonds $\mathcal{D}_1, \dots, D_s$ of $\mathcal{L}$ are vertex-disjoint and so the long paths cannot interfere with each other.

To check that the condition $2dm/7 \leqslant \ell \leqslant (1 - 2\varepsilon^{1/6})2m$ from Lemma 5.5 is satisfied, observe first the following:

(a) $s \geqslant f(\varepsilon)\frac{n}{N}\frac{t(\varepsilon_0, t_0)}{4}$, where $f(\varepsilon) = (1 - 500\varepsilon^{1/6})^{-2}$, by Lemma 2.3;
(b) $t \leqslant t(\varepsilon_0, t_0)$ and $m \geqslant (1 - \varepsilon)\frac{N}{t} > (1 - 500\varepsilon^{1/6})\frac{N}{t}$ by Theorem 2.1;
(c) $c \leqslant 3r \leqslant 6t < \varepsilon m < \varepsilon n$ by Lemma 2.2(3);
(d) $N < \frac{5}{4}n(1 + \eta) < \frac{3}{2}n$ (see Step 1).

From (a) and (b), it follows that

$$\frac{1}{s}\frac{n}{2} < \left(1 - 500\varepsilon^{1/6}\right)2m < \left(1 - 2\varepsilon^{1/6}\right)2m - 2r. \tag{5.1}$$

Then, using (5.1) and (c), we obtain $\ell \leqslant \lfloor \frac{1}{s}(\frac{n}{2} - c)\rfloor < (1 - 2\varepsilon^{1/6})2m$ as required.

Of course we also have $s \leqslant t$ by Lemma 4.2, so (c) implies $\ell > \frac{1}{s}(\frac{n}{2} - c) - 2r > \frac{n}{t}(\frac{1}{2} - \varepsilon) - \varepsilon m$. But then this is at least $\frac{2}{3}\frac{N}{t}(\frac{1}{2} - \varepsilon) - \varepsilon m > \frac{2m}{7}$ by (d). Thus ℓ falls into the correct range for Lemma 5.5.

Therefore the assumptions of Lemma 5.5 are satisfied, and we can find a path joining v_i and v_{i+1} of length ℓ. For the last diamond we choose ℓ appropriately so that the resulting cycle has length exactly $n/2$, i.e. it is a copy of $\mathcal{C}_n$. This is possible since we would need to adjust the basic length by at most $2r$, and as above

$$\frac{2m}{7} < \frac{1}{s}\left(\frac{n}{2} - c\right) - 2r < \frac{1}{s}\left(\frac{n}{2} - c\right) + 2r \leqslant \left(1 - 2\varepsilon^{1/6}\right)2m.$$

This completes the proof. $\square$

Acknowledgments

This work was done as a group project of the authors, as part of a Focused Research Group meeting at the Banff International Research Station in May 2003.

The first author was partially supported by NSERC, the fourth author by NSF Grant DMS-0300529, the fifth author by KBN Grant P03A 015 23, and the seventh author by NSF Grant INT-0305793 and NSA Grant H98230-04-1-0035.

The authors would like to thank an anonymous referee for many helpful suggestions on the presentation of the paper.

References

[1] J.A. Bondy, P. Erdős, Ramsey numbers for cycles in graphs, J. Combin. Theory Ser. B 14 (1973) 46–54.

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

[3] R. Faudree, R. Schelp, All Ramsey numbers for cycles in graphs, Discrete Math. 8 (1974) 313–329.

[4] A. Figaj, T. Łuczak, The Ramsey number for a triple of long even cycles, submitted for publication.

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

[6] P. Frankl, V. Rödl, Extremal problems on set systems, Random Structures Algorithms 20 (2002) 131–164.

[7] P. Haxell, T. Łuczak, Y. Peng, V. Rödl, A. Ruciński, M. Simonovits, J. Skokan, The Ramsey number for hypergraph cycles II, manuscript.

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

[9] H.J. Prömel, A. Steger, Excluding induced subgraphs III. A general asymptotic, Random Structures Algorithms 3 (1992) 19–31.

[10] V. Rosta, On a Ramsey-type problem of J.A. Bondy and P. Erdős, I and II, J. Combin. Theory Ser. B 15 (1973) 94–104, 105–120.

[11] E. Szemerédi, Regular partitions of graphs, in: Problèmes Combinatoires et Théorie des Graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS n. 260, CNRS, Paris, 1978, pp. 399–401.