Loose Hamilton cycles in 3-uniform hypergraphs of high minimum degree

Daniela Kühn Deryk Osthus

Abstract

We say that a 3-uniform hypergraph has a Hamilton cycle if there is a cyclic ordering of its vertices such that every pair of consecutive vertices lies in a hyperedge which consists of three consecutive vertices. Also, let $\mathcal{C}_4$ denote the 3-uniform hypergraph on 4 vertices which contains 2 edges. We prove that for every $\varepsilon > 0$ there is an n_0 such that for every $n \geq n_0$ the following holds: Every 3-uniform hypergraph on n vertices whose minimum degree is at least $n/4 + \varepsilon n$ contains a Hamilton cycle. Moreover, it also contains a perfect $\mathcal{C}_4$-packing. Both these results are best possible up to the error term εn.

1 Introduction

1.1 Hamilton cycles

A classical theorem of Dirac states that every graph on n vertices with minimum degree at least $n/2$ contains a Hamilton cycle. If one seeks an analogue of this result for uniform hypergraphs, then several alternatives suggest themselves. In the following, we will first restrict ourselves to 3-uniform hypergraphs and then discuss the r-uniform case at the beginning of Section 1.3.

A natural way to extend the notion of the minimum degree of a graph to that of a 3-uniform hypergraph $\mathcal{H}$ is the following. Given two distinct vertices x and y of $\mathcal{H}$, the *neighbourhood* $N(x,y)$ *of* (x,y) in $\mathcal{H}$ is the set of all those vertices z which form a hyperedge together with x and y. The *minimum degree* $\delta(\mathcal{H})$ is defined to be the minimum $|N(x,y)|$ over all pairs of vertices of $\mathcal{H}$.

We say that a 3-uniform hypergraph $\mathcal{C}$ is a *cycle of order* n if there exists a cyclic ordering $v_1, \dots, v_n$ of its vertices such that every consecutive pair v_iv_{i+1} lies in a hyperedge of $\mathcal{C}$ and such that every hyperedge of $\mathcal{C}$ consists of 3 consecutive vertices. Thus the cyclic ordering of the vertices of $\mathcal{C}$ induces a cyclic ordering of its hyperedges. A cycle is *tight* if every three consecutive vertices form a hyperedge. Thus, up to isomorphisms, there is exactly one tight cycle of order n and every cycle of order n is a subhypergraph of the tight cycle of order n. A cycle of order n is *loose* if it has the minimum possible number of hyperedges among all cycles on n vertices. Thus if the number n of vertices in a loose cycle $\mathcal{C}$ is even and at least 6, then consecutive hyperedges in $\mathcal{C}$ have exactly one vertex in common and the number of hyperedges in $\mathcal{C}$ is exactly $n/2$. If $n \geq 5$ is odd, then exactly one pair of consecutive hyperedges in a loose cycle $\mathcal{C}$ have two vertices in common and the number of hyperedges in $\mathcal{C}$ is exactly

$\lceil n/2 \rceil$. A *Hamilton cycle* of a 3-uniform hypergraph $\mathcal{H}$ is a subhypergraph of $\mathcal{H}$ which is a cycle containing all its vertices.

Theorem 1.1 *For each $\sigma > 0$ there is an integer $n_0 = n_0(\sigma)$ such that every 3-uniform hypergraph $\mathcal{H}$ with $n \geq n_0$ vertices and minimum degree at least $n/4 + \sigma n$ contains a loose Hamilton cycle.*

The bound on the minimum degree in Theorem 1.1 is best possible up to the error term σn. In fact, Proposition 2.1 shows that if the minimum degree is less than $\lceil n/4 \rceil$, then we cannot even guarantee *any* Hamilton cycle. Moreover, if we had an algorithmic version of the Regularity Lemma for 3-uniform hypergraphs (Lemma 5.8), then our proof of Theorem 1.1 would yield a polynomial time randomized algorithm which finds the Hamilton cycle guaranteed by Theorem 1.1 with high probability, see Section 3.2 for more details.

Recently, Rödl, Ruciński and Szemerédi [21] proved that if the minimum degree is at least $n/2 + \sigma n$ and n is sufficiently large, then one can even guarantee a tight Hamilton cycle. Their bound is best possible up to the error term σn. In [21] they also announced that they can eliminate this error term. The proofs of both our Theorem 1.1 and the result in [21] rely on the recent Regularity Lemma for 3-uniform hypergraphs due to Frankl and Rödl [7]. However, in [21] the authors make extensive use of the fact that the intersection of the neighbourhoods of any two pairs of vertices is nonempty, which is far from true in our case. For this reason, our argument has a rather different structure.

In fact, if we assume that our hypergraph has minimum degree $n/2 + \sigma n$ and the number of vertices is divisible by four, then our result is much easier to prove: consider a random partition of the vertex set of $\mathcal{H}$ into three parts X, Y and Z so that $|Z| = 2|X| = 2|Y|$. Choose a Hamilton cycle in the complete bipartite graph spanned by X and Y uniformly at random. Consider a bipartite auxiliary graph H where the first class C consists of all the pairs of consecutive vertices in the Hamilton cycle and the other class is Z. Join a pair in C to a vertex in Z if together they form a hyperedge in $\mathcal{H}$. A standard application of a large deviation bound for the hypergeometric distribution shows that with high probability every pair in C is adjacent to at least half of the vertices in Z. By bounding the number of perfect matchings in a graph with given degrees, one can also show that with high probability, every vertex in Z is adjacent to at least half of the pairs in C—see [16] for a similar argument. Thus the minimum degree of H is at least $|Z|/2 = |C|/2$ and hence $\mathcal{H}$ contains a perfect matching. This in turn corresponds to the desired loose Hamilton cycle in $\mathcal{H}$.

A much weaker notion of a Hamilton cycle was also considered in Bermond et al. [2]. There an r-uniform graph was said to have a Hamilton cycle if there is a cyclic ordering of its vertices such that each consecutive pair of vertices lies in some hyperedge and all these hyperedges are distinct for distinct consecutive pairs. Related problems are also studied in Katona and Kierstead [12].

1.2 Perfect packings

In the case of graphs, there are many results determining the minimum degree which guarantees the existence of spanning substructures other than Hamilton cycles. For instance given two graphs H and G, an *H-packing in G* is a collection of vertex-disjoint copies of H in G. It is *perfect* if all of the vertices of G are covered. Komlós, Sárközy and Szemerédi [14] proved that given a graph H of chromatic number χ, there exists a constant c such that every sufficiently large graph G whose order n is divisible by $|H|$ and whose minimum degree is at least $(1-1/\chi)n+c$ has a perfect H-packing. For hypergraphs no analogue of this result exists so far.

The following theorem gives a perfect packing result for loose cycles of sufficient length (but bounded when compared to the order of the host hypergraph). It follows easily from Theorem 1.1 using a random vertex partition argument, see Section 12.2. We denote a loose cycle on k vertices by $\mathcal{C}_k$.

Theorem 1.2 *For any $\gamma > 0$ there is an integer $k_0 = k_0(\gamma)$ such that the following holds for all $k \geq k_0$. Suppose that $\mathcal{H}$ is a 3-uniform hypergraph whose number n of vertices is divisible by k and whose minimum degree is at least $n/4 + \gamma n$, then $\mathcal{H}$ contains a perfect $\mathcal{C}_k$-packing.*

As with Theorem 1.1, the bound on the minimum degree is best possible up to the error term γn (see Proposition 2.2). It is also best possible in the sense that the condition that $k \geq k_0(\gamma)$ is needed if k is not divisible by 4 (see Proposition 2.3).

An easy modification of the proof of Theorem 1.1 yields the following result about $\mathcal{C}_4$-packings, see Section 12.1. (Note the loose cycle $\mathcal{C}_4$ on 4 vertices has 2 hyperedges and these hyperedges share 2 vertices.)

Theorem 1.3 *For any $\gamma > 0$ there is an integer $n_1 = n_1(\gamma)$ such that every 3-uniform hypergraph $\mathcal{H}$ whose number $n \geq n_1$ of vertices is divisible by 4 and whose minimum degree is at least $n/4 + \gamma n$ contains a perfect $\mathcal{C}_4$-packing.*

In [16], we proved that every r-uniform hypergraph whose order n is sufficiently large and divisible by r and whose minimum degree is at least $n/2 + 3r^2\sqrt{n \log n}$ contains a perfect matching. This bound is best possible up to the error term $3r^2\sqrt{n \log n}$. (With a larger error term, the 3-uniform case of this also follows from the result in [21].) In view of this, Theorems 1.2 and 1.3 might be surprising at first sight since they imply that as far as the minimum degree is concerned, it is much harder to find a perfect matching in a 3-uniform hypergraph than a perfect $\mathcal{C}_4$-packing or a perfect $\mathcal{C}_k$-packing where k is sufficiently large.

1.3 Open problems

Very recently, the Regularity Lemma for 3-uniform hypergraphs due to Frankl and Rödl [7] was generalized to r-uniform hypergraphs independently by Gowers [9] and Rödl and Skokan [22]. In principle, this opens up the possibility of generalizing the above results to r-uniform hypergraphs (in which case we believe that the bound $n/4$ would be replaced by the bound $\frac{n}{2(r-1)}$, see the remark at the end of Section 2.2). However, it seems that such an extension of our results will not be quite straightforward. For example, we use some results of Dementieva et al. [5] for which analogues for r-uniform hypergraphs do not exist yet.

Another obvious open problem is of course whether the error term σn in Theorem 1.1 can be removed.

Finally, it would be very desirable to obtain more general results on perfect packings in hypergraphs. We are not aware of any results here except those mentioned in the previous two subsections. For instance the minimum degree which is necessary to force a perfect packing of the complete 3-uniform hypergraph $\mathcal{K}_4^{(3)}$ on 4 vertices is not yet known. In fact it is not even known what minimum degree forces a single copy of $\mathcal{K}_4^{(3)}$, see Czygrinow and Nagle [4] for a discussion of these two problems. To make some progress on packing problems for hypergraphs, it would be very helpful to obtain some analogue of the so-called 'Blow-up Lemma' of Komlós, Sárközy and Szemerédi [13] which was used in [14]. Roughly speaking this lemma ensures the existence of an arbitrary spanning subgraph H in a pseudo-random graph G, provided that H has bounded maximum degree. Unfortunately, so far no analogue for hypergraphs is known even for special hypergraphs $\mathcal{H}$. One advantage of our proof of Theorem 1.1 is that it gives such blow-up results in three special cases: Our argument shows that we can guarantee the existence of a spanning subhypergraph $\mathcal{H}$ in a pseudo-random hypergraph in the case when $\mathcal{H}$ is a perfect matching, when $\mathcal{H}$ is a loose Hamilton cycle and when $\mathcal{H}$ consists of a perfect $\mathcal{C}_4$-packing (see Section 11 for the Hamilton cycle and Section 12.1 for the $\mathcal{C}_4$-packing and the perfect matching).

