Technical Report

TR-2004-022

A Dirac-type theorem for 3-uniform hypergraphs

by

Vojtech Rodl, Andrzej Rucinski, Endre Szemeredi

Mathematics and Computer Science

EMORY UNIVERSITY

A Dirac-type theorem for 3-uniform hypergraphs

Vojtech Rödl*
Emory University,
Atlanta, GA
rodl@mathcs.emory.edu

Andrzej Ruciński†
A. Mickiewicz University
Poznań, Poland
rucinski@ amu.edu.pl

Endre Szemerédi‡
Rutgers University
New Brunswick
szemered@cs.rutgers.edu

May 13, 2004

Abstract

A hamiltonian cycle in a 3-uniform hypergraph is a cyclic ordering of the vertices in which every three consecutive vertices form an edge. In this paper we prove an approximate and asymptotic version of an analog of celebrated Dirac's theorem for graphs: For each $\gamma > 0$ there exists n_0 such that every 3-uniform hypergraph on $n \geq n_0$ vertices, in which each pair of vertices belongs to at least $(1/2 + \gamma)n$ edges, contains a hamiltonian cycle.

1 Introduction

A substantial amount of research in graph theory continues to concentrate on the existence of hamiltonian cycles. A classic theorem of Dirac states that a sufficient

*Research supported by NSF grant DMS-0300529.
†Research supported by KBN grant 2 P03A 015 23. Part of research performed at Emory University, Atlanta.
‡

condition for an n-vertex graph to be hamiltonian is that the minimum degree is at least $n/2$, and there are obvious counterexamples showing that this is best possible.

The study of hamiltonian cycles in hypergraphs was initiated in [3] where, however, a different definition than the one considered here was introduced. From now on, by a hypergraph we will always mean a 3-uniform hypergraph.

Definition 1.1 *A cycle of order k* is a hypergraph C on k vertices and k edges, whose vertices can be labeled $v_1, \ldots, v_k$ in such a way that for each $i = 1, \ldots, k-2$, $\{v_i, v_{i+1}, v_{i+2}\} \in C$ as well as $\{v_{k-1}, v_k, v_1\} \in C$ and $\{v_k, v_1, v_2\} \in C$ (there are $2k$ such labelings). By a *hamiltonian cycle* in an n-vertex hypergraph we mean a subhypergraph which is a cycle of order n. In other words, we say that a hypergraph H with $|V(H)| = n$ is *hamiltonian* if its vertices can be labeled $v_1, \ldots, v_n$ in such a way that for each $i = 1, \ldots, n$, $\{v_i, v_{i+1}, v_{i+2}\} \in H$ as well as $\{v_{n-1}, v_n, v_1\} \in H$ and $\{v_n, v_1, v_2\} \in H$.

This notion and its generalizations have a potential to be applicable in many contexts which still need to be explored. An application in the relational database theory can be found in [4]. As observed in [8], the square of a (graph) hamiltonian cycle naturally coincides with a hamiltonian cycle in a hypergraph built on top of the triangles of the graph. More precisely, given a graph G, let $Tr(G)$ be the set of triangles in G. Define a hypergraph $H^{Tr}(G) = (V(G), Tr(G))$. Then there is a one-to-one correspondence between hamiltonian cycles in $H^{Tr}(G)$ and the squares of hamiltonian cycles in G. For results about the existence of squares of hamiltonian cycles see, e.g., [9].

Example 1.1 Consider a robot walking through a tough terrain with the task to visit n designated locations and return to the base (one may view these locations as fuel providers). In order for the robot to move from one location to another, after reaching any one of them it has to be able to "see" the next one. To optimize, we do not want the robot to visit a location more than once. So far, this is just the standard traveling salesman problem, but suppose that in order to speed up the motion, or to smooth out the trajectory, we request that the robot "sees" the next two locations. Then our problem becomes that of finding a hamiltonian cycle in $H^{Tr}(G)$, where G is the graph of those pairs of n locations which can "see" each other.

Of course, a reader with strong imaginary skills can replace the robotics terminology with something else, like mountain hiking or the traveling salesman problem with an option of skipping a town.

Our next example cannot be formulated in terms of $H^{Tr}(G)$ for any graph G.

Example 1.2 Consider a patient taking 24 different pills on a daily basis, one at a time every hour. Certain combinations of three pills can be deadly if taken within 2.5 hour. Let D be the set of deadly triplets of pills. Then any safe schedule corresponds to a hamiltonian cycle in the hypergraph which is precisely the complement of D.

In [8] the authors gave a sufficient condition for a hypergraph to have a hamiltonian cycle. They proved that if every pair of vertices belongs to more than $\frac{5}{6}(n-1)+1$ edges, then the hypergraph contains a hamiltonian cycle. They also conjectured that, in fact, a much stronger result is true, namely that $\frac{5}{6}(n-1)+1$ can be replaced by $n/2$. If true this would be in close analogy with Dirac's degree condition for graphs. Some support for this conjecture stems from a construction of an edge-maximal hypergraph with each pair degree at least $\lfloor n/2 \rfloor - 1$, but not containing a hamiltonian cycle (see [8], Theorem 3). In this paper we prove an approximate and asymptotic version of this conjecture.

We say that a hypergraph H is an (n,γ)-graph if H has n vertices and every pair of vertices belongs to at least $(1/2+\gamma)n$ edges.

Theorem 1.1 *For each $\gamma > 0$ there exists n_0 such that every (n,γ)-graph with $n \geq n_0$ is hamiltonian.*

Remark 1.1 Note that an (n,γ)-graph is also an (n,γ')-graph for all $\gamma' < \gamma$. Therefore it is enough to prove Theorem 1.1 only for sufficiently small γ.

2 Preliminary lemmas

All statements in this section assume that $0 < \gamma < 1$ is sufficiently small (see Remark 1.1), n is sufficiently large and H is an (n,γ)-graph on a vertex set V.

Definition 2.1 A *k-path* is a hypergraph P on k vertices and $k-2$ edges, whose vertices can be labeled $v_1,\ldots,v_k$ in such a way that for each $i = 1,\ldots,k-2$, $\{v_i, v_{i+1}, v_{i+2}\} \in P$ (there are two such labelings). We say that P connects the (ordered) pairs $v_1v_)$ and v_kv_{k-1}, which will be referred to as *the endpairs of P*. Note that by saying that ab is an endpair of a hyperpath, we always mean that a is the first (or the last) vertex on the path, while b is the second (or penultimate). We will often call a hyperpath, simply, a path.

For two paths P and Q, let ab be an endpair of P and ba be an endpair of Q, and assume further that $V(P) \cap V(Q) = \{a,b\}$. By $P \circ Q$ we denote the path obtained

(in a unique way) as a concatenation of P and Q. This definition extends naturally to more than two paths.

Lemma 2.1 (Connecting Lemma) *For every two disjoint and ordered pairs of vertices xy and cd there is a k-path in H, $k \leq 4/\gamma$, which connects xy and cd.*

Proof: We construct sets $A_0, A_1, A_2, \ldots$ and bipartite graphs $G_1, G_2, \ldots$, where $V(G_i) = A_{i-1} \cup A_i$, as follows. Let $A_0 = \{y\}$ and $A_1 = \{z : xyz \in H\}$, and let G_1 be the star with y as the center and A_1 as the set of its leaves. Note that $|A_1| \geq (1/2 + \gamma)n$. Further, let

$$A'_2 = \{w : \exists z \in A_1 \text{ such that } yzw \in H\} \text{ and } G'_2 = \{zw : z \in A_1, w \in A'_2, yzw \in H\}.$$

Observe that for every edge $zw \in G'_2$ with $w \neq x$ the vertices $xyzw$ form a 4-path in H. Also, for each $z \in A_1$, we have $deg_{G'_2}(z) \geq (1/2 + \gamma)n$.

Let $A^0_2 = \{w \in A'_2 : deg_{G'_2}(w) < \sqrt{n}\}$, $A_2 = A'_2 \setminus A^0_2$ and $G_2 = G'_2[A_1 \cup A_2]$. Note that

$$|A_1|(1/2 + \gamma)n \leq |G'_2| \leq n^{3/2} + |A_2||A_1|$$

which implies that $|A_2| > n/2$.

Having constructed $A_0, A_1, \ldots, A_j$ and $G_1, \ldots G_j$, $j \geq 2$, consider, for every $w \in A_j$, an auxiliary bipartite graph B^j_w between the neighbors of w in G_j and all vertices in V, where a pair zu, $z \in N_{G_j}(w)$, $u \in V$, is an edge of B^j_w if $zwu \in H$. Define

$$A'_{j+1} = \{u : \exists w \in A_j \text{ such that } deg_{B^j_w}(u) \geq n^{1/4}\}$$

and

$$G'_{j+1} = \{wu : \text{ such that } w \in A_j \text{ and } deg_{B^j_w}(u) \geq n^{1/4}\}.$$

Finally, let

$$A^0_{j+1} = \{w \in A'_{j+1} : deg_{G'_{j+1}}(w) < \sqrt{n}\},$$

$$A_{j+1} = A'_{j+1} \setminus A^0_{j+1}$$

and

$$G_{j+1} = G'_{j+1}[A_j \cup A_{j+1}].$$

Notice that some sets A_j may intersect or even coincide (in fact, at some point the construction starts to repeat itself forever). Nevertheless, for the sake of our construction, we treat them as disjoint, cloning the vertices as much as necessary. Let us call the entire structure, consisting of the sets $A_0, A_1, A_2 \ldots$ and the graphs $G_1, G_2 \ldots$, an *xy-cascade*.

We had to alter our construction for $j \geq 3$ and require $deg_{B_w^j}(u) \geq n^{1/4}$ and not just $deg_{B_w^j}(u) \geq 1$, in order to be able to return from any edge of G_j back to xy by a legitimate hyperpath, on which all vertices must be distinct. With the above definition, in any xy-cascade, there is always a hyperpath from any edge of G_j going backward all the way down to xy as long as $j < n^{1/4}$. Indeed, when choosing a next (backward) vertex, we can avoid any given set of vertices of size less than $n^{1/4}$. In particular, we can avoid all vertices which are already on the path, as well as x and y. (In fact, we will need this property only to avoid sets of size $O(1)$.)

A vertex $u \in A_j$ is called *heavy* if $deg_{G_j}(u) \geq (1/2 + \gamma/2)n$.

Claim 2.1 *There exists an index $j \leq j_0 = \lceil 1/\gamma \rceil + 1$ such that A_j contains at least one heavy vertex.*

Proof: We will first show that for $j \geq 2$ every vertex $w \in A_j$ has in G'_{j+1} degree at least $(1/2 + \gamma)n - n^{3/4}$. Indeed, let s be the number of vertices $u \in V$ with $deg_{B_w^j}(u) < n^{1/4}$. Then

$$sn^{1/4} + (n-s)|N_{G_j}(w)| \geq |B_w^j| \geq |N_{G_j}(w)|(1/2+\gamma)n$$

which yields, using $|N_{G_j}(w)| = deg_{G_j}(w) \geq \sqrt{n}$ and $s \leq n$,

$$n - s \geq (1/2+\gamma)n - \frac{sn^{1/4}}{|N_{G_j}(w)|} \geq (1/2+\gamma)n - n^{3/4}.$$

Note also that the total number of edges of G'_{j+1} incident to the vertices of A^0_{j+1} is smaller than $n^{3/2}$.

Now suppose that the claim is not true. Then, using the above estimates, for each $j = 2, \ldots, j_0$

$$|A_{j-1}|(1/2+\gamma)n - n^{7/4} - n^{3/2} \leq |G_j| \leq |A_j|(1/2+\gamma/2)n$$

and, consequently, since $|A_1| \geq (1/2+\gamma)n$, we have

$$|A_{j_0}| > \frac{1+2\gamma}{1+\gamma}|A_{j_0-1}| - O(n^{3/4}) > \left(\frac{1+2\gamma}{1+\gamma}\right)^{\lceil 1/\gamma \rceil} \frac{n}{2} > n,$$

a contradiction. (For the last inequality we used the fact that $(1-x)e^x \leq 1$ with $x = \gamma/(1+2\gamma)$ and assumed that $(1+2\gamma)\ln 2 \leq 2$.) ■

Given two disjoint, ordered pairs of vertices xy and cd, consider the xy-cascade $(A_j^{(1)}, G_j^{(1)})$ and the cd-cascade $(A_j^{(2)}, G_j^{(2)})$. For $i = 1, 2$, let $b^{(i)} \in A^{(i)}_{j^{(i)}}$ be a heavy

vertex, where $j^{(i)} \leq j_0$. Then, there exists $a^{(i)} \in A^{(i)}_{j^{(i)}-1}$ such that $a^{(i)}b^{(i)} \in G^{(1)}_{j^{(i)}}$ and $a^{(i)}b^{(i)}b^{(3-i)} \in H$. Moreover, by the definition of the xy-cascade, there is a $(j^{(1)}+2)$-path $P^{(1)}$ connecting xy and $b^{(1)}a^{(1)}$ and, by the definition of the cd-cascade, there is a $(j^{(2)}+2)$-path $P^{(2)}$, disjoint from $P^{(1)}$, connecting cd and $b^{(2)}a^{(2)}$. Hence, for some

$$k = (j^{(1)}+2) + (j^{(2)}+2) \leq 2(j_0+2) \leq 4/\gamma,$$

there is a k-path in H which connects xy and cd (for the last inequality we have assumed that $\gamma \leq 1/4$).

■

Lemma 2.2 (Absorbing Lemma) *There is an l-path A in H with $l = |V(A)| \leq 20\gamma^2 n$, such that for every subset $U \subset V \setminus V(A)$ of size at most $\gamma^5 n$ there is a path A_U in H with $V(A_U) = V(A) \cup U$ and such that A_U has the same endpairs as A.*

In other words, this lemma asserts that there is *one* not too long path such that *every* not too large subset can be "absorbed" into this path by creating a longer path with the same endpairs. Consequently, if this path happens to be a segment of a cycle C of order at least $(1-\gamma^5)n$ then, setting $U = V \setminus V(C)$, the path A_U is hamiltonian. We will use this fact at the end of our proof of Theorem 1.1.

Proof: Given a vertex v we say that a 4-tuple of vertices x, y, z, w *absorbs* v if $xyz, yzw, xyv, yvz, vzw \in H$. A 4-tuple is called *absorbing* if it absorbs a vertex. This terminology reflects the fact that the path $xyzw$ can be extended by inserting (or absorbing) vertex v to create the path $xyvzw$. Note that both paths have the same set of endpairs.

Claim 2.2 *For every $v \in V$ there are at least $2\gamma^2 n^4$ 4-tuples absorbing v.*

Proof: Because H is an (n,γ)-graph, there are at least $(n-1)(1/2+\gamma)n$ ordered pairs yz such that $vyz \in H$. For each such pair there are at least $2\gamma n$ common neighbors x of vy and yz, and at least $2\gamma n - 1$ common neighbors w of vz and yz, yielding together at least

$$(n-1)(1/2+\gamma)n2\gamma n(2\gamma n - 1) > 2\gamma^2 n^4$$

4-tuples absorbing v. ■

For each $v \in V$, let $\mathcal{A}_v$ be the family of all 4-tuples absorbing v. The next claim is obtained by the probabilistic method.

Claim 2.3 *There exists a family $\mathcal{F}$ of at most $2\gamma^3 n$ disjoint, absorbing 4-tuples of vertices of H such that for every $v \in V$, $|\mathcal{A}_v \cap \mathcal{F}| > \gamma^5 n$.*

Proof: We first select a family $\mathcal{F}'$ of 4-tuples at random by including each of $n(n-1)(n-2)(n-3) \sim n^4$ of them independently with probability $\gamma^3 n^{-3}$ (some of the selected 4-tuples may not be absorbing at all). By Chernoff's inequality (see, e.g., [7]), with probability $1 - o(1)$, as $n \to \infty$,

- $|\mathcal{F}'| < 2\gamma^3 n$, and
- for each $v \in V$, $|\mathcal{A}_v \cap \mathcal{F}'| > \frac{3}{2}\gamma^5 n$.

Moreover, the expected number of intersecting pairs of 4-tuples in $\mathcal{F}'$ is at most

$$n^4 \times 4 \times 4 \times n^3 \times (\gamma^3 n^{-3})^2 = 16\gamma^6 n,$$

and so, by Markov's inequality, with probability at least $1/17$

- there are at most $17\gamma^6 n$ pairs of intersecting 4-tuples in $\mathcal{F}'$.

Thus, with positive probability, a random family $\mathcal{F}'$ possesses all three properties marked by the bullets above, and hence there exists at least one such family which, with a little abuse of notation, we also denote by $\mathcal{F}'$. After deleting from $\mathcal{F}'$ all 4-tuples intersecting other 4-tuples in $\mathcal{F}'$, as well as those which do not absorb any vertex, we obtain a subfamily $\mathcal{F}$ of $\mathcal{F}'$ consisting of disjoint and absorbing 4-tuples and such that for each $v \in V$,

$$|\mathcal{A}_v \cap \mathcal{F}| > \frac{3}{2}\gamma^5 n - 34\gamma^6 n > \gamma^5 n.$$

■

Set $f = |\mathcal{F}|$ and let $F_1, \ldots, F_f$ be the elements of $\mathcal{F}$. For each $i = 1, \ldots, f$, F_i is absorbing and thus spans a 4-path in H. We will further denote these paths also by F_i and set $F = \bigcup_{i=1}^{f} F_i$.

Our next task is to connect all these 4-paths into one, not too long path A. To this end, we will repeatedly apply Lemma 2.1 and, for each $i = 1, \ldots, f-1$, connect the endpairs of F_i and F_{i+1} by a short path. Recall that the operation $P \circ Q$ has been defined at the beginning of this section.

Claim 2.4 *There exists a path A in H of the form*

$$A = F_1 \circ C_1 \cdots \circ F_{f-1} \circ C_{f-1} \circ F_f$$

where the paths $C_1, \ldots, C_{f-1}$ have each at most $8/\gamma$ vertices.