2 Basic definitions and extremal examples

In this section, we give some very basic definitions and discuss the extremal hypergraphs which show that Theorems 1.1–1.3 are essentially best possible.

2.1 Basic definitions

Given a 3-uniform hypergraph $\mathcal{H}$ we denote the vertex set of $\mathcal{H}$ by $V(\mathcal{H})$ and the set of hyperedges by $E(\mathcal{H})$. The order $|V(\mathcal{H})|$ of $\mathcal{H}$ is denoted by $|\mathcal{H}|$. Given a pair x, y of distinct vertices in $\mathcal{H}$, we let $d(x, y) := |N(x, y)|$.

A 3-uniform hypergraph $\mathcal{P}$ is a *loose path* if there is a linear ordering of its hyperedges such that every hyperedge intersects its predecessor in exactly one vertex u and its successor in exactly one vertex $v \neq u$ and does not intersect any other hyperedge. Thus the number of vertices in $\mathcal{P}$ must be odd and $\mathcal{P}$ can be obtained from a loose cycle of even order by deleting one hyperedge e and the unique vertex w which is contained in e and not in any other hyperedge. The *length* of $\mathcal{P}$ is the number of its hyperedges. A *starting point* of $\mathcal{P}$ is a vertex which lies in its first hyperedge but not in the second one. Thus, if $\mathcal{P}$ has at least 5 vertices, then it has precisely 2 starting points. Similarly, an *endpoint* of P is a vertex which lies in its last hyperedge but not in the last but one.

2.2 Extremal examples

We first prove that the bound on the minimum degree in Theorem 1.1 is best possible up to the error term σn. In fact, Proposition 2.1 shows that if the minimum degree is less than $\lceil n/4 \rceil$, then we cannot even guarantee any Hamilton cycle.

Proposition 2.1 *For every integer $n \geq 3$ there is a 3-uniform hypergraph $\mathcal{H}$ with n vertices and minimum degree $\lceil n/4 \rceil - 1$ which does not contain a Hamilton cycle.*

Proof. Clearly, we may assume that $n \geq 5$. We construct the hypergraph $\mathcal{H}$ as follows. The vertex set of $\mathcal{H}$ is the disjoint union of two sets A and B, where $|B| = \lceil n/4 \rceil - 1$ and $A = n - |B|$. $\mathcal{H}$ contains exactly those triples of vertices whose intersection with B is non-empty. Then $\delta(\mathcal{H}) = |B|$. Suppose that $\mathcal{H}$ contains a Hamilton cycle $\mathcal{C}$. Let $v_1, \ldots, v_n$ be a cyclic ordering of the vertices of $\mathcal{C}$ as in the definition of a cycle. Consider any 4 consecutive vertices $v_i, \ldots, v_{i+3}$. We claim that at least one of these lies in B. (This would show that $|B| \geq n/4$, a contradiction.) To prove the claim, simply note that the definition of a cycle implies that the pair $v_{i+1}v_{i+2}$ lies in some hyperedge of $\mathcal{C}$ and the third vertex of this hyperedge must be v_i or v_{i+3}. The claim now follows since every hyperedge meets B. □

The following result implies that the bounds on the minimum degree in Theorems 1.2 and 1.3 are also best possible up to the error term γn. Its proof is essentially the same as that of the previous proposition.

Proposition 2.2 *Let $k, n \geq 3$ be integers such that n is divisible by k. There is a 3-uniform hypergraph $\mathcal{H}$ with n vertices and minimum degree $\lceil n/4 \rceil - 1$ which does not contain a perfect $\mathcal{C}_k$-packing.* □

Since $\lceil k/4 \rceil / k > 1/4$ for any k which is not divisible by 4, the next proposition implies that for any such k the condition that the length k of the cycle in Theorem 1.2 has to satisfy $k \geq k_0(\gamma)$ is also necessary. In the case where k is divisible by 4, we conjecture that this restriction is not necessary.

Proposition 2.3 *Let $k, n \geq 3$ be integers such that k is not divisible by 4 and n is divisible by k. There is a 3-uniform hypergraph $\mathcal{H}$ with minimum degree $n\lceil k/4 \rceil / k - 1$ which does not contain a perfect $\mathcal{C}_k$-packing.*

Proof. Write $k = 4\ell + r$, where $\ell, r \in \mathbb{N}$ and $r \leq 3$. (Thus $\ell = \lfloor k/4 \rfloor$.) Consider a hypergraph $\mathcal{H}$ as in the proof of Proposition 2.1 except that now we set $|B| := (n/k)(\ell + 1) - 1 = n\lceil k/4 \rceil / k - 1$. Then the argument there shows that every copy of $\mathcal{C}_k$ in $\mathcal{H}$ contains at least $\lceil k/4 \rceil = \ell + 1$ vertices in B. Thus $\mathcal{H}$ does not contain a perfect $\mathcal{C}_k$-packing. □

Given an integer $r \geq 4$, we say that an r-uniform hypergraph $\mathcal{C}$ is a *cycle of order* n if there exists a cyclic ordering $v_1, \dots, v_n$ of its vertices such that every consecutive pair $v_i v_{i+1}$ lies in a hyperedge of $\mathcal{C}$ and such that every hyperedge of $\mathcal{C}$ consists of r consecutive vertices. It is easy to see that the above examples generalize to r-uniform hypergraphs if we replace the term $n/4$ by $n/(2(r-1))$ in Propositions 2.1 and 2.2 and replace $k/4$ by $k/(2(r-1))$ in Proposition 2.3.

3 Overview of the proof of Theorem 1.1 and algorithmic aspects

In the first subsection, we give an overview of the proof of Theorem 1.1 which is divided into several subsections. In the second subsection, we discuss algorithmic aspects of this proof.

3.1 Overview of the proof of Theorem 1.1

In order to make the organization of the proof of Theorem 1.1 clearer, each of the following subsections in our overview of the proof corresponds to a separate section of the proof itself.

3.1.1 Basic probabilistic estimates

In Section 4 we describe a standard large deviation bound for the hypergeometric distribution which we will need very frequently.

3.1.2 The Regularity Lemma

Roughly speaking, the Regularity Lemma for graphs states that the vertex set of every dense graph can be partitioned into a bounded number of clusters V_i such that most of the bipartite graphs (V_i, V_j) induced by two of the clusters are pseudo-random in the sense that all of its sufficiently large induced subgraphs have a similar density as the original bipartite graph (V_i, V_j) (i.e. they are ε-regular). In Section 5, we introduce the Regularity Lemma for 3-uniform hypergraphs due to Frankl and Rödl [7]. There are several earlier weak versions (see e.g. the survey [15]) which state that the vertex set of every dense hypergraph can be partitioned into a bounded number of clusters V_i such that most of the tripartite hypergraphs (V_i, V_j, V_k) induced by three of the clusters are pseudo-random in the sense that all of its sufficiently large induced subhypergraphs have a similar density as the original hypergraph (V_i, V_j, V_k). However, for our purposes this notion of pseudo-randomness or hypergraph-regularity is not strong enough, so we apply the version in [7]. It states that for every dense hypergraph there is a partition of its vertices into a bounded number of clusters V_i and a partition of the edge sets of the complete bipartite graphs spanned by the pairs V_i, V_j into a bounded number of mostly ε-regular bipartite graphs P^{ij} such that in most cases, the tripartite graphs P^{ijk} consisting the union of P^{ij}, P^{ik}, P^{jk} have the following hypergraph-regularity property: for every sufficiently large induced subgraph of P^{ijk}, the proportion of triangles which correspond to hyperedges of $\mathcal{H}$ is almost the same as that of P^{ijk} (see Lemma 5.8 for the precise statement).

One advantage of this stronger notion of the pseudo-randomness (or regularity) of a hypergraph is for instance that typically the so called link graph of a vertex $x \in V_i$ in P^{ijk} is sufficiently large and $(\alpha d, \xi)$-regular (the link graph is the set of all edges in P^{jk} which form a hyperedge in $\mathcal{H}$ together with x). We will make essential use of this later on.

Given a partition of a 3-uniform hypergraph $\mathcal{H}$ as guaranteed by the Regularity Lemma, one can define a *reduced hypergraph* $\mathcal{R}$ which captures the coarse structure of $\mathcal{H}$. The vertices of $\mathcal{R}$ are the clusters V_i and the hyperedges of $\mathcal{R}$ are those triples of clusters $V_iV_jV_k$ for which there is a tripartite graph P^{ijk} which is hypergraph-regular in the sense described above and which is 'dense' in the sense that a significant proportion of its triangles correspond to hyperedges of $\mathcal{H}$ (see Definition 5.17). Thus $\mathcal{R}$ is again a 3-uniform hypergraph.

3.1.3 Finding an almost perfect cover of the reduced hypergraph $\mathcal{R}$

Given a 3-uniform hypergraph $\mathcal{H}$ of minimum degree at least $n/4 + \sigma n$ and its reduced graph $\mathcal{R}$, we can show that almost all pairs of vertices of $\mathcal{R}$ have at least $|\mathcal{R}|/4$ neighbours (see Proposition 6.1). The analogue of this result for graphs is well known and very simple to prove, to obtain the hypergraph version given here one has to work a little harder.

We define $\mathcal{H}_8$ to be the 3-uniform hypergraph on 8 vertices a_i, b_i, c_j (where $1 \leq i \leq 3$ and $1 \leq j \leq 2$) which contains all hyperedges of the form $a_ib_ic_j$. In Proposition 6.2 we show that every 3-uniform hypergraph in which most pairs of vertices form a hyperedge together with at least a quarter of the other vertices contains an almost perfect packing consisting of copies of $\mathcal{H}_8$. So in particular, this applies to $\mathcal{R}$. The fact that we seek an almost perfect packing of $\mathcal{R}$ with copies of $\mathcal{H}_8$ and not of any other hypergraph is mainly for convenience: Another candidate instead of $\mathcal{H}_8$ would have been $\mathcal{C}_4$, but we do not have an elementary proof of an analogue of Proposition 6.2 for this.