Proof: We will prove by induction on i that for each $i = 1, \dots f$, there exists a path A_i in H of the form $A_1 = F_1$ and, for $i \geq 2$,

$$A_i = F_1 \circ C_1 \cdots \circ F_{i-1} \circ C_{i-1} \circ F_i,$$

where the paths $C_1, \dots, C_{i-1}$ have each at most $8/\gamma$ vertices. Then $A = A_f$.

There is nothing to prove for $i = 1$. Assume the statement is true for some $1 \leq i \leq f-1$. Let ab be an endpair of A_i and let cd be an endpair of F_{i+1}. Denote by H_i the subhypergraph induced in H by the set of vertices $V_i = (V \setminus V(F \cup A_i)) \cup \{a, b, c, d\}$. Since

$$|V(F \cup A_i)| < |\mathcal{F}|(4 + 8/\gamma) < 10f/\gamma < 20\gamma^2 n,$$

H_i is a $(|V_i|, \gamma/2)$-graph, where $0 < n - |V_i| < 20\gamma^2 n$. By Lemma 2.1 applied to H_i and the pairs ba and dc, there is a path $C_i \subset H_i$ of length at most $4/(\gamma/2) = 8/\gamma$, connecting these pairs. Note that $V(C_i) \setminus \{a, b, c, d\}$ is disjoint from $V(F \cup A_i)$, and thus,

$$A_{i+1} = A_i \circ C_i \circ F_{i+1}$$

is the desired path. ■

Claim 2.4 states that we may connect all 4-paths in $\mathcal{F}$ into one path A of length at most $f(4 + 8/\gamma) < 20\gamma^2 n$. It remains to show that A has the absorbing property. Let $U \subset V \setminus V(A)$, $|U| \leq \gamma^5 n$. Because for every $v \in U$ we have $|\mathcal{A}_v \cap \mathcal{F}| > \gamma^5 n$, that is, there are at least $\gamma^5 n$ disjoint, v-absorbing 4-tuples in A, we can insert all vertices of U into A one by one, each time using a fresh absorbing 4-tuple. ■

Given $U \subseteq V$ and $x, y \in V$, let

$$deg_H(xy, U) = |\{z \in U : xyz \in H\}|,$$

and, in particular, $deg_H(xy) = deg_H(xy, V)$. Note that in an (n, γ)-graph H we have $deg_H(xy) \geq (1/2 + \gamma)n$ for all pairs of vertices $x, y \in V$.

Lemma 2.3 (Reservoir Lemma) *For every subset $W \subset V$, $|W| < \gamma n/4$ there exists a subset $R \subset V \setminus W$ (a reservoir) such that $|R| = \lceil \gamma^5 n/2 \rceil$ and for every pair of vertices $x, y \in V$*

$$deg_H(xy, R) \geq (1/2 + \gamma/2)(|R| + 4).$$

In particular, for every $S \subset V \setminus R$, $|S| = 4$, $H[R \cup S]$ is an $(|R| + 4, \gamma/2)$-graph.

Proof: Set $r = \gamma^5 n/2$ (to avoid irrelevant complications, we assume that r is an integer). We choose R randomly out of all $\binom{n-|W|}{r}$ possibilities and apply the probabilistic method. The random variable $X = X_{xy}$ has the hypergeometric distribution

with expectation $\mathbb{E}X$ satisfying

$$r \geq \mathbb{E}X = \frac{deg_H(xy)}{n-|W|} r \geq \left(\frac{1}{2} + \frac{3}{4}\gamma\right) r.$$

Hence, by Chernoff's bound ([7], (2.6) on page 26 and Thm. 2.10 on page 29)

$$\mathbb{P}\left(X < \left(\frac{1}{2} + \frac{1}{2}\gamma\right)(r+4)\right) \leq \mathbb{P}\left(X \leq \mathbb{E}X - \frac{1}{4}\gamma r + 2\right) \leq \exp\left\{-\frac{\gamma^2 r}{33}\right\},$$

and consequently,

$$\mathbb{P}\left(\exists xy : X_{xy} < \left(\frac{1}{2} + \frac{1}{2}\gamma\right)(r+4)\right) \leq \binom{n}{2} \exp\left\{-\frac{\gamma^7 n}{66}\right\} = o(1).$$

■

A matching in a hypergraph is a set (subhypergraph) of disjoint edges. Our last lemma in this section guarantees an almost perfect matching in a special class of hypergraphs. This lemma will be later applied to the so called cluster graph resulting from the Regularity Lemma (Lemma 4.1).

Given a hypergraph K, let G_K be the graph of all pairs xy of vertices which belong to less than $|V(K)|/2$ edges of K, that is, for which $deg_K(xy) < |V(K)|/2$.

Lemma 2.4 *If a hypergraph K on $t \geq 24$ vertices satisfies the inequality*

$$\Delta(G_K) \leq \frac{t}{12},$$

then K contains a matching M covering all but at most $\max(2, \Delta(G_K)) + 1$ *vertices.*

Proof: Let $M = e_1, \ldots e_g$, $g < t/3$, be a largest matching in K and suppose that a set U of more than $u = \max(2, \Delta(G_K)) + 1$ vertices remains uncovered. Note that

$$t - 3g = |U| \geq u + 1 \geq 4. \tag{1}$$

Call a pair of vertices of K big if it belongs to at least $t/2$ edges of K, that is, it is an edge of the complement G_K^c of the graph G_K. Let $v_1, v_2, v_3, v_4 \in U$ be four vertices of U, where v_3v_4 forms a big pair. (The existence of a big pair in U is guaranteed, because for any vertex $v \in U$ its degree in G_K is at most $u-1$, while there are at least u vertices in U besides v.)

The set $N(v_3, v_4)$ of neighbors of the pair v_3v_4 in K intersects at least $t/6$ edges of M (since it is disjoint from U). Then, because $2(u-1) < t/6$, there exists an edge

of the matching M, say $e_g = \{x_1, x_2, x_3\}$, such that $v_3v_4x_3 \in K$ and both, v_1x_1 and v_2x_2 are big pairs in K.

Note that if for some $s = 1, 2$, there is a vertex $v \in U \setminus \{v_3, v_4\}$ such that $vv_sx_s \in K$, then the edges $e_1, \ldots, e_{g-1}$, vv_sx_s and $v_3v_4x_3$ would form a matching of K larger than M – a contradiction. Hence, for each $s = 1, 2$ we have $N(v_sx_s) \cap U \subseteq \{v_3, v_4\}$. Similarly, $v_1x_1x_2 \notin K$ and $v_2x_2x_1 \notin K$. Hence, for each $i = 1, 2$, all but at most three neighbors of v_sx_s belong to $M - e_g$.

Let a_q^s, $s = 1, 2$, $q = 1, \ldots g - 1$, be the number of vertices of e_q which together with v_sx_s form an edge of K. Then

$$\sum_{q=1}^{g-1}(a_q^1 + a_q^2) \geq 2\left(\frac{t}{2} - 3\right) = t - 6.$$

By averaging, there must be an index q, $1 \leq q \leq g - 1$, such that

$$a_q^1 + a_q^2 \geq \frac{t-6}{g-1} > 3,$$

where the last inequality follows from (1). This means, however, that there are $y, z \in e_q$, $y \neq z$, such that $yv_1x_1 \in K$ and also $zv_2x_2 \in K$. But then

$$e_1, \ldots, e_{q-1}, e_{q+1}, \ldots, e_{g-1}, v_3v_4x_3, yv_1x_1, zv_2x_2$$

is a larger matching than M – a contradiction. ■

3 Proof of Theorem 1.1

We first outline the forthcoming proof. Let H be an (n, γ)-graph.

- By Lemma 2.2 fix an absorbing l-path A, $l = |V(A)| \leq 20\gamma^2 n$.
- By Lemma 2.3 with $W = V(A)$, fix a reservoir set $R \subset V \setminus V(A)$, $|R| = \lceil \frac{1}{2}\gamma^5 n \rceil$.
- Set $H_1 = H[V \setminus (V(A) \cup R)]$ and cover all but at most $\frac{1}{2}\gamma^5 n$ vertices of H_1 by disjoint paths $P_1, \ldots, P_p$, where $p \leq \gamma^8 n$. Denote the set of uncovered vertices by T.
- By $p+1$ applications of Lemma 2.1 and by the property of R, connect all paths $P_1, \ldots, P_p$, as well as A, into one cycle C in H, leaving only a leftover subset R' of R and the trash set T outside C. Note that $|R' \cup T| \leq \gamma^5 n$.

- Using the absorbing property of A insert $R' \cup T$ into C, obtaining a hamiltonian cycle in H.

It remains to explain the third task on the above list: how to cover almost all vertices of H_1 by disjoint paths $P_1, \ldots, P_p$. This will be taken care of by Lemma 3.1 stated below. However, as this lemma relies heavily on a regularity lemma for hypergraphs and some related results, we devote two separate sections, Section 4 and Section 5, to its proof.

Lemma 3.1 (Path-Cover Lemma) *For every $\gamma > 0$ there exists n_0 such that every (n, γ)-graph, $n > n_0$, contains a family of at most $\gamma^8 n$ vertex-disjoint paths, covering all but at most $\gamma^5 n/2$ vertices.*

Proof of Theorem 1.1: Let us assume that $\gamma < \frac{1}{400}$ (see Remark 1.1) and let A be an absorbing l-path in H, $l = |V(A)| \leq 20\gamma^2 n$, whose existence is guaranteed by Lemma 2.2. By Lemma 2.3 applied to H with $W = V(A)$, there exists a reservoir set $R \subset V \setminus V(A)$ of size $|R| = \frac{1}{2}\gamma^5 n$ (for simplicity we are assuming that this is an integer) with the property described in that lemma. Set $H_1 = H[V \setminus (V(A) \cup R)]$ and note that H_1 is an (n_1, γ_1)-graph, where

$$n - 20\gamma^2 n - \frac{1}{2}\gamma^5 n \leq n_1 \leq n$$

and

$$0 < \gamma - 20\gamma^2 - \frac{1}{2}\gamma^5 < \gamma_1 < \gamma.$$

We apply Lemma 3.1 to H_1, obtaining a family of $p \leq \gamma_1^8 n_1 \leq \gamma^8 n$ vertex-disjoint paths $P_1, \ldots, P_p$, and the set $T \subset V(H_1)$ of vertices not covered by these paths of size $|T| \leq \gamma_1^5 n_1/2 \leq \gamma^5 n/2$.

To connect all these paths as well as the path A into one cycle C, we successively apply Lemma 2.1 to ever shrinking subhypergraphs of the form $H[R_i \cup S]$, where a subset $R_i \subseteq R$ will be defined below in (2), while S consists of all four vertices from the endpairs of the two paths to be connected at the current stage. Thus, new vertices of the connecting paths will be entirely contained in the set R. The next claim, very similar to Claim 2.4, describes the procedure of connecting together all path $P_1, \ldots, P_p$ into one path L.

Claim 3.1 *There exists a path L in H of the form*

$$L = P_1 \circ C_1 \cdots \circ P_p \circ C_p \circ A,$$

where the paths $C_1, \ldots, C_p$ have each at most $20/\gamma$ vertices and are such that

$$V(C_1 \cup \cdots \cup C_p) \setminus V(P_1 \cup \cdots \cup P_p \cup A) \subset R.$$

Proof: Set $P_{p+1} = A$ to unify notation. We will prove by induction on i that for every $i = 1, \ldots, p+1$, there exists a path L_i in H of the form

$$L_i = P_1 \circ C_1 \cdots \circ P_{i-1} \circ C_{i-1} \circ P_i,$$

where the paths $C_1, \ldots, C_{i-1}$ have each at most $20/\gamma$ vertices and are such that

$$V(C_1 \cup \cdots \cup C_{i-1}) \setminus V(P_1 \cup \cdots \cup P_i) \subset R.$$

Then $L = L_{p+1}$. There is nothing to prove for $i = 1$. Assume that the statement is true for some $1 \leq i \leq p$. Let ab be an endpair of L_i and cd be an endpair of P_{i+1}. Since

$$|V(C_1 \cup \cdots \cup C_{i-1})| \leq (i-1)(20/\gamma) < 20p/\gamma < 20\gamma^7 n,$$

the remaining subset

$$R_i = R \setminus V(C_1 \cup \cdots \cup C_{i-1}) \tag{2}$$

of R still maintains the property described in Lemma 2.3 but with $\frac{2}{5}\gamma$ instead of $\gamma/2$. Indeed, since $\gamma < \frac{1}{400}$, for every pair $x, y \in V$

$$deg_H(xy, R_i) \geq \left(\frac{1}{2} + \frac{1}{2}\gamma\right)(|R| + 4) - 20\gamma^7 n \leq \left(\frac{1}{2} + \frac{2}{5}\gamma\right)(|R_i| + 4).$$

Hence, the subhypergraph $H[R_i \cup \{a, b, c, d\}]$ is an $(|R_i| + 4, \frac{2}{5}\gamma)$-graph. By Lemma 2.1 applied to $H[R_i \cup \{a, b, c, d\}]$ and the pairs ba and dc there is a path C_i in $H[R_i \cup \{a, b, c, d\}]$ of length at most $20/\gamma$ connecting these pairs. Note that $V(C_i) \setminus \{a, b, c, d\}$ is disjoint from $V(C_1 \cup \cdots \cup C_{i-1}) \cup V(P_1 \cup \cdots \cup P_{p+1})$ and thus,

$$L_{i+1} = L_i \circ C_i \circ P_{i+1}$$

is the desired path. ■

To obtain the cycle C, let ab and cd be the two endpairs of $L = L_{p+1}$ (we follow the notation from the proof of Claim 3.1. Again, applying Lemma 2.1 to $H[R_{p+1} \cup \{a, b, c, d\}]$ and the pairs ba and dc we obtain a path C_{p+1} of length at most $20/\gamma$ connecting these pairs (and thus forming the desired cycle C) and such that $V(C_{p+1}) \setminus \{a, b, c, d\} \subseteq R_{p+1}$. Set $R_{p+2} = R_{p+1} - V(C_{p+1})$. There are at most $\gamma^5 n$ vertices left outside C. Indeed, we have

$$n - |V(C)| = |T| + |R_{p+2}| < |T| + |R| < \frac{1}{2}\gamma^5 n + \frac{1}{2}\gamma^5 n = \gamma^5 n.$$

Finally, we absorb these remaining vertices into the path A, which is now part of the cycle C. Set $U = V - V(C)$. By Lemma 2.2 there is a path A_U with the same endpairs as A and such that $V(A_U) = V(A) \cup U$. Then $A_U \cup C$ is a hamiltonian cycle in H. ■

4 Regularity of Hypergraphs

In the previous section we stated Lemma 3.1, so crucial for the proof of our main result. Here we make thorough preparations towards its proof which is contained in Section 5. Our proof will be based on a regularity lemma for hypergraphs from [5].

4.1 Regularity of graphs

We say that a bipartite graph G with bipartition $V(G) = X \cup Y$ is (d, ε)-*regular* if for all $A \subseteq X$ and $B \subseteq Y$ with $|A| > \varepsilon|X|$ and $|B| > \varepsilon|Y|$, we have

$$|d_G(A, B) - d| < \varepsilon,$$

where

$$d_G(A, B) = \frac{e_G(A, B)}{|A||B|}$$

is the *density* of the pair (A, B) and $e_G(A, B)$ is the number of edges in G with one endpoint in A and the other in B. We will write d_G or $d(G)$ for $d_G(X, Y)$. We say that G is ε-*regular* if it is (d, ε)-regular for some d.

Note that the (bipartite) complement of a (d, ε)-regular graph is itself $(1 - d, \varepsilon)$-regular. Also, if G_i is (d_i, ε_i)-regular, $i = 1, 2$, and G_1 and G_2 have the same vertex set (and the same bipartition), but are edge-disjoint, then their union $G_1 \cup G_2$ is $(d_1 + d_2, \varepsilon_1 + \varepsilon_2)$-regular.

A triple $\mathcal{T} = (P^{12}, P^{13}, P^{23})$ of bipartite graphs with vertex sets $V_1 \cup V_2$, $V_1 \cup V_3$ and $V_2 \cup V_3$ will be referred to as a *triad*. Let $tr(\mathcal{T})$ stand for the number of triangles in $P = P^{12} \cup P^{13} \cup P^{23}$. It is easy to estimate the number of triangles in a triad consisting of ε-regular graphs (see, e.g., Fact A in [5]). Here we will need a slight extension of that result, assuming that only two out of the three bipartite graphs are ε-regular. We include a simple proof for completeness.

Fact 4.1 *Let $\mathcal{T} = (P^{12}, P^{13}, P^{23})$ be a triad, where for some $0 \leq d_{13}, d_{23}, \varepsilon \leq 1$, the graphs P^{13} and P^{23} are, respectively, (d_{13}, ε)-regular and (d_{23}, ε)-regular. Then*

$$d(P^{12})d_{13}d_{23} - 4\varepsilon < \frac{tr(\mathcal{T})}{|V_1||V_2||V_3|} < d(P^{12})d_{13}d_{23} + 6\varepsilon.$$

In particular, if all three graphs are (d, ε)-regular, then

$$d^3 - 5\varepsilon < \frac{tr(\mathcal{T})}{|V_1||V_2||V_3|} < d^3 + 7\varepsilon.$$

Proof: Assume for simplicity that $|V_i| = n$ for each $i = 1, 2, 3$. For $v \in V_1$, let $N(v)$ be the set of neighbors of v in P^{13}, and for $v \in V_1$ and $u \in V_2$, let $N(v, u)$ be the subset of $N(v)$ consisting of the neighbors of u in P^{23} (which are thus also neighbors of v in P^{13}).

Let U^+ and U^ε be the sets of those vertices $v \in V_1$ for which, respectively, $|N(v)| > (d_{13} + \varepsilon)n$ and $|N(v)| \leq \varepsilon n$. By the ε-regularity of P^{13} we have $|U^+| \leq \varepsilon n$. If $v \in U^\varepsilon$, then, clearly, $|N(v, u)| \leq \varepsilon n$ for every $u \in V_2$. For each $v \in U = V_1 \setminus (U^+ \cup U^\varepsilon)$, let U_v^+ be the set of those vertices $u \in V_2$ for which $|N(v, u)| > (d_{23} + \varepsilon)|N(v)|$. Then, by the ε-regularity of P^{23} we have $|U_v^+| \leq \varepsilon n$.