3.1.4 Almost covering $\mathcal{H}$ by triples and tidying them up

For each copy of $\mathcal{H}_8$ in the almost perfect packing of $\mathcal{R}$ we do the following: In Section 7.1 we first split the 8 clusters corresponding to this copy into 18 smaller ones to obtain triples (X_k, Y_k, Z_k) of clusters (where $1 \leq k \leq 6$) such that each of the triples is hypergraph-regular in the sense described above and such that $|X_k| = |Y_k| = |Z_k|/2$. Each of these triples corresponds to one of the 6 edges of $\mathcal{H}_8$. Thus we have covered almost all vertices of $\mathcal{H}$ by triples (X_k, Y_k, Z_k) of vertex sets.

Recall that, when defining the reduced hypergraph $\mathcal{R}$, to each hyperedge in $E(\mathcal{R}) \supseteq E(\mathcal{H}_8)$ there belongs one of the hypergraph-regular tripartite graphs P^{ijk} with the property that many triangles in P^{ijk} correspond to hyperedges of $\mathcal{H}$. So each triple (X_k, Y_k, Z_k) induces a tripartite subgraph of some P^{ijk} (where P^{ijk} is the tripartite graph belonging to the hyperedge of $\mathcal{H}_8 \subseteq \mathcal{R}$ corresponding to (X_k, Y_k, Z_k)). We denote by $P_{X_kY_k}$, $P_{Y_kZ_k}$ and $P_{X_kZ_k}$ the bipartite graphs forming this subgraph of P^{ijk}.

In Section 7.2 we then tidy up the triples (X_k, Y_k, Z_k) in the sense that we delete some of their vertices in order to obtain triples (X'_k, Y'_k, Z'_k) which are even more 'regular'. All the deleted vertices are collected in a set of so called *exceptional vertices*. The properties of the modified triples (X'_k, Y'_k, Z'_k) are summarized in Proposition 7.16. For instance, one of the properties is that for *every* vertex $z \in Z'_k$ the link graph L_z is $(\alpha d, \xi)$-regular. (The edges of L_z are now those pairs of vertices $x \in X'_k$ and $y \in Y'_k$ with $xy \in P_{X_k Y_k}$ which form a hyperedge together with z.)

3.1.5 Incorporating the exceptional vertices and connecting the triples by bridges

To the exceptional vertices mentioned above we also add those vertices of $\mathcal{H}$ lying in clusters which are not contained in the almost perfect $\mathcal{H}_8$-packing of $\mathcal{R}$. In Section 8.1 we then find a loose path $\mathcal{L}$ in $\mathcal{H}$ which contains all the exceptional vertices as well as some of the vertices of the Z'_k. We choose $\mathcal{L}$ in such a way that when removing its vertices from the sets Z'_k we do not destroy too much of the regularity properties of the triples (X'_k, Y'_k, Z'_k).

In Section 8.2 we find for each successive pair of triples a hyperedge which contains one vertex from both of the triples (its third vertex is taken from some other triple). We call the hyperedge connecting the kth triple to the $(k+1)$th triple the kth bridge. The first bridge will not be a single hyperedge but a longer path which contains $\mathcal{L}$ as a subpath. If n is even then this path will be loose (Fig. 1). If n is odd then it will be loose except for two successive hyperedges which have two vertices in common.

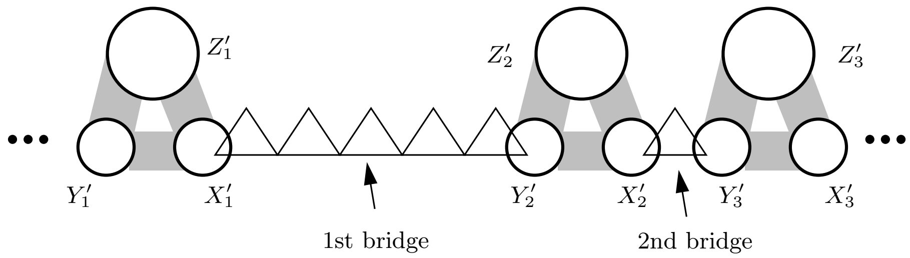

Figure 1: Connecting the triples by bridges in the case when n is even

We would now like to find a loose Hamilton path in each of the triples (X'_k, Y'_k, Z'_k) which starts in an endpoint of the $(k-1)$th bridge and ends in a starting point of the kth bridge. (These two vertices in $X_k \cup Y_k$ are called the bridge vertices of (X'_k, Y'_k, Z'_k).) For this, a trivial necessary condition is that each (X'_k, Y'_k, Z'_k) contains an odd number of vertices. (Here we count the bridge vertices but not any other vertices lying in bridges.) This can be obtained by deleting some further vertices from the triples which are chosen such that they form a loose path extending the first bridge. We still denote each modified triple by (X'_k, Y'_k, Z'_k).

3.1.6 Finding the equalizing paths and augmenting the bridges

Consider some triple (X'_k, Y'_k, Z'_k). In Section 9 we find a loose path $\mathcal{Q}_k$ which contains only vertices from (X'_k, Y'_k, Z'_k), begins in the starting point of the kth bridge and has the following additional property: Let $(X^{**}_k, Y^{**}_k, Z^{**}_k)$ denote the subtriple obtained from (X'_k, Y'_k, Z'_k) by deleting the endpoint of the $(k-1)$th bridge as well as all the vertices in $\mathcal{Q}_k$. Then $2|X^{**}_k| + 1 = 2|Y^{**}_k| + 1 = |Z^{**}_k|$. Moreover, the sets X^{**}_k, Y^{**}_k and Z^{**}_k will be small compared to X'_k, Y'_k and Z'_k. We call $\mathcal{Q}_k$ an equalizing path and augment the kth bridge into a longer path by adding $\mathcal{Q}_k$ to it. The existence of $\mathcal{Q}_k$ is shown using a greedy argument and the hypergraph-regularity of (X'_k, Y'_k, Z'_k).

3.1.7 Perfect matchings in superregular graphs

In Section 10 we collect results from [17] about (random) perfect matchings in ε-regular graphs. Theorem 10.3 is the main result of this section. This will be an important tool when choosing the path R^* in Section 11 (see below).

3.1.8 Finding a loose Hamilton path in the remainder of each triple

In fact, before we incorporate the exceptional vertices and glue the triples together as described in Section 3.1.5, from each triple (X'_k, Y'_k, Z'_k) we set aside a subtriple (X''_k, Y''_k, Z''_k) using a probabilistic argument. Thus (X''_k, Y''_k, Z''_k) is very regular. Moreover, it will satisfy $2|X''_k| = 2|Y''_k| = |Z''_k|$ and X''_k will be much larger than each of X^{**}_k, Y^{**}_k and Z^{**}_k. The latter property ensures that the triple $W_k := (X^{**}_k \cup X''_k, Y^{**}_k \cup Y''_k, Z^{**}_k \cup Z''_k)$ is still rather regular. (We cannot guarantee this for $(X^{**}_k, Y^{**}_k, Z^{**}_k)$ since it was left over by a greedy argument. This is the reason to set aside the random subtriple (X''_k, Y''_k, Z''_k) earlier.) In Section 11 we will find a Hamilton path R^*_k in the bipartite subgraph G^*_k of $P_{X_k Y_k}$ induced by $X^{**}_k \cup X''_k$ and $Y^{**}_k \cup Y''_k$ (Fig. 2).

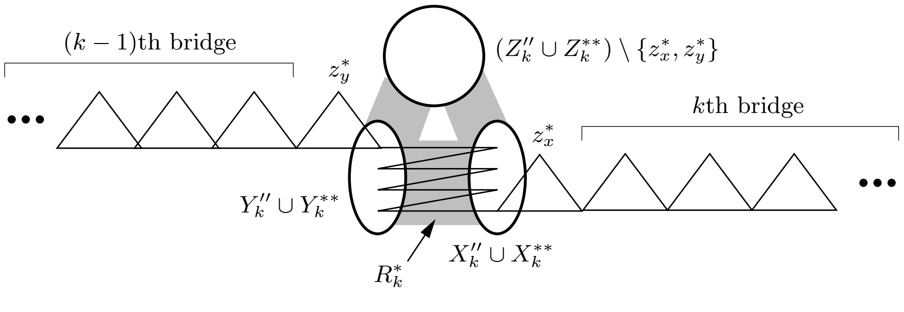

Figure 2: The Hamilton path R^*_k which forms the 'base' of $\mathcal{Q}^*_k$.

This Hamilton path R^*_k will have the following properties: Firstly, there are vertices $z^*_x, z^*_y \in Z''_k$ such that z^*_y forms a hyperedge together with one endvertex of R^*_k and an endpoint of the $(k-1)$th bridge while z^*_x forms a hyperedge together with the one endvertex of R^*_k and a starting point of the kth bridge. Secondly, consider the auxiliary bipartite graph H^*_k whose vertex classes are $Z^{**}_k \cup (Z''_k \setminus \{z^*_x, z^*_y\})$ and the edges of R^*_k and where an edge of R^*_k

is joined to a vertex z if they form a hyperedge in $\mathcal{H}$. Then H_k^* will contain a perfect matching. The perfect matching in H_k^* obviously corresponds to a loose Hamilton path $\mathcal{Q}_k^*$ in W_k which connects the $(k-1)$th bridge to the kth bridge. All these loose paths $\mathcal{Q}_k^*$ together with all the bridges then form our desired loose Hamilton cycle of $\mathcal{H}$.

The existence of the graph Hamilton path R_k^* is (very roughly) shown as follows: We consider the 2-factor obtained by choosing two random perfect matchings in G_k^* and turn it into a Hamilton cycle by changing a few of its edges. This enables us to use Theorem 10.3 to provide information about the auxiliary graph H_k^*: For instance, we apply Theorem 10.3 with $H = L_z[X_k^{**} \cup X_k'' \cup Y_k^{**} \cup Y_k'']$, where L_z denotes the link graph of a given vertex in $Z_k^{**} \cup Z_k''$ (i.e. those edges of G_k^* which form a hyperedge together with z). In this case the lemma implies that R_k^* contains many edges of L_z and thus that z has many neighbours in H_k^*. The latter property will be useful when showing that H_k^* satisfies Hall's condition and thus contains a perfect matching.

3.2 Algorithmic aspects

As mentioned in the introduction, if we had an algorithmic version of the Regularity Lemma for 3-uniform hypergraphs (Lemma 5.8), then our proof of Theorem 1.1 would yield a polynomial time randomized algorithm which finds the Hamilton cycle guaranteed by Theorem 1.1 with high probability. Such an algorithmic version is known for the case $r = 1$ of Lemma 5.8 (see [6]). As mentioned after Lemma 5.19, the only place where we use the (δ_*, r)-regularity instead of just the $(\delta_*, 1)$-regularity is Lemma 5.19 due to Dementieva et al. [5]. In fact, in [5] the authors conjecture that Lemma 5.19 even holds for $r = 1$, i.e. if we only assume that the triad P is $(\delta_*, 1)$-regular. Thus if either this conjecture is true or if there is an algorithmic version of Lemma 5.8 then the proof of Theorem 1.1 translates into a randomized polynomial time algorithm which finds the loose Hamilton cycle guaranteed by the theorem with high probability (the same also holds for the packings guaranteed by Theorems 1.2 and 1.3).

With two exceptions, the translation of the rest of our argument into an algorithm is elementary:

The first exception is in Section 10.1 where we need the result of Jerrum, Sinclair and Vigoda [11] (based on rapidly mixing Markov chains) that in an arbitrary bipartite graph one can sample a random perfect matching in polynomial time so that the distribution is almost uniform. More precisely, given ε there is a randomized algorithm whose running time is polynomial in the number of vertices of the input graph G and which does the following: the algorithm returns a random perfect matching M in G so that for any perfect matching M' in G we have $1/(2K) \leq \mathbb{P}(M = M') \leq 2/K$, where K denotes the number of perfect matchings in G. Thus with high probability, in Section 10.1 we can find a random perfect matching M which is almost uniformly distributed. Since the error probabilities in Section 10.1 are all exponential, the extra factor 2 in the distribution of M is not a problem.

The second exception is the nontrivial but well-known fact that a maximum matching in a bipartite graph can be found in polynomial time (see e.g. [18]).

Thus the perfect matchings guaranteed in Sections 10.2 and 11.2 can be found in polynomial time.

4 Basic probabilistic estimates

Given a positive number ε and sets $A, Q \subseteq T$, we say that A *is split ε-fairly by Q* if

$$\left| \frac{|A \cap Q|}{|Q|} - \frac{|A|}{|T|} \right| \leq \varepsilon.$$

Thus, if ε is small and A is split ε-fairly by Q, then the proportion of all those elements of T which lie in A is almost equal to the proportion of all those elements of Q which lie in A. We will use the following version of the well-known fact that if Q is random then it tends to split large sets ε-fairly.

Proposition 4.1 *For each $0 < \varepsilon < 1$ there exists an integer $q_0 = q_0(\varepsilon)$ such that the following holds. Given $t \geq q \geq q_0$ and a set T of size t, let Q be a subset of T which is obtained by successively selecting q elements uniformly at random without repetitions. Let $\mathcal{A}$ be a family of at most q^{10} subsets of T such that $|A| \geq \varepsilon t$ for each $A \in \mathcal{A}$. Then with probability at least $1/2$ every set in $\mathcal{A}$ is split ε-fairly by Q.*

To prove Proposition 4.1 we will use the following large deviation bound for the hypergeometric distribution (see e.g. [10, Thm. 2.10 and Cor. 2.3]).

Lemma 4.2 *Given $q \in \mathbb{N}$ and sets $A \subseteq T$ with $|T| \geq q$, let Q be a subset of T which is obtained by successively selecting q elements of T uniformly at random without repetitions. Let $X := |A \cap Q|$. Then for all $0 < \varepsilon < 1$ we have*

$$\mathbb{P}(|X - \mathbb{E}X| \geq \varepsilon \mathbb{E}X) \leq 2\mathrm{e}^{-\frac{\varepsilon^2}{3}\mathbb{E}X}.$$

Proof of Proposition 4.1. Given $A \in \mathcal{A}$, Lemma 4.2 implies that

$$\begin{aligned}\mathbb{P}(A \text{ is not split } \varepsilon\text{-fairly by } Q) &\leq \mathbb{P}(|\,|A \cap Q| - q|A|/t\,| \geq \varepsilon q|A|/t) \\ &\leq 2\mathrm{e}^{-\frac{\varepsilon^2}{3}\frac{q|A|}{t}} \leq 2\mathrm{e}^{\frac{-\varepsilon^3 q}{3}}.\end{aligned}$$

Hence, if q_0 is sufficiently large compared with ε, the probability that there is an $A \in \mathcal{A}$ which is not split ε-fairly is at most $2q^{10}\mathrm{e}^{-\varepsilon^3 q/3} < 1/2$, as required. □

5 The Regularity Lemma

5.1 Regular bipartite graphs

In this subsection, we introduce some well known definitions and facts about ε-regular graphs. Given a bipartite graph $G = (A, B)$ with vertex classes A and B, we denote the edge set of G by $E(A, B)$ and let $e(G) := e(A, B) :=$

$|E(A,B)|$. We write $N_G(x)$ for the neighbourhood of a vertex x in G and let $d_G(x) := |N_G(x)|$. The *density* of a bipartite graph $G = (A, B)$ is defined to be

$$d(A,B) := \frac{e(A,B)}{|A||B|}.$$

Given $\varepsilon > 0$, we say that G is *ε-regular* if for all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ we have $|d(A,B) - d(X,Y)| < \varepsilon$. Given $0 < \varepsilon, d \leq 1$, we say that G is *(d,ε)-regular* if for all sets $X \subseteq A$ and $Y \subseteq B$ with $|X| \geq \varepsilon|A|$ and $|Y| \geq \varepsilon|B|$ we have $(1-\varepsilon)d < d(X,Y) < (1+\varepsilon)d$. These two concepts are more or less equivalent. Indeed, every (d,ε)-regular graph is $2\varepsilon d$-regular and thus also 2ε-regular. Conversely, if $d := d(A,B) \geq \sqrt{\varepsilon}$, then every ε-regular bipartite graph (A,B) is $(d,\sqrt{\varepsilon})$-regular. Given $d \in [0,1]$, we say that G is *(d,ε)-superregular* if it is (d,ε)-regular and, furthermore, if $(1-\varepsilon)d|B| < d_G(a) < (1+\varepsilon)d|B|$ for all vertices $a \in A$ and $(1-\varepsilon)d|A| < d_G(b) < (1+\varepsilon)d|A|$ for all $b \in B$. Note that the latter two notions of regularity also make sense if we allow ε to be larger than d.

We will often use the following simple fact.

Proposition 5.1 *Given a (d,ε)-regular bipartite graph (A,B) and a set $X \subseteq A$ with $|X| \geq \varepsilon|A|$, there are less than $\varepsilon|B|$ vertices in B which have at most $(1-\varepsilon)d|X|$ neighbours in X. Similarly, there are less than $\varepsilon|B|$ vertices in B which have at least $(1+\varepsilon)d|X|$ neighbours in X.*

Using Proposition 5.1 it is easy to prove the following well-known fact which states that, given $d \gg \varepsilon$ and a 3-partite graph P formed by (d,ε)-regular bipartite graphs, one can slightly modify the vertex classes of P such that each of the 3 bipartite graphs becomes superregular.

Proposition 5.2 *Let $d \geq 8\sqrt{\varepsilon}$ and let P be a 3-partite graph with vertex classes V_1, V_2 and V_3 such that $|V_1| = |V_2| = |V_3| =: m$. Suppose that each of the bipartite graphs $P[V_1 \cup V_2]$, $P[V_2 \cup V_3]$, and $P[V_1 \cup V_3]$ is (d,ε)-regular. Then each V_i contains a subset V_i' of size $(1-4\varepsilon)m$ such that all the restricted bipartite graphs $P[V_1' \cup V_2']$, $P[V_2' \cup V_3']$ and $P[V_1' \cup V_3']$ are $(d,\sqrt{\varepsilon})$-superregular.*

5.2 The Regularity Lemma for 3-uniform hypergraphs

The main purpose of this section is to introduce the Regularity Lemma for 3-uniform hypergraphs due to Frankl and Rödl [7]. Before we can state it, we will collect the necessary definitions.

Consider a 3-uniform hypergraph $\mathcal{H}$. Suppose that $V_0, V_1, \dots, V_t$ is a partition of $V(\mathcal{H})$ such that $|V_1| = \dots = |V_t| = \lfloor |V|/t \rfloor$. Furthermore, suppose that for each pair $1 \leq i < j \leq t$ we are given a family $(P_\alpha^{ij})_{\alpha=0}^{\ell_{ij}}$ of $\ell_{ij}+1$ edge-disjoint bipartite graphs with vertex classes V_i and V_j such that the union $\bigcup_{\alpha=0}^{\ell_{ij}} P_\alpha^{\ell_{ij}}$ of all these graphs is the complete bipartite graph with vertex classes V_i and V_j.

We will refer to each 3-partite graph of the form $P_\alpha^{ij} \cup P_\beta^{jk} \cup P_\gamma^{ik}$ (where $1 \leq i < j < k \leq t$) as a *triad*. More precisely, given a triple $V_iV_jV_k$ ($1 \leq i < j < k \leq t$), we refer to each triad of the form $P_\alpha^{ij} \cup P_\beta^{jk} \cup P_\gamma^{ik}$ as a *triad* *belonging to the triple* $V_iV_jV_k$. We say that P^{ij}_α, P^{jk}_β and P^{ik}_γ are the *bipartite graphs forming the triad* $P^{ij}_\alpha \cup P^{jk}_\beta \cup P^{ik}_\gamma =: P$. We write $T(P)$ for the set of all triangles contained in P and let $t(P)$ denote the number of these triangles.