Let us express P^{12} as a union of four edge-disjoint subgraphs, $P^{12} = F_1 \cup F_2 \cup F_3 \cup F_4$, where F_1 consists of all edge vu with $v \in U^+$, F_2 – with $v \in U^\varepsilon$, F_3 – with $v \in U$ and $u \in U_v^+$, and, finally, F_4 consists of all edges vu with $v \in U$ and $u \notin U_v^+$.

By the above estimates, $\sum_{vu \in F_i} |N(v, u)| \leq \varepsilon n^3$, $i = 1, 2, 3$, while

$$\sum_{vu \in F_4} |N(v, u)| \leq |P^{12}|(d_{13} + \varepsilon)(d_{23} + \varepsilon)n.$$

Altogether,

$$tr(\mathcal{T}) = \sum_{vu \in P^{12}} |N(v, u)| \leq 3\varepsilon n^3 + d(P^{12})(d_{13} + \varepsilon)(d_{23} + \varepsilon)n^3 < (d(P^{12})d_{13}d_{23} + 6\varepsilon)n^3.$$

For the lower bound, we may assume that $\min(d(P^{12}), d_{13}, d_{23}) > 4\varepsilon$ and consider the set U^- of all $v \in V_1$ for which $|N(v)| < (d_{13} - \varepsilon)n$, and for each $v \in V_1 \setminus U^-$, the set U_v^- of those vertices $u \in V_2$ for which $|N(v, u)| < (d_{23} - \varepsilon)|N(v)|$. We have $|U^-| \leq \varepsilon n$ and, for all $v \in V_1 - U^-$, $|U_v^-| \leq \varepsilon n$, because $|N(v)| \geq (d_{13} - \varepsilon)n > \varepsilon n$. Thus, for all but at most $2\varepsilon n^2$ pairs v, u, we have $|N(v, u)| > (d_{13} - \varepsilon)(d_{23} - \varepsilon)n$. Consequently, similarly as for the upper bound,

$$tr(\mathcal{T}) \geq (d(P^{12}) - 2\varepsilon)(d_{13} - \varepsilon)(d_{23} - \varepsilon)n^3 > (d(P^{12})d_{13}d_{23} - 4\varepsilon)n^3.$$

■

4.2 Regularity of hypergraphs

For a triad $\mathcal{T} = (P^{12}, P^{13}, P^{23})$ with $tr(\mathcal{T}) > 0$ and a 3-uniform, 3-partite hypergraph H with vertex set $V(H) = V_1 \cup V_2 \cup V_3$ we define the *density of H over $\mathcal{T}$* as

$$d_H(\mathcal{T}) = \frac{|H \cap Tr(\mathcal{T})|}{tr(\mathcal{T})},$$

where $Tr(\mathcal{T})$ is the set of triplets formed by the vertex sets of all triangles in P. (If $tr(\mathcal{T}) = |Tr(\mathcal{T})| = 0$ then we set $d_H(\mathcal{T}) = 0$.)

Definition 4.1 Let $\delta > 0$. We will say that a hypergraph H is *δ-regular* with respect to the triad $\mathcal{T} = (P^{12}, P^{13}, P^{23})$ if for every triad $\mathcal{S} = (Q^{ij}, Q^{ik}, Q^{ik})$ such that $Q^{ij} \subseteq P^{ij}$, $1 \leq i < j \leq 3$, and $tr(\mathcal{S}) > \delta tr(\mathcal{T})$, we have $|d_H(\mathcal{S}) - d_H(\mathcal{T})| < \delta$. A triad with respect to which a hypergraph is not δ-regular will be called *δ-irregular*.

The hereditary nature of regularity is captured by the following, simple fact.

Fact 4.2 *Let $\mathcal{T} = (P^{12}, P^{13}, P^{23})$ be a triad with vertex sets V_1, V_2, V_3, all of equal size n, and let H be a hypergraph, $V(H) = V_1 \cup V_2 \cup V_3$. Furthermore, for $0 < \eta < 1$, let $U_i \subseteq V_i$, $|U_i| > \eta n$, $i = 1, 2, 3$, and $Q^{ij} = P^{ij}[U_1, U_2]$, $1 \leq i < j \leq 3$.*

(a) If P^{12} is (d, ε)-regular with $d > \varepsilon$ and $\varepsilon < \eta < 1$ then Q^{12} is ε/η-regular with density $d - \varepsilon < d_{Q^{12}}(U_1, U_2) < d + \varepsilon$.

(b) If all graphs P^{ij}, $1 \leq i < j \leq 3$, are (d, ε)-regular , $d^3 > 11\varepsilon/\eta$, and H is δ-regular with respect to $\mathcal{T}$, where $\delta < \eta^3/3$, then H is $3\delta/\eta^3$- regular with respect to the triad $\mathcal{S} = (Q^{12}, Q^{13}, Q^{23})$ and has density $d_H(\mathcal{S})$ satisfying $|d_H(\mathcal{S}) - d_H(\mathcal{T})| < \delta$.

Proof: Part (a) is obvious. For part (b), note that by Fact 4.1, $tr(\mathcal{T}) < n^3(d^3 + 7\varepsilon)$ and, similarly,

$$tr(\mathcal{S}) > (\eta n)^3(d^3 - 5\varepsilon/\eta) > \frac{1}{3}\eta^3(d^3 + 7\varepsilon)n^3 > \frac{1}{3}\eta^3 tr(\mathcal{T}),$$

where for the middle inequality we used the assumption that $d^3 > 11\varepsilon/\eta$. Thus, if $\mathcal{R}$ is a subtriad of $\mathcal{S}$ with $tr(\mathcal{R}) \geq (3\delta/\eta^3)tr(\mathcal{S})$ then, since $\eta^3/3 > \delta$, we have $tr(\mathcal{R}) > \delta tr(\mathcal{T})$ and by the δ-regularity of H with respect to $\mathcal{T}$,

$$|d_H(\mathcal{R}) - d_H(\mathcal{T})| < \delta.$$

Since the above applies in particular to $\mathcal{R} = \mathcal{S}$, we conclude that $|d_H(\mathcal{R}) - d_H(\mathcal{S})| < 2\delta < 3\delta/\eta^3$, which proves that H is $3\delta/\eta^3$-regular with respect to $\mathcal{S}$. ■

We now state the regularity lemma for 3-uniform hypergraphs in a simplified form, suitable for our needs. Set $K(U,W)$ for the complete bipartite graph with vertex sets U and W.

Lemma 4.1 (Regularity Lemma for Hypergraphs [5]) *For every $\delta > 0$, an integer t_0 and for all sequences $0 < \varepsilon(l) \le 1/(20l^3)$, there exist constants T_0, L_0 and N_0 such that every 3-uniform hypergraph H with at least N_0 vertices admits a partition Π of $\binom{V(H)}{2}$ consisting of an auxiliary vertex set partition $V(H) = V_0 \cup V_1 \cup \cdots \cup V_t$, where $t_0 \le t < T_0$, $|V_0| < t$ and $|V_1| = |V_2| = \cdots = |V_t|$, and, for each pair i,j, $1 \le i < j \le t$, of a partition $K(V_i,V_j) = \bigcup_{a=1}^{l} P_a^{ij}$, where $1 \le l < L_0$, satisfying the following conditions:*

(i) all graphs P_a^{ij} are $(1/l, \varepsilon(l))$- regular,

(ii) for all but at most $\delta l^3 t^3$ triads $\mathcal{T}_{abc}^{hij} = (P_a^{hi}, P_b^{hj}, P_c^{ij})$, the hypergraph H is δ-regular with respect to $\mathcal{T}_{abc}^{hij}$.

There are three essential differences between the original version from [5] and the one stated above. Firstly, we consider only the case $r(l,t) \equiv 1$. Secondly, instead of counting triangles contained in irregular triads, in (ii) we count the irregular triads themselves, which is essentially equivalent (cf. Proposition 4.6 in [11] and the outline below).

Thirdly, and most importantly, we have no exceptional graphs P_a^{ij} whatsoever. Moreover, the number of graphs P_a^{ij} between each pair (V_i,V_j) of clusters, which varied in the original setting, is now, conveniently, precisely l. Below we outline how our version follows from the original statement, Theorem 3.5 in [5] (see [14] for more details).

Outline of how Lemma 4.1 follows from Theorem 3.5 in [5]: Given $\delta > 0$, t_0 and $0 < \varepsilon(l) \le 1/(20l^3)$, apply Theorem 3.5 in [5] to H with $\delta' = \delta/2$, $\varepsilon_1 = \delta^4/8$, t_0, $l_0 = 1$, $r(l,t) = 1$ and $\varepsilon_2(l) = \delta\varepsilon(l)/(6l)$ obtaining a partition Π' as in Theorem 3.5 of [5], consisting of a vertex partition $V(H) = V_0 \cup V_1 \cup \cdots \cup V_t = m$ and, for each pair i,j, $1 \le i < j \le t$, of a partition $K(V_i,V_j) = \bigcup_{a=0}^{l_{ij}} \bar{P}_a^{ij}$, $l_{ij} \le l$. Set

$$\varepsilon = \varepsilon(l) \qquad \text{and} \qquad \varepsilon_2 = \varepsilon_2(l).$$

For each pair i,j, $1 \le i < j \le t$, call a graph $\bar{P}_a^{ij}$ *good* if it is $(1/l,\varepsilon_2)$-regular. Let $s = s_{ij} \le l$ be the number of good graphs $\bar{P}_a^{ij}$ and assume that these are the graphs $\bar{P}_a^{ij}$, $a = 1,\ldots,s$. Define

$$R^{ij} = K(V_i,V_j) - \bigcup_{a=1}^{s} \bar{P}_a^{ij}.$$

Note that R^{ij}, as a complement of a union of s $(1/l, \varepsilon_2)$-regular graphs, is $(1-s/l, s\varepsilon_2)$-regular.