We will usually consider triads P which are formed by (d,ε)-regular bipartite graphs. For such triads, one can easily estimate $t(P)$ using Proposition 5.1:

Proposition 5.3 *Let* $0 < 2\varepsilon^{1/6} \le d \le 1$ *be given constants. Suppose that* $P = P^{ij}_\alpha \cup P^{jk}_\beta \cup P^{ik}_\gamma$ *is a triad formed by* (d,ε)*-regular bipartite graphs. Then*

$$(1-\sqrt{\varepsilon})(1-\varepsilon)^3 d^3 |V_i||V_j||V_k| < t(P) < (1+\sqrt{\varepsilon})(1+\varepsilon)^3 d^3 |V_i||V_j||V_k|.$$

Definition 5.4 (densities $d_{\mathcal{H}}(P)$ and $d_{\mathcal{H}}(\vec{Q})$) The *density of a triad* $P = P^{ij}_\alpha \cup P^{jk}_\beta \cup P^{ik}_\gamma$ *with respect to* $\mathcal{H}$ is defined by

$$d_{\mathcal{H}}(P) := \begin{cases} |E(\mathcal{H}) \cap T(P)|/t(P) & \text{if } t(P) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

In other words, $d_{\mathcal{H}}(P)$ denotes the proportion of all those triangles in P which are hyperedges of $\mathcal{H}$. More generally, suppose that we are given an r-tuple $\vec{Q} = (Q(1),\dots,Q(r))$ of subtriads of P, where $Q(s) = Q^{ij}_\alpha(s) \cup Q^{jk}_\beta(s) \cup Q^{ik}_\gamma(s)$, $Q^{ij}_\alpha(s) \subseteq P^{ij}_\alpha$, $Q^{jk}_\beta(s) \subseteq P^{jk}_\beta$ and $Q^{ik}_\gamma(s) \subseteq P^{ik}_\gamma$ for all $s = 1,\dots,r$. Put

$$t(\vec{Q}) := \left| \bigcup_{s=1}^{r} T(Q(s)) \right|.$$

The *density of* $\vec{Q}$ *with respect to* $\mathcal{H}$ is defined to be

$$d_{\mathcal{H}}(\vec{Q}) := \begin{cases} |E(\mathcal{H}) \cap \bigcup_{s=1}^{r} T(Q(s))|/t(\vec{Q}) & \text{if } t(\vec{Q}) > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Note that in Definition 5.4 the sets $T(Q(s))$ of triangles need not necessarily be disjoint.

Definition 5.5 ((δ_*, r)-regularity of triads with respect to $\mathcal{H}$) Given an integer r and a constant $\delta_* > 0$, we say that a triad P is (δ_*, r)*-regular* with respect to $\mathcal{H}$ if for every r-tuple $\vec{Q} = (Q(1),\dots,Q(r))$ of subtriads of P with

$$t(\vec{Q}) \ge \delta_* \cdot t(P)$$

we have

$$|d_{\mathcal{H}}(P) - d_{\mathcal{H}}(\vec{Q})| < \delta_*.$$

Definition 5.6 ($(\ell, t, \varepsilon_1, \varepsilon_2)$-partition) Let V be a set. An $(\ell, t, \varepsilon_1, \varepsilon_2)$*-partition* $\mathcal{P}$ *of* V is a partition into $V_0, V_1, \dots, V_t$ together with families $(P^{ij}_\alpha)^{\ell_{ij}}_{\alpha=0}$ $(1 \le i < j \le t)$ of edge-disjoint bipartite graphs such that

(i) $|V_1| = \dots = |V_t| = \lfloor |V|/t \rfloor =: m_*$,

(ii) $\ell_{ij} \leq \ell$ for all pairs $1 \leq i < j \leq t$,

(iii) $\bigcup_{\alpha=0}^{\ell_{ij}} P_\alpha^{ij}$ is the complete bipartite graph with vertex classes V_i and V_j (for all pairs $1 \leq i < j \leq t$),

(iv) all but at most $\varepsilon_1 \binom{t}{2} m_*^2$ edges of the complete t-partite graph $K(V_1, \dots, V_t)$ with vertex classes $V_1, \dots, V_t$ lie in some ε_2-regular graph P_α^{ij},

(v) for all but at most $\varepsilon_1 \binom{t}{2}$ pairs V_i, V_j $(1 \leq i < j \leq t)$ we have $e(P_0^{ij}) \leq \varepsilon_1 m_*^2$ and

$$|d_{P_\alpha^{ij}}(V_i, V_j) - 1/\ell| \leq \varepsilon_2$$

for all $\alpha = 1, \dots, \ell_{ij}$.

Definition 5.7 ((δ_*, r)-regular $(\ell, t, \varepsilon_1, \varepsilon_2)$-partition) Suppose that $\mathcal{H}$ is a 3-uniform hypergraph and that $V_0, V_1, \dots, V_t$ is an $(\ell, t, \varepsilon_1, \varepsilon_2)$-partition of the vertex set $V(\mathcal{H})$ of $\mathcal{H}$. Set $m_* := |V_1| = \cdots = |V_t|$. Recall that a triad is a 3-partite graph of the form $P = P_\alpha^{ij} \cup P_\beta^{jk} \cup P_\gamma^{ik}$. We say that the partition $V_0, V_1, \dots, V_t$ is (δ_*, r)-*regular* if

$$\sum_{\text{irregular}} t(P) < \delta_* |\mathcal{H}|^3,$$

where $\sum_{\text{irregular}}$ denotes the sum over all triads P which are not (δ_*, r)-regular with respect to $\mathcal{H}$.

We can now state the Regularity Lemma for 3-uniform hypergraphs which was proved by Frankl and Rödl [7].

Lemma 5.8 (Regularity Lemma for 3-uniform hypergraphs) *For all δ_* and ε_1 with $0 < \varepsilon_1 \leq 2\delta_*^4$, for all $t_0, \ell_0 \in \mathbb{N}$ and for all integer-valued functions $r = r(t, \ell)$ and all decreasing functions $\varepsilon_2(\ell)$ with $0 < \varepsilon_2(\ell) \leq 1/\ell$, there exist integers T_0, L_0 and N_0 such that the vertex set of any 3-uniform hypergraph $\mathcal{H}$ of order $|\mathcal{H}| \geq N_0$ admits a $(\delta_*, r(t, \ell))$-regular $(\ell, t, \varepsilon_1, \varepsilon_2(\ell))$-partition for some t and ℓ satisfying $t_0 \leq t \leq T_0$ and $\ell_0 \leq \ell \leq L_0$.*

The elements $V_1, \dots, V_t$ of the $(\ell, t, \varepsilon_1, \varepsilon_2(\ell))$-partition given by Lemma 5.8 are called *clusters*. V_0 is the *exceptional set*. As we shall see later on, the main difficulty when working with the above lemma arises from the fact that the density of the regular triads (which is roughly $1/\ell$) is very small compared to δ_*, the parameter which measures the hypergraph regularity of the partition guaranteed by the lemma.

5.3 Definition of the reduced hypergraph

When we apply the Regularity Lemma to a graph G, we often consider the so called reduced graph, whose vertices are the clusters V_i and whose edges correspond to those pairs of clusters which induce an ε-regular bipartite graph of sufficient density. Analogously, we will now define a 3-uniform reduced hypergraph. But before we can do this, we need some more definitions.

Also, throughout the rest of the paper we fix positive constants satisfying the following hierarchy:

$$1/\delta_* \ll \ell \ll r(t,\ell), \tag{1}$$

$$\varepsilon_2(\ell) \ll \min\{\varepsilon_1, 1/\ell\}, \tag{2}$$

$$\max\{\varepsilon_1, 1/t_0\} \ll \delta_* \ll \varepsilon_3 \ll \xi \ll \eta_1 \ll \eta \ll \alpha_* \ll \sigma. \tag{3}$$

The results and observations in this section are then valid for any choice of constants $\varepsilon_1, \varepsilon_2 = \varepsilon_2(\ell), \varepsilon_3, \xi, \delta_*, \alpha_*, r = r(\ell,t), t_0$ satisfying the above and for any hypergraph $\mathcal{H}$ to which we apply the Regularity Lemma (Lemma 5.8) with these $\varepsilon_1, \varepsilon_2 = \varepsilon_2(\ell), \delta_*, r = r(\ell,t), t_0$. We then let ℓ and t be as defined in Lemma 5.8.

We also define the following constants:

$$\varepsilon' := 4\varepsilon_2^{1/4},\ d := 1/\ell,\ \delta := \sqrt{\delta_*},$$

(which first occur in Section 7.1),

$$\delta_0 := \delta^{1/4},\ \delta_1 := 24\delta_0/\delta^{1/5},$$

(which first occur in Section 7.2) and

$$\varepsilon := 50\sqrt{\varepsilon'},\ \nu := 10^5\sqrt{\xi},$$

(which first occur in Section 11.1). Together with (1)–(3) this implies that

$$\varepsilon' \ll d \ll \delta \ll \xi \ll \eta_1 \ll \eta \ll \alpha_*. \tag{4}$$

Definition 5.9 (good pair V_iV_j) We call a pair V_iV_j $(1 \le i < j \le t)$ of clusters *good* if it satisfies the following two properties:

- $e(P_0^{ij}) \le \varepsilon_1 m_*^2$ and $|d_{P_\alpha^{ij}}(V_i,V_j) - 1/\ell| \le \varepsilon_2$ for all $\alpha = 1,\dots,\ell_{ij}$. (This means that V_iV_j does not belong to the at most $\varepsilon_1\binom{t}{2}$ exceptional pairs described in Definition 5.6(v).)
- at most $\varepsilon_3\ell/6$ of the bipartite graphs P_α^{ij} $(1 \le \alpha \le \ell_{ij})$ are not $(1/\ell, \sqrt{\varepsilon_2})$-regular.

Let us now estimate the number of bad pairs. It is easy to see that if V_iV_j satisfies the first condition in Definition 5.9, then every ε_2-regular P_α^{ij} is $(1/\ell, \sqrt{\varepsilon_2})$-regular. Thus for each bad pair which satisfies the first condition in Definition 5.9, at least $\varepsilon_3\ell/6$ of the bipartite graphs P_α^{ij} are not ε_2-regular. So for each bad pair V_iV_j which satisfies the first condition in Definition 5.9 there are at least $(\varepsilon_3\ell/6)(1/\ell - \varepsilon_2)m_*^2 \ge \varepsilon_3 m_*^2/8$ edges running between V_i and V_j which are not contained in some ε_2-regular bipartite graph P_α^{ij}. But by Definition 5.6(iv) at most $\varepsilon_1\binom{t}{2}m_*^2$ edges of $K(V_1,\dots,V_t)$ are not contained in ε_2-regular graphs P_α^{ij}. Thus there are at most

$$\frac{\varepsilon_1\binom{t}{2}m_*^2}{\varepsilon_3 m_*^2/8} \le \frac{4\varepsilon_1 t^2}{\varepsilon_3}$$

such bad pairs V_iV_j. Together with Definition 5.6(v) this immediately implies the following

Proposition 5.10 *At most $5\varepsilon_1 t^2/\varepsilon_3$ pairs V_iV_j of clusters are bad.* □

Definition 5.11 (good triple $V_iV_jV_k$) We call a triple $V_iV_jV_k$ ($1 \le i < j < k \le t$) of clusters *good* if both of the following hold:

- each of the pairs V_iV_j, V_jV_k and V_iV_k is good,
- at most $\varepsilon_3\ell^3$ of the triads belonging to $V_iV_jV_k$ are not (δ_*, r)-regular with respect to $\mathcal{H}$.

The next proposition states that only a small fraction of the triples $V_iV_jV_k$ is bad (i.e. not good).