Partition each R^{ij} with $s < l$ into $l - s$ $(1/l, \varepsilon)$-regular graphs P_a^{ij}, $a = s+1, \ldots, l$ (a random partition will do), and set $P_a^{ij} = \bar{P}_a^{ij}$ for $a = 1, \ldots, s$. If $s = l$, set $P_1^{ij} = \bar{P}_1^{ij} \cup R^{ij}$ and $P_a^{ij} = \bar{P}_a^{ij}$ for $a = 2, \ldots, s$. Then we have $K(V_i, V_j) = \bigcup_{a=1}^{l} P_a^{ij}$, for all $1 \le i < j \le t$, and all graphs P_a^{ij}, $a = 1, \ldots, l$, are $(1/l, \varepsilon)$-regular. This is the required partition Π.

It remains to estimate the number of δ-irregular triads in Π. According to Theorem 3.5 in [5], only at most $\delta' n^3$ triplets belonged to (as triangles of) δ'-irregular triads of Π'. The changes that lead from partition Π' to the new partition Π affected only the edges of $\bigcup R^{ij}$, that is, at most $\varepsilon_1 \binom{t}{2} m^2/2 < \varepsilon_1 n^2/2$ edges belonging to ε_2-irregular graphs of Π', at most $\varepsilon_1 n^2/2$ edges between the exceptional pairs (V_i, V_j), and at most $\varepsilon_1 n^2/2$ edges in graphs $\bar{P}_0^{ij}$ with $|\bar{P}_0^{ij}| < \varepsilon_1 m^2$. Altogether,

$$\Big| \bigcup_{1 \le i < j \le j} R^{ij} \Big| \le 2\varepsilon_1 n^2.$$

The edges of $\bigcup R^{ij}$ could enlarge the set of existing $\delta' n^3$ triplets belonging to δ-irregular triads by at most $2\varepsilon_1 n^3 < \delta n^3/4$ triplets. Using Fact 4.1 and our assumption on ε, it is easy to check that a set of at most $\frac{3}{4}\delta n^3$ triplets may form at most $\delta l^3 t^3$ triads.

Note that when $s_{ij} = l$, each δ'-regular triad of Π' which contained the graph $\bar{P}_1^{ij}$, turned, after "swallowing" the graph R^{ij} into a $(\delta' + 3l\varepsilon_2(l))$-regular, hence also δ-regular triad, and thus the edges of $\bar{P}_1^{ij}$ do not need to be accounted for. ■

4.3 Regularity and hyperpaths

In order to cover a δ-regular hypergraph by many disjoint paths, it suffices to construct one path, remove it, and use the heredity. The lemma below is a very special case of Theorem 3.1.1 in [10], where all we want is just one path. In its full version, Theorem 3.1.1 in [10], under a stronger assumption of so called (δ, r)-regularity of H, guaranties the right number of copies of any fixed hypergraph, but has, therefore, a much more complicated proof which requires the strongest form of Lemma 4.1 (see [5]).

Lemma 4.2 *For all integers $k \ge 3$ and $l \ge 1$, real numbers $\alpha > 0$, and all*

$$\delta < \frac{1}{2}\min(k^{-1}, \alpha) \qquad \textit{and} \qquad \varepsilon \le \delta/(15l^3)$$

the following holds. Suppose that

- *H is a k-partite 3-uniform hypergraph with vertex partition $(V_1, \ldots, V_k)$, $|V_1| = \cdots = |V_k| = m$,*
- *P^i is a bipartite graph between V_i and V_{i+1}, $1 \leq i \leq k-1$, and Q^i is a bipartite graph between V_i and V_{i+2}, $1 \leq i \leq k-2$,*
- *all graphs P^i and Q^i are $(1/l, \varepsilon)$-regular, and*
- *for all $1 \leq i \leq k-2$, H is δ-regular with respect to the triad $\mathcal{T}^i = (P^i, P^{i+1}, Q^i)$ and has density $d_H(\mathcal{T}^i) \geq \alpha$.*

Then, H contains a path on k vertices, one from each set V_i.

Proof: Define graphs $B^{k-1}, \ldots, B^1$ recursively. Set $B^{k-1} = \emptyset$ and for each $i = k-2, \ldots, 1$, let B^i be the set of those edges $uv \in P^i$, $u \in V_i$, $v \in V_{i+1}$, for which there is no $w \in V_{i+2}$ such that $uvw \in H$ and $vw \notin B^{i+1}$.

Claim 4.1 *For each $i = 1, \ldots, k-1$, we have $|B^i| < 2(k-i)\delta|P^i|$.*

Observe that once we have proved this claim, we are done. Indeed, we may create a required path as follows. Select any edge $v_1v_2 \in P^1 - B^1$, where $v_1 \in V_1$ and $v_2 \in V_2$. This is possible, because $|P^1| > |B^1|$. Then choose $v_3 \in V_3$ so that $v_1v_2v_3 \in H$ and $v_2v_3 \in P^2 - B^2$. Continue until an entire path on k vertices is found.

Proof of Claim 4.1: We prove the claim by backward induction on $i = k-1, \ldots, 1$. It is clearly true for $i = k-1$ (recall that $B^{k-1} = \emptyset$). Assume that for some $1 \leq i < k-1$, we have

$$|B^{i+1}| < 2(k-i-1)\delta|P^{i+1}| \qquad \text{but} \qquad |B^i| \geq 2(k-i)\delta|P^i|.$$

Set

$$\mathcal{T} = \mathcal{T}^i = (P^i, P^{i+1}, Q^i) \qquad \text{and} \qquad \mathcal{T}_0 = (B^i, P^{i+1} - B^{i+1}, Q^i).$$

Our goal is to prove that

$$tr(\mathcal{T}_0) > \delta tr(\mathcal{T}), \tag{3}$$

and thus, by the δ- regularity of H with respect to $\mathcal{T}$, that

$$d_H(\mathcal{T}_0) > d_H(\mathcal{T}) - \delta > \alpha - 2\delta > 0.$$

As $d_H(\mathcal{T}_0) = |H \cap Tr(\mathcal{T}_0)/tr(\mathcal{T}_0)$, this would imply that there is an edge $uvw \in H$ with $uv \in B^i$ and $vw \in P^{i+1} - B^{i+1}$ – a contradiction with the definition of B^i.

To prove (3), note first that, by Fact 4.1,

$$tr(\mathcal{T}) < \left(\frac{1}{l^3} + 7\varepsilon\right) m^3 \tag{4}$$

and, trivially,

$$tr(\mathcal{T}_0) \geq tr(B^i, P^{i+1}, Q^i) - tr(P^i, B^{i+1}, Q^i).$$

Now, apply Fact 4.1 two more times, first the lower bound to $tr(B^i, P^{i+1}, Q^i)$, then the upper bound to $tr(P^i, B^{i+1}, Q^i)$, obtaining

$$tr(B^i, P^{i+1}, Q^i) > \left(\frac{2(k-i)\delta|P^i|}{l^2m^2} - 4\varepsilon\right) m^3,$$

$$tr(P^i, B^{i+1}, Q^i) < \left(\frac{2(k-i-1)\delta|P^i|}{l^2m^2} - 6\varepsilon\right) m^3,$$

and, finally,

$$tr(\mathcal{T}_0) > \left(\frac{2\delta}{l^3} - 10\varepsilon - \frac{2\varepsilon}{l^2}\right) m^3 > \left(\frac{1}{l^3} + 7\varepsilon\right) \delta m^3 > \delta tr(\mathcal{T}),$$

by our assumption on ε and (4). ■

It is now relatively easy to show that a δ-regular hypergraph can be almost covered by disjoint paths of a given length.

Corollary 4.1 *For all $k \geq 1$, $l \geq 1$, $0 < \alpha < 1$,*

$$\delta \leq \frac{\alpha^4}{k^{16}18^4} \qquad \textit{and} \qquad \varepsilon \leq \frac{\sqrt{\delta}}{11kl^3}$$

the following is true. If $\mathcal{T} = (P, Q, R)$ is a triad of $(1/l, \varepsilon)$-regular graphs and a hypergraph H is δ-regular with respect to $\mathcal{T}$ and has density $d_H(\mathcal{T}) \geq \alpha$, then at least $(1 - \delta^{1/4})|V(H)|$ vertices of H can be covered by vertex-disjoint $3k$-paths.