Proposition 5.12 *At most $6\delta_* t^3/\varepsilon_3$ triples $V_iV_jV_k$ of clusters are bad.*

Proof. Proposition 5.10 implies that at most $5\varepsilon_1 t^3/\varepsilon_3$ triples violate the first condition in Definition 5.11. Thus it remains to consider the number of all those triples that satisfy the first condition of Definition 5.11 but violate the second condition. Let $V_iV_jV_k$ be such a triple (if it exists). Then by Definition 5.9, at least $\varepsilon_3\ell^3 - 3\cdot\varepsilon_3\ell^3/6 = \varepsilon_3\ell^3/2$ of the triads $P^{ij}_\alpha \cup P^{jk}_\beta \cup P^{ik}_\gamma$ have the property that they are not (δ_*, r)-regular with respect to $\mathcal{H}$ but each of their bipartite subgraphs P^{ij}_α, P^{jk}_β and P^{ik}_γ is $(1/\ell, \sqrt{\varepsilon_2})$-regular. Proposition 5.3 now implies that all these triads together contain at least $(1-\varepsilon_2^{1/4})(1-\sqrt{\varepsilon_2})^3\ell^{-3}m_*^3\cdot\varepsilon_3\ell^3/2 \ge \varepsilon_3 m_*^3/4$ triangles of the complete 3-partite graph $K(V_i, V_j, V_k) \subseteq K(V_1, \dots, V_t)$. On the other hand, by Definition 5.7, at most $\delta_*|\mathcal{H}|^3$ triangles of the complete t-partite graph $K(V_1, \dots, V_t)$ do not lie in triads which are (δ_*, r)-regular with respect to $\mathcal{H}$. This shows that at most

$$\frac{\delta_*|\mathcal{H}|^3}{\varepsilon_3 m_*^3/4} \le \frac{5\delta_* t^3}{\varepsilon_3}$$

triples $V_iV_jV_k$ satisfy the first condition in Definition 5.11 but violate the second condition. Thus in total there are at most

$$\frac{5\varepsilon_1 t^3}{\varepsilon_3} + \frac{5\delta_* t^3}{\varepsilon_3} \le \frac{6\delta_* t^3}{\varepsilon_3}$$

bad triples. □

Definition 5.13 (useful pair V_iV_j) We call a pair V_iV_j of clusters *useful* if there are at most $\varepsilon_3 t$ clusters V_k such that $V_iV_jV_k$ is a bad triple.

Note that by definition, a useful pair is always good. Moreover, since each useless pair belongs to at least $\varepsilon_3 t$ bad triples and each bad triple is counted at most 3 times in this way, Proposition 5.12 immediately implies the following

Proposition 5.14 *At most $18\delta_* t^2/\varepsilon_3^2$ pairs V_iV_j of clusters are useless.* □

Definition 5.15 (useful triad P) We call a triad P *useful* if

(i) P is formed by $(1/\ell, \sqrt{\varepsilon_2})$-regular bipartite graphs,

(ii) P is (δ_*, r)-regular with respect to $\mathcal{H}$,

(iii) $d_{\mathcal{H}}(P) \geq \alpha_*$.

Definition 5.16 (useful triple $V_iV_jV_k$) We call a triple $V_iV_jV_k$ of clusters *useful* if it is good and if there is a useful triad P belonging to $V_iV_jV_k$.

We are now ready to define the reduced hypergraph $\mathcal{R}$.

Definition 5.17 (Reduced hypergraph) The vertices of the *reduced hypergraph* $\mathcal{R}$ are all the clusters $V_1, \dots, V_t$. The hyperedges of $\mathcal{R}$ are precisely the useful triples $V_iV_jV_k$.

Thus, similarly as $\mathcal{H}$, also $\mathcal{R}$ is a 3-uniform hypergraph.

For each hyperedge $V_iV_jV_k$ of the reduced graph $\mathcal{R}$ we now fix a useful triad P belonging to $V_iV_jV_k$. We denote the bipartite graphs forming P by P^{ij}, P^{jk} and P^{ik}.

Proposition 6.1 in Section 6 shows that the minimum degree of $\mathcal{H}$ is 'almost inherited' by its reduced hypergraph $\mathcal{R}$ in the sense that only a small fraction of the vertex pairs in $\mathcal{R}$ have a neighbourhood whose size is significantly smaller than $\frac{\delta(\mathcal{H})}{|\mathcal{H}|} \cdot |\mathcal{R}|$.

Definition 5.18 (Link and colink graphs) Given a hyperedge $V_iV_jV_k$ of $\mathcal{R}$ and a vertex $x \in V_i$, we define the *link graph L_x of x* to be the bipartite graph whose vertex classes are $N_{P^{ij}}(x) \subseteq V_j$ and $N_{P^{ik}}(x) \subseteq V_k$ and in which $y \in N_{P^{ij}}(x)$ and $z \in N_{P^{ik}}(x)$ are joined by an edge if and only if xyz is a hyperedge in $E(\mathcal{H}) \cap T(P)$. Thus L_x is a subgraph of P^{jk}. The link graphs of vertices lying in V_j or V_k are defined analogously.

Given distinct vertices $x, x' \in V_i$, we define the *colink graph $L_{xx'}$ of x, x'* to be the bipartite graph whose vertex classes are $N_{P^{ij}}(x) \cap N_{P^{ij}}(x') \subseteq V_j$ and $N_{P^{ik}}(x) \cap N_{P^{ik}}(x') \subseteq V_k$ and whose edge set is $E(L_x) \cap E(L_{x'})$. The colink graphs for pairs of vertices from V_j or from V_k are defined analogously.

The following lemma shows that almost all link graphs and almost all colink graphs are very regular and have the density one would expect. Part (i) of Lemma 5.19 was proved by Frankl and Rödl [7] (see also Nagle and Rödl [19]), part (ii) is due to Dementieva et al. [5].

Lemma 5.19 *For all $\alpha_*, \xi_* > 0$ there exists $\delta_* > 0$ such that for all positive integers ℓ, there are integers m_0, r and a constant $\varepsilon > 0$ such that the following holds. Suppose that $\mathcal{H}$ is a 3-uniform hypergraph and V_1, V_2, V_3 are disjoint subsets of vertices of $\mathcal{H}$ such that $|V_1| = |V_2| = |V_3| =: m_* \geq m_0$. Moreover, suppose that $P = P^{12} \cup P^{23} \cup P^{13}$ is a 3-partite graph with vertex classes V_1, V_2 and V_3 which satisfies the following properties:*

- *P is formed by $(1/\ell, \sqrt{\varepsilon})$-regular bipartite graphs,*

- *P is (δ_*, r)-regular with respect to $\mathcal{H}$,*
- *$d_{\mathcal{H}}(P) =: \alpha \geq \alpha_*$.*

Then the following holds for each $s = 1, 2, 3$:

(i) *for all but at most $\xi_* m_*$ vertices $x \in V_s$, the link graph L_x is $(\alpha/\ell, \xi_*)$-regular,*

(ii) *for all but at most $\xi_* m_*^2$ pairs of vertices $x, x' \in V_s$, the colink graph $L_{xx'}$ is $(\alpha^2/\ell, \xi_*)$-regular.*

In [5] the authors stated their result with $\alpha = \alpha_*$, but their proof shows that the constants δ_* etc. guaranteed for α_* and ξ_* also work for any $\alpha \geq \alpha_*$. Lemma 5.19(ii) is the reason why we needed to introduce the notion of (δ_*, r)-regularity of hypergraphs instead of just the simpler notion of δ_*-regularity (which corresponds to $(\delta_*, 1)$-regularity), i.e. we only needed the (δ_*, r)-regularity in order to ensure that (ii) holds .

Remark 5.20 *Definitions* 5.15 *and* 5.17 *together with inequalities* (1)–(3) *imply that we may (and will) assume that each hyperedge $V_iV_jV_k$ of $\mathcal{R}$ satisfies* (i) *and* (ii) *of Lemma* 5.19 *where ξ^2 plays the role of ξ_* and V_i, V_j, V_k play the role of V_1, V_2, V_3.*

6 Finding an almost perfect cover of $\mathcal{R}$

We define $\mathcal{H}_8$ to be the hypergraph on 8 vertices a_1, a_2, a_3, b_1, b_2, b_3, c_1 and c_2 whose hyperedges are all the triples of the form $a_i b_i c_j$ ($i = 1, 2, 3$, $j = 1, 2$).

In this section, we will show that the reduced hypergraph $\mathcal{R}$ of the original hypergraph $\mathcal{H}$ contains an almost perfect $\mathcal{H}_8$-packing (i.e. a collection of disjoint copies of $\mathcal{H}_8$ which covers almost all vertices of $\mathcal{R}$). This will follow immediately from the following two more general propositions. The first one asserts that the minimum degree of an arbitrary 3-uniform hypergraph $\mathcal{H}$ is 'almost inherited' by its reduced hypergraph $\mathcal{R}$ in the sense that the neighbourhood of most vertex pairs in $\mathcal{R}$ is not significantly smaller than $\frac{\delta(\mathcal{H})}{|\mathcal{H}|}|\mathcal{R}|$ (there may be a small proportion of vertex pairs whose neighbourhood can be arbitrarily small – this is a difference to the graph case). The second proposition shows that any hypergraph where most of the pairs of vertices are adjacent to at least a quarter of the other vertices has an almost perfect $\mathcal{H}_8$-packing (so in particular this applies to the reduced hypergraph $\mathcal{R}$ of our given hypergraph $\mathcal{H}$).

Proposition 6.1 *Let $\mathcal{H}$ be a* 3*-uniform hypergraph and let $c \geq \sigma$ be positive constants such that σ satisfies* (3)*. Suppose that $d_{\mathcal{H}}(x, y) \geq cn$ for all pairs of vertices $x, y \in \mathcal{H}$. Let $\mathcal{R}$ be the reduced hypergraph obtained by an application of Lemma* 5.8 *with constants satisfying* (1)*–*(3)*. Then $d_{\mathcal{R}}(V_i, V_j) \geq (c - \sigma)t$ for all useful pairs V_iV_j of clusters. In particular, all but at most $18\delta_* t^2/\varepsilon_3^2$ pairs V_iV_j of clusters satisfy $d_{\mathcal{R}}(V_i, V_j) \geq (c - \sigma)t$.*

Proof. Given a useful pair V_iV_j, let $Z_{\mathcal{R}}$ be the set of all clusters V_k such that the triple $V_iV_jV_k$ is not useful. Thus $Z_{\mathcal{R}} := V(\mathcal{R}) \setminus N_{\mathcal{R}}(V_i, V_j)$. Let $Z_{\mathcal{H}} \subseteq V(\mathcal{H})$ be the union of all clusters belonging to $Z_{\mathcal{R}}$. We will prove the following claim.

There exist vertices $x \in V_i$ and $y \in V_j$ for which there are at most $\sigma n/2$ vertices $z \in Z_{\mathcal{H}}$ such that $xyz \in \mathcal{H}$. $(*)$

Note that $(*)$ implies the proposition since

$$\begin{aligned} cn \le d_{\mathcal{H}}(x,y) &\le |N_{\mathcal{H}}(x,y) \cap (V_0 \cup V_i \cup V_j)| + |N_{\mathcal{H}}(x,y) \cap Z_{\mathcal{H}}| + d_{\mathcal{R}}(V_i,V_j) \cdot m_* \\ &\le 3n/t + \sigma n/2 + d_{\mathcal{R}}(V_i,V_j) \cdot m_* \overset{(3)}{\le} \sigma n + d_{\mathcal{R}}(V_i,V_j) \cdot m_* \end{aligned}$$

and thus $d_{\mathcal{R}}(V_i,V_j) \ge (c-\sigma)n/m_* \ge (c-\sigma)t$.

Thus it remains to show $(*)$. Suppose that $(*)$ is not true and let h be the number of hyperedges of $\mathcal{H}$ having one vertex in each of V_i, V_j and $Z_{\mathcal{H}}$. Thus

$$h \ge m_*^2 \cdot \sigma n/2. \tag{5}$$

Let h_1 be the number of all those hyperedges $xyz \in \mathcal{H}$ with $x \in V_i$, $y \in V_j$, $z \in V_k \in Z_{\mathcal{R}}$ for which $V_iV_jV_k$ is a bad triple. Similarly, let h_2 be the number those hyperedges $xyz \in \mathcal{H}$ with $x \in V_i$, $y \in V_j$, $z \in V_k \in Z_{\mathcal{R}}$ for which $V_iV_jV_k$ is a good triple but not useful. Thus $h = h_1 + h_2$.

Since V_iV_j is a useful pair, there are at most $\varepsilon_3 t$ clusters $V_k \in Z_{\mathcal{R}}$ for which $V_iV_jV_k$ is a bad triple (Definition 5.13). Thus

$$h_1 \le \varepsilon_3 t m_*^3 \le \varepsilon_3 m_*^2 n. \tag{6}$$

We will now give an upper bound for h_2. To do this, let $V_k \in Z_{\mathcal{R}}$ be any cluster such that $V_iV_jV_k$ is good but not useful. Since $V_iV_jV_k$ is good, almost all triads belonging to $V_iV_jV_k$ are (δ_*, r)-regular with respect to $\mathcal{H}$ and are formed by $(1/\ell, \sqrt{\varepsilon_2})$-regular bipartite graphs. Since $V_iV_jV_k$ is not useful, this means that all these triads have only a very small density with respect to $\mathcal{H}$ and thus they induce only a small number of hyperedges of $\mathcal{H}$. Hence h_2 cannot be too large.

More precisely, we can argue as follows. Since $V_iV_jV_k$ is good, at most $3\varepsilon_1 m_*^3$ hyperedges of $\mathcal{H}$ lie in a triad including one of P_0^{ij}, P_0^{jk} or P_0^{ik}. Thus let us now count the number of hyperedges of $\mathcal{H}$ which lie in some triad P which does not include one of P_0^{ij}, P_0^{jk} or P_0^{ik}. Note that, since each of the pairs V_iV_j, V_jV_k, V_iV_k is good, each such P is formed by bipartite graphs of density at most $1/\ell + \varepsilon_2$. Moreover, at most $3 \cdot \varepsilon_3\ell/6 = \varepsilon_3\ell/2$ of the bipartite graphs P_α^{ij}, P_β^{jk}, P_γ^{ik} are not $(1/\ell, \sqrt{\varepsilon_2})$-regular (cf. Definition 5.9). Thus altogether, these bipartite graphs contain at most $(1/\ell + \varepsilon_2)m_*^2 \cdot \varepsilon_3\ell/2$ edges. Since each such edge lies in at most m_* triangles, this gives room for at most $\varepsilon_3 m_*^3$ hyperedges of $\mathcal{H}$.

All the remaining hyperedges of $\mathcal{H}$ having one vertex in each of V_i, V_j and V_k lie in some triad P which is formed by $(1/\ell, \sqrt{\varepsilon_2})$-regular graphs. If such a triad P is also (δ_*, r)-regular with respect to $\mathcal{H}$ then, by Proposition 5.3, P contains at most

$$d_{\mathcal{H}}(P) \cdot t(P) \le d_{\mathcal{H}}(P)(1 + \varepsilon_2^{1/4})^4 m_*^3/\ell^3 \le d_{\mathcal{H}}(P) \cdot \frac{2m_*^3}{\ell^3} \le \frac{2\alpha_* m_*^3}{\ell^3}$$

hyperedges of $\mathcal{H}$. (Here $d_{\mathcal{H}}(P) \le \alpha_*$ since P cannot be useful.) Thus at most $\ell^3 \cdot 2\alpha_* m_*^3/\ell^3 = 2\alpha_* m_*^3$ hyperedges of $\mathcal{H}$ lie in such triads P.

Thus it remains to bound the number of hyperedges of $\mathcal{H}$ having one vertex in each of V_i, V_j and V_k which lie in some triad P which is not (δ_*, r)-regular with respect to $\mathcal{H}$ but which is formed by $(1/\ell, \sqrt{\varepsilon_2})$-regular graphs. Since $V_iV_jV_k$ is good, there are at most $\varepsilon_3\ell^3$ such triads and by Proposition 5.3 each contains at most $(1+\varepsilon_2^{1/4})^4 m_*^3/\ell^3$ triangles. So this gives room for at most $2\varepsilon_3 m_*^3$ hyperedges of $\mathcal{H}$. Thus we have shown that there are at most

$$3\varepsilon_1 m_*^3 + \varepsilon_3 m_*^3 + 2\alpha_* m_*^3 + 2\varepsilon_3 m_*^3 \leq 3\alpha_* m_*^3$$

hyperedges $xyz \in \mathcal{H}$ with $x \in V_i$, $y \in V_j$ and $z \in V_k$. Summing over all clusters $V_k \in Z_{\mathcal{R}}$ for which $V_iV_jV_k$ is good but not useful implies that

$$h_2 \leq 3\alpha_* m_*^3 t \leq 3\alpha_* m_*^2 n.$$

Together with (5) and (6) this gives

$$\sigma m_*^2 n/2 \leq h = h_1 + h_2 \leq \varepsilon_3 m_*^2 n + 3\alpha_* m_*^2 n,$$

a contradiction since $\sigma \gg \alpha_* \gg \varepsilon_3$ by (3). □

Proposition 6.2 *For every positive ν' there exists an integer $n_0 = n_0(\nu')$ such that every 3-uniform hypergraph $\mathcal{G}$ of order $n \geq n_0$ for which $d_{\mathcal{G}}(x,y) \geq n/4$ for all but at most $\nu' n^2$ pairs of vertices $x, y \in \mathcal{G}$ contains an $\mathcal{H}_8$-packing which covers all but at most $4\sqrt{\nu'}n$ vertices of $\mathcal{G}$.*

Proof. Set $\overline{\nu} := 4\sqrt{\nu'}$. Let $\mathcal{M}$ be an $\mathcal{H}_8$-packing in $\mathcal{G}$ of maximum size. Denote the copies of $\mathcal{H}_8$ in $\mathcal{M}$ by $H_1, H_2, \ldots$. Let X be the set of uncovered vertices. Suppose that $|X| \geq \overline{\nu}n$. We will prove the following claim.

There is a set F of $\overline{\nu}n/20$ disjoint pairs of vertices in X such that, firstly, each of these pairs has at most $\overline{\nu}n/8$ of its neighbours in X and, secondly, each of these pairs xy satisfies $d_{\mathcal{G}}(x,y) \geq n/4$. $(**)$

Before we prove $(**)$, let us show how it implies the proposition. So let F be as in $(**)$. We say that $H_i \in \mathcal{M}$ *is good for a pair $xy \in F$* if xy forms a hyperedge with at least three vertices of H_i. Let n_{xy} be the number of H_i which are good for xy. Then

$$n/4 - \overline{\nu}n/8 \leq d_{\mathcal{G}}(x,y) - \overline{\nu}n/8 \leq 2(|\mathcal{M}| - n_{xy}) + 8n_{xy} \leq 2((1-\overline{\nu})n/8 - n_{xy}) + 8n_{xy}.$$

Thus $n_{xy} \geq \overline{\nu}n/48 =: n^*$. Now we double-count all those triples (H_i, H_j, xy) for which $i < j$, $xy \in F$ and for which both H_i and H_j are good for xy. On the one hand, the number of such triples is at least $|F|\binom{n^*}{2} \geq \overline{\nu}^3 n^3/10^5$. On the other hand, let a_{ij} be the number of pairs $xy \in F$ such that both H_i and H_j are good for xy. Put $a := \max_{i<j} a_{ij}$. Then the number of triples (H_i, H_j, xy) as above is (crudely) at most an^2. By combining these inequalities, we have $a \geq \overline{\nu}^3 n/10^5$. So it follows that there are H_i, H_j and $\overline{\nu}^3 n/10^5$ pairs in F such that both H_i and H_j are good for all these pairs. This in turn implies that there must be vertices v_1, v_2 and v_3 in H_i, vertices w_1, w_2 and w_3 in H_j and

pairs $x_1y_1, \dots, x_9y_9 \in F$ such that v_k as well as w_k forms a hyperedge with each $x_{k'}y_{k'}$ ($k = 1, 2, 3$, $k' = 1, \dots, 9$).

This enables us to find an $\mathcal{H}_8$-packing $\mathcal{M}'$ which is larger than the original $\mathcal{M}$ as follows: First remove H_i and H_j from $\mathcal{M}$. Then form three vertex disjoint copies of $\mathcal{H}_8$, where the kth copy contains the vertices v_k and w_k as well as three of the pairs $x_{k'}y_{k'}$. These copies are disjoint from the current $\mathcal{H}_8$-packing. Thus they can be added to form a new $\mathcal{H}_8$-packing $\mathcal{M}'$ which contains one more copy of $\mathcal{H}_8$, a contradiction.

Thus it remains to prove $(**)$. Choose a partition of X into X_1 and X_2 uniformly at random from all partitions with $|X_1| = 7|X_2|$ (by removing up to 7 vertices from X if necessary we may assume that $|X|$ is divisible by 8). By Lemma 4.1 we may assume that

$$|N_{\mathcal{G}}(x, y) \cap X_2| \geq |N_{\mathcal{G}}(x, y) \cap X|/10 \tag{7}$$

for all those pairs x, y of vertices of $\mathcal{G}$ for which $d_{\mathcal{G}}(x, y) \geq n/4$. Now greedily choose disjoint pairs $x_1y_1, \dots, x_sy_s$ in X_1 such that $d_{\mathcal{G}}(x_i, y_i) \geq n/4$ for all i. It is easy to see that we can choose $s := |X_2| \geq \overline{\nu}n/10$ such pairs, as altogether there are at most $\nu' n^2$ pairs whose degree is smaller. Clearly, $(**)$ holds if at least half of the pairs x_iy_i have at most $\overline{\nu}n/8$ of their neighbours in X. (In this case we can take F to be the set of these pairs). So suppose that at least half of the pairs x_iy_i have at least $\overline{\nu}n/8$ of their neighbours in X. Define an auxiliary bipartite graph B as follows. The vertex classes of B are X_2 and the set A of all pairs $x_1y_1, \dots, x_sy_s$. A pair x_iy_i is joined to $x \in X_2$ by an edge in B if x_iy_ix forms a hyperedge of $\mathcal{G}$. Note that the vertex classes of B are of equal size. Also, our assumption and inequality (7) together imply that at least half the vertices in A have at least $\overline{\nu}n/80 \geq \overline{\nu}|X_2|/80$ neighbours in X_2. If n is sufficiently large this in turn implies that B contains a complete bipartite graph $K_{2,3}$ where the vertex class of size two is contained in X_2 (see e.g. Chapter VI, Lemma 2.1. in [3]). But this $K_{2,3}$ corresponds to a copy of $\mathcal{H}_8$ in $\mathcal{G}$ that is disjoint from $\mathcal{M}$, which contradicts the assumption that $\mathcal{M}$ is of maximum size and thus proves $(**)$. $\square$

7 Almost covering $\mathcal{H}$ by triples and tidying them up

7.1 Almost covering $\mathcal{H}$ by triples

In this subsection, we will choose a cover of almost all vertices of $\mathcal{H}$ by boundedly many disjoint triples $T_k = (X_k, Y_k, Z_k)$ of vertex sets such that $|X_k| = |Y_k| = |Z_k|/2$ for every k. To each T_k there will be assigned a triad P_k such that P_k is highly regular and sufficiently dense with respect to $\mathcal{H}$ and such that the 3 bipartite graphs forming P_k are highly regular too. Roughly speaking, our aim later on is to find a loose Hamilton path in each of these triples and to glue all these paths together into a loose Hamilton cycle of $\mathcal{H}$. However, before we can do this, we need to 'tidy up' the T_k by deleting a few carefully chosen vertices in Section 7.2. The triples T_k are obtained as follows. Recall from Section 5.3

that

$$\varepsilon' := 4\varepsilon_2^{1/4},\ d := 1/\ell \text{ and } \delta := \sqrt{\delta_*}.$$

Apply Propositions 6.1 and 6.2 to the reduced graph $\mathcal{R}$ to obtain an $\mathcal{H}_8$-packing $\mathcal{M}$ which covers all but at most $20\sqrt{\delta_*}t/\varepsilon_3$ vertices of $\mathcal{R}$. Consider any element $\mathcal{H}' \in \mathcal{M}$. Denote the clusters of $\mathcal{H}$ which correspond to the vertices of $\mathcal{H}'$ by A_i, B_i and C_j in such a way that the hyperedges of $\mathcal{H}'$ have the form $A_iB_iC_j$, where $1 \leq i \leq 3$ and $1 \leq j \leq 2$. For each hyperedge $A_iB_iC_j$ of $\mathcal{H}'$ let $P(A_iB_iC_j)$ denote the useful triad which we fixed for $A_iB_iC_j$ after Definition 5.17. Recall that by Definition 5.15(i), this triad is formed by $(d, \sqrt{\varepsilon_2})$-regular bipartite graphs. Proposition 5.2 implies that by removing $4\sqrt{\varepsilon_2}|\mathcal{H}| = 4\sqrt{\varepsilon_2}n$ vertices from the clusters and adding these vertices to the exceptional set V_0 we may assume that each $P(A_iB_iC_j)$ is even formed by $(d, \varepsilon'/4)$-superregular bipartite graphs. Moreover, by putting at most $2t$ further vertices into V_0 if necessary, we may assume that the number of vertices in each of the clusters is divisible by three (and it is the same for all the clusters). We denote this number by m. We also put all those vertices of $\mathcal{H}$ into V_0 which do not lie in a cluster which is contained in an element of $\mathcal{M}$. The Hypergraph Regularity Lemma 5.8 implied that we had $|V_0| \leq t$ originally, so now we have

$$|V_0| \leq t + 20\sqrt{\delta_*}n/\varepsilon_3 + 4\sqrt{\varepsilon_2}n + 2t \overset{(3)}{\leq} \xi n. \tag{8}$$

Now partition both C_1 and C_2 randomly into three sets of equal size and call the resulting sets $Y_1, \dots, Y_6$. Next partition each of the A_i randomly into two parts X_i and Z_{i+3} and partition B_i randomly into two parts X_{i+3} and Z_i such that

$$2m/3 = |Z_i| = 2|X_i| = 2|Y_i|$$

for all i with $1 \leq i \leq 6$.

Altogether this gives us six triples $T_k = (X_k, Y_k, Z_k)$ of vertex sets. Note that for each T_k there is a (unique) hyperedge $A_iB_iC_j \in E(\mathcal{R})$ which corresponds to this triple, i.e. such that either $X_k \subseteq A_i$, $Y_k \subseteq C_j$ and $Z_k \subseteq B_i$ or else $X_k \subseteq B_i$, $Y_k \subseteq C_j$ and $Z_k \subseteq A_i$. Let P_k denote the subtriad of $P(A_iB_iC_j)$ induced by $X_k \cup Y_k \cup Z_k$. We proceed similarly for every element of the $\mathcal{H}_8$-packing $\mathcal{M}$ to obtain

$$N := 6|\mathcal{M}| \leq |\mathcal{R}| \tag{9}$$

such triples $T_k = (X_k, Y_k, Z_k)$.

In the Section 8, the following simple fact will be helpful when incorporating the exceptional set V_0.

Proposition 7.1 *For every pair of vertices $v, w \in V(\mathcal{H})$ there are at least $\sigma N/8$ indices k for which $|N_{\mathcal{H}}(v,w) \cap Z_k| \geq \sigma|Z_k|/8$.*

Proof. Recall that m_* denotes the size of the original clusters of $\mathcal{R}$. Let N' denote the number of elements $\mathcal{H}'$ in the $\mathcal{H}_8$-packing $\mathcal{M}$ which contain at least 3 original clusters whose intersection with $N_{\mathcal{H}}(v,w)$ is least $\sigma m_*/6$. Then

$$n/4 + \sigma n \leq |N_{\mathcal{H}}(v,w)| \leq 8m_*N' + |\mathcal{M}|(2m_* + 6\sigma m_*/6).$$

Since $n \geq 8|\mathcal{M}|m_*$ this implies that $N' \geq 7\sigma|\mathcal{M}|/8$.

Consider any $\mathcal{H}' \in \mathcal{M}$ which contains at least 3 (original) clusters whose intersection with $N_{\mathcal{H}}(v,w)$ has size at least $\sigma m_*/6$. As at the beginning of this section, denote the clusters of $\mathcal{H}'$ which correspond to the vertices of $\mathcal{H}'$ by A_i, B_i and C_j, where $1 \leq i \leq 3$ and $1 \leq j \leq 2$. Then either the intersection of $N_{\mathcal{H}}(v,w)$ with some A_i has size at least $\sigma m_*/6$ or the intersection of $N_{\mathcal{H}}(v,w)$ with some B_i has size at least $\sigma m_*/6$. But both A_i and B_i contain some set of the form Z_k. Lemma 4.1 (applied with $\mathcal{A}$ being the set of neighbourhoods of vertex pairs in $\mathcal{H}$) implies that we may assume that for the pair v, w the intersection of $N_{\mathcal{H}}(v,w)$ with this Z_k has size at least $\sigma|Z_k|/8$. Thus we have found at least $N' \geq 3\sigma|\mathcal{M}|/4 = \sigma N/8$ sets Z_k, each containing at least $\sigma|Z_k|/8$ neighbours of the pair v, w, as required.

□

7.2 Tidying up the triples

In the remainder of this section we will consider any (sub-)triad P_k together with its vertex sets X_k, Y_k and Z_k. For simplicity, we will call this triad P and its vertex sets X, Y and Z. We denote the bipartite subgraphs forming P by P_{XY}, P_{YZ} and P_{XZ}. Recall that $T(P)$ denotes the set of triangles in P.

Definition 7.2 ($\mathcal{H}_P$) We write $\mathcal{H}_P$ for the 3-partite subhypergraph of $\mathcal{H}$ whose vertex set is $X \cup Y \cup Z$ and whose set of hyperedges is $E(\mathcal{H}) \cap T(P)$.

The main aim of this section is to find sets $X' \subseteq X$, $Y' \subseteq Y$ and $Z' \subseteq Z$ and a subgraph $P^{rich}_{X'Y'}$ of P_{XY} with vertex classes X' and Y' satisfying the following (see Proposition 7.16): For all $z \in Z'$ the link graph in $P^{rich}_{X'Y'}$ (which consists of those edges of $P^{rich}_{X'Y'}$ forming a hyperedge with z) is regular (the exact parameters are defined below) and all edges of $P^{rich}_{X'Y'}$ are rich in the sense that they form hyperedges with a significant proportion of the vertices in Z'. This will be essential in the final section where we define a bipartite auxiliary graph H^* whose vertex classes consist of a random subset of Z' (together with a few other vertices) and a cycle R^* in $P^{rich}_{X'Y'}$ and which has an edge between a vertex in Z' and an edge of the cycle R^* if together they form a hyperedge of $\mathcal{H}_P$. The above properties will be used to show that this auxiliary graph has large minimum degree and that it contains a perfect matching.

The main technical difficulty in proving Proposition 7.16 is that when deleting all those edges between X and Y which are not rich (we will call them poor) we also have to delete some vertices in X and Y to ensure that the subgraph