Proof: For clarity of exposition let the three vertex sets of $\mathcal{T}$ be U, V, W and $|U| = |V| = |W| = m$, where k divides m. We break each set U, V, W arbitrarily into k subsets of equal size: $U = U_1 \cup U_2 \cup \cdots \cup U_k$, etc., and claim that there are at least $(1 - \delta^{1/4})m/k$ vertex-disjoint $3k$-paths, each containing precisely one vertex from each set $U_1, V_1, W_1, U_2, V_2, W_2, \ldots, U_k, V_k, W_k$ (we will call such paths *transversal*).

Indeed, consider any family $\mathcal{Q}$ of less than $(1 - \delta^{1/4})m/k$ vertex-disjoint transversal $3k$-paths. Let subsets $U_i' \subset U_i$, $V_i' \subset V_i$, and $W_i' \subset W_i$, $i = 1, \ldots, k$, consist of all vertices not covered by these paths. Since for all $h, i, j = 1, \ldots, k$

$$|U_h'| = |V_i'| = |W_i'| > \delta^{1/4}m/k,$$

by Fact 4.2 with $\eta = \delta^{1/4}/k$ (note that $(1/l)^3 > 11\varepsilon/\eta$ and $\delta < \eta^3/3$), the subtriad $\mathcal{T}'$ induced by sets U'_h, V'_i, W'_j is such that the three bipartite subgraphs,

$$P[U'_h, V'_i], \quad Q[U'_h, W'_j], \quad R[V'_i, W'_j],$$

are all $(1/l, \varepsilon')$-regular with $\varepsilon' = \varepsilon/\eta$, and H is δ'-regular with respect to $\mathcal{T}'$ with $\delta' = 3\delta^{1/4}k^3$ and has density $d_H(\mathcal{T}') \geq \alpha - \delta \geq \alpha/2$. Since $\varepsilon' \leq \delta'/(15l^3)$ and $\delta' < \frac{1}{2}\min((3k)^{-1}, \alpha/2)$, by Lemma 4.2 applied to the subhypergraph

$$H' = H[U'_1 \cup V'_1 \cup W'_1 \cup U'_2 \cup V'_2 \cup W'_2 \cdots \cup U'_k \cup V'_k \cup W'_k]$$

with $3k$ and $\alpha/2$ in place of k and α, there is a transversal $3k$-path in H'. This path can be added to $\mathcal{Q}$. ■

5 Proof of The Path-Cover Lemma

In this final section we prove Lemma 3.1, a crucial ingredient of the proof of our main Theorem 1.1. The outline of the proof of Lemma 3.1 goes as follows. With a suitable choice of δ, t_0, and $\varepsilon(l)$,

- Obtain a partition Π with respect to H as in Lemma 4.1.
- Select from Π a system of bipartite graphs $\mathcal{P} = \{P^{i,j}, 1 < i \leq j < t\}$, such that the corresponding cluster hypergraph $K = K(\mathcal{P})$ of dense and δ-regular triads preserves essentially the property of H (see Claim 5.2 below).
- Relying on Lemma 2.4, choose from K a subsystem M of vertex-disjoint triads which cover most of the clusters.
- To each of the triads of M apply Corollary 4.1 and cover most of its vertices by disjoint paths.

To achieve the second task above, we will randomly choose one graph P_a^{ij} between each pair (V_i, V_j) and benefit from the assumption that H is an (n, γ)-graph. The obtained system K will contain a large subsystem K', $|V(K)| - |V(K')| < 12\delta^{1/4}t$, with $\Delta(G_{K'}) \leq \delta^{1/4}t$.

Proof of the Path-Cover Lemma 3.1: Let $0 < \gamma < 1/2$, n be sufficiently large, and let H be an (n, γ)-graph. Let

$$\delta = \frac{\gamma^{132}}{(12 \times 13)^4} \qquad \text{and} \qquad \varepsilon(l) = \frac{\gamma^{74}}{11 \times 6^4 l^3}$$

and $t_0 > 200/\gamma$ such that for all $t \geq t_0$ we have

$$t^2 \exp\left\{-\frac{1}{32}\gamma^2 t\right\} < \frac{1}{2}.$$

Assuming $n > N_0(\delta, t_0, \varepsilon(l))$, apply Lemma 4.1 to H with the above choice of δ, t_0 and $\varepsilon(l)$, to obtain a partition Π of $\binom{V(H)}{2}$ satisfying conditions (i) and (ii) of that lemma. Set

$$\varepsilon = \varepsilon(l) \qquad \text{and} \qquad |V_1| = \cdots = |V_t| = m,$$

and recall that $|V_0| < t$.

A triad $\mathcal{T}$ in Π is said to be *dense* if $d_H(\mathcal{T}) \geq \gamma/2$. Using the assumption that every pair of vertices in H belongs to at least $(1/2+\gamma)n$ hyperedges, we will now show that every graph P_a^{ij} belongs to nearly as large fraction of dense triads.

Claim 5.1 *In the partition Π every graph P_a^{ij} belongs to at least $(1/2+\gamma/3)tl^2$ dense triads.*

Proof: For clarity of exposition, we assume that $V_0 = \emptyset$, or, equivalently, that t divides n. Suppose to the contrary that a graph $P = P_a^{ij}$ belongs to less than $(1/2+\gamma/3)tl^2$ dense triads of Π. Let S be the set of hyperedges of H which contain an edge of P and a third vertex outside $V_i \cup V_j$. First note that, because every pair of vertices of H belongs to at least $(1/2+\gamma)n$ hyperdges of H, we have

$$|S| \geq \left(\frac{1}{l} - \varepsilon\right) m^2 \left(\frac{1}{2}n + \gamma n - 2m\right).$$

Since, $m = n/t \leq n/t_0$, $t_0 > 200/\gamma$ and $0.99\gamma < 1/2$, this leads to the bound

$$|S| > \left(\frac{1}{2} + 0.99\gamma - \varepsilon l\right) \frac{n^3}{t^2 l}.$$

We will find an upper bound on $|S|$ contradicting the above lower bound. Let us split the hyperedges of S into two classes. Let S_1 consist of those hyperedges of H which contain an edge of P, an edge of P_b^{ih} and an edge of P_c^{jh}, for some $h \in [t] \setminus \{i,j\}$ and $1 \leq b, c \leq l$, such that the triad $\mathcal{T}^{h,b,c} = (P, P_b^{ih}, P_c^{jh})$ is dense. For each such triad we will estimate the number of hyperedges of H by the number $tr(\mathcal{T}^{h,b,c})$ of triangles in (P, P_b^{ih}, P_c^{jh}), which, by Fact 4.1, is at most $(1/l^3 + 7\varepsilon)m^3$. Since we assumed that there are at most $(1/2+\gamma/3)tl^2$ such triads, we infer that,

$$|S_1| < \left(\frac{1}{2} + \frac{1}{3}\gamma\right) tl^2 \left(\frac{1}{l^3} + 7\varepsilon\right) m^3 < \left(\frac{1}{2} + \frac{1}{3}\gamma + 7\varepsilon l^3\right) \frac{n^3}{t^2 l}.$$

Let $S_2 = S \setminus S$ consist of those hyperedges of H which contain an edge of P, an edge of P_b^{ih} and an edge of P_c^{jh}, for some $h \in [t] \setminus \{i,j\}$ and $1 \le b, c \le l$, such that the triad $\mathcal{T}^{h,b,c} = (P, P_b^{ih}, P_c^{jh})$ is not dense. In this case, conversely, we estimate their number by tl^2 – the total number of triads containing P, while the number of hyperedges of H is now at most $\frac{1}{2}\gamma t r(\mathcal{T}^{h,b,c})$. Hence,

$$|S_2| < \frac{1}{2}\gamma t l^2 \left(\frac{1}{l^3} + 7\varepsilon\right) m^3 < \left(\frac{1}{2}\gamma + 7\varepsilon l^3\right) \frac{n^3}{t^2 l}$$

and

$$|S| = |S_1| + |S_2| < \left(\frac{1}{2} + \frac{5}{6}\gamma + 14\varepsilon l^3\right) \frac{n^3}{t^2 l}.$$

As $14\varepsilon l^3 + \varepsilon l < \gamma/7 < (0.99 - 5/6)\gamma$, this is a contradiction with the previously established lower bound on $|S|$. ■

Using the probabilistic method we will now select one graph from each set $\{P_a^{ij} : a = 1, \ldots, l\}$, $1 \le i < j \le t$, which (almost) maintains the property established in Claim 5.1, and in which most triads are δ-regular.

Claim 5.2 *There exists a family $\mathcal{P}$ of $\binom{t}{2}$ bipartite graphs $P^{ij} = P_{a_{ij}}^{ij}$ between pairs (V_i, V_j) such that*

(a) every graph of $\mathcal{P}$ belongs to at least $(1/2 + \gamma/12)t$ dense triads in $\mathcal{P}$,

and

(b) all but at most $2\delta t^3$ triads of $\mathcal{P}$ are δ-regular.

Proof: We apply the probabilistic method and Chernoff's and Markov's inequalities. For all $1 \le i < j \le t$, choose an index $a_{ij} \in \{1, 2, \ldots, l\}$ independently and uniformly at random. The selected indices determine a (random) family $\mathcal{P}$ of $\binom{t}{2}$ bipartite graphs.

For each $a = 1, 2, \ldots, l$, let $I_a^{ij} = 1$ if $a_{ij} = a$ and 0 otherwise. For convenience, we will abbreviate $P_{a_{ij}}^{ij} = P^{ij}$. Further, let X_a^{ij} be the number of indices $h \in [t] - \{i,j\}$ such that $(P_a^{ij}, P^{ih}, P^{jh})$ is a dense triad. Note that I_a^{ij} and X_a^{ij} are independent random variables and that X_a^{ij} is a sum of independent $0-1$ random variables with

$$\mathbb{E}(X_a^{ij}) \ge \left(\frac{1}{2} + \frac{1}{3}\gamma\right) t l^2 \left(\frac{1}{l}\right)^2 = \left(\frac{1}{2} + \frac{1}{3}\gamma\right) t.$$

Thus, by Chernoff's inequality ([7], (2.6) on page 26 and Thm. 2.8 on page 29)

$$\mathbb{P}\left\{X_a^{ij} < \left(\frac{1}{2} + \frac{1}{12}\gamma\right) t\right\} \le \mathbb{P}\left\{X_a^{ij} < \mathbb{E}X_a^{ij} - \frac{1}{4}\gamma t\right\} \le \exp\left\{-\frac{1}{32}\gamma^2 t\right\}.$$

Further, let Z_a^{ij}=1 if both $I_a^{ij} = 1$ and $X_a^{ij} < (1/2 + \gamma/12)\, t$. Then

$$\mathbb{P}\left\{\sum_{a,i,j} Z_a^{ij} > 0\right\} \leq \sum_{a,i,j} E(Z_a^{ij}) = \sum_{a,i,j} \mathbb{E}(I_a^{ij})\mathbb{P}\left\{X_a^{ij} < \left(\frac{1}{2} + \frac{1}{12}\gamma\right) t\right\}$$
$$\leq t^2 l \frac{1}{l} \exp\left\{-\frac{1}{32}\gamma^2 t\right\} = t^2 \exp\left\{-\frac{1}{32}\gamma^2 t\right\} < \frac{1}{2}$$

by our assumption on t_0.

On the other hand, by condition (ii) of Lemma 4.1, the expected number of δ-irregular triads of $\mathcal{P}$ is at most $\delta t^3 l^3 (1/l)^3 = \delta t^3$, and hence, by Markov's inequality, the probability that there are more than $2\delta t^3$ such triangles in G' is less than $1/2$. Thus, there exists a selection $\mathcal{P}$ which satisfies both (a) and (b). ■

To proceed we define the *cluster hypergraph* $K = K(\mathcal{P})$ as the 3-uniform hypergraph consisting of all triplets ijh, $1 \leq i < j < h \leq t$, corresponding to dense and δ-regular triads $\mathcal{T}^{ijh} = (P^{ij}, P^{ih}, P^{jh})$, where recall that "dense" means that $d_H(\mathcal{T}^{ijh}) \geq \gamma/2$.

Consider also the auxilliary hypergraphs $D = D(\mathcal{P})$ and $IR = IR(\mathcal{P})$ of all triplets ijh for which the triad $\mathcal{T}^{ijh} = (P^{ij}, P^{ih}, P^{jh})$ is, respectively, dense and δ-irregular. Thus, for $\mathcal{P}$ satisfying the conclusion of Claim 5.2, we have that D is a $(t, \gamma/12)$-graph, while $|IR| \leq 2\delta t^3$. Note that $K = D - IR$ does not necessarily satisfies the assumption of Lemma 2.4. To remedy this we will select carefully a large subhypergraph K' of K as follows.

Call a pair i, j of vertices of K *malicious* if it belongs to more than $\sqrt{\delta}t$ triplets of IR, that is if $deg_{IR}(ij) > \sqrt{\delta}t$. Then at most $6\sqrt{\delta}t^2$ pairs are malicious, since otherwise the sum of pair degrees would be larger than $3|IR|$ – a contradiction. Let B be the graph of malicious pairs. In turn, call a vertex i *malicious* if $deg_B(i) > \delta^{1/4}t$. At most $12\delta^{1/4}t$ vertices are malicious, since otherwise the sum of vertex degrees in B would be larger than $2|B|$ – a contradiction, again. Remove all malicious vertices obtaining a subhypergraph D' and a subgraph B', both on the same set of at least $t - 12\delta^{1/4}t$ vertices. Note that $\Delta(B') \leq \delta^{1/4}t$.

In the hypergraph $K' = D' - IR$, every pair i, j which is not an edge of B' has degree

$$deg_{B'}(ij) \geq \left(\frac{1}{2} + \frac{1}{12}\gamma - 12\delta^{1/4} - \sqrt{\delta}\right) t \geq \frac{t}{2},$$

where the last inequality follows by our choice of δ. Hence, the graph $G_{K'}$, consisting of those pairs i, j for which $deg_{K'}(ij) < t/2$, is a subgraph of B' and thus $\Delta(G_{K'}) \leq \Delta(B') \leq \delta^{1/4}t$. By Lemma 2.4, there is in K' a matching M covering all but at most $\Delta(G_{K'}) + 1 \leq 2\delta^{1/4}t$ vertices of K'.

Set $k = \lfloor \gamma^{-8} \rfloor$ and $\alpha = \gamma/2$, and note that each triad corresponding to a triplet of K' (and thus, also of M) satisfies the assumptions of Corollary 4.1. Apply Corollary 4.1 to each triad corresponding to a triplet of M to conclude that all but $2\delta^{1/4}$-fraction of the vertices of each such triad can be covered by vertex-disjoint $3k$-paths. Hence, by our choice of δ, altogether there are only at most

$$|V_0| + (12 + 2 + 1)\delta^{1/4} n < 16\delta^{1/4} n < \frac{1}{2}\gamma^{33} n < \frac{1}{2}\gamma^5 n$$

vertices not covered by this system of disjoint $3k$-paths. Clearly, as the paths in the cover have each $3k > \gamma^{-8}$ vertices, the total number of these paths does not exceed $n/3k < \gamma^8 n$. ■

6 Concluding Remarks

Remark 6.1 Using a recent result from [6] and standard derandomization techniques our proof can be turned into a polynomial time algorithm constructing a hamiltonian cycle in every (n, γ)-graph with sufficiently many vertices.

Remark 6.2 It is possible to generalize Theorem 1.1 to k-uniform hypergraphs, $k \geq 4$, in which every $(k-1)$-tuple of vertices belongs to at least $(1/2 + \gamma)n$ edges. (This is work in progress – see [12].) Moreover, a much more refined and complicated argument allows us to strengthen Theorem 1.1 to the case $\gamma = 0$, which was originally conjectured in [8]. (This is work in progress too – see [13].)

Remark 6.3 After completing their proof, the authors realized that the Frankl-Rödl Regularity Lemma for 3-uniform hypergraphs from [5] could be replaced by an application of a weaker form of the lemma, which is a straightforward generalization of the Szemerédi Regularity Lemma for graphs in [15] (see, e.g., [2] for a precise statement).

The method presented in this paper will be, however, more suitable for approaching problems of this kind which require more structure than just a path, such as hypergraph extensions of the results in [1] and [9]. This is because, unlike the weak regularity lemma, the Frankl-Rödl Regularity Lemma provides a partition into δ-regular blocks which contain many copies of any fixed subhypergraph (see the comment prior to Lemma 4.2).

References

[1] N. Alon and R. Yuster, H-factors in dense graphs, J. Combinatorial Theory, Ser. B 66 (1996), 269-282.

[2] A. Czygrinow, V. Rödl, An algorithmic regularity lemma for hypergraphs. SIAM J. Comput. 30 (2000), no. 4, 1041–1066.

[3] J. C. Bermond et al., Hypergraphes hamiltoniens, *Prob. Comb. Theorie Graph Orsay* 260 (1976) 39- 43.

[4] J. Demetrovics, G. O. H. Katona and A. Sali, Design type problems motivated by database theory, *Journal of Statistical Planning and Inference* 72 (1998) 149-164.

[5] P. Frankl and V. Rödl, Extremal problems on set systems, *Random Struct. Algorithms* 20 (2002) 131-164.

[6] P. Haxell, B. Nagle and V. Rödl, An algorithmic version of the 3-graph regularity method, in preparation.

[7] S. Janson, T. Łuczak and A. Ruciński, *Random Graphs*, John Wiley and Sons, New York (2000).

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

[9] J. Komlós, G. N. Sárközy and E. Szemerédi, On the Pósa-Seymour conjecture, *J. Graph Theory* 29 (1998) 167-176.

[10] B. Nagle and V. Rödl, Regularity properties for triple systems, *Random Struct. Algorithms* 23 (2003) 264-332.

[11] V. Rödl and A. Ruciński, Ramsey properties of random hypergraphs, *JCT A* 81 (1998) 1- 33.

[12] V. Rödl, A. Ruciński and E. Szemerédi, A Dirac-type theorem for k-uniform hypergraphs, in preparation.

[13] V. Rödl, A. Ruciński and E. Szemerédi, Dirac theorem for 3-uniform hypergraphs, in preparation.

[14] V. Rödl and M. Schacht, Regular partitions of hypergraphs, in preparation.

[15] E. Szemerédi, Regular partitions of graphs. Problemes combinatoires et theorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), pp. 399–401, Colloq. Internat. CNRS, 260, CNRS, Paris, 1978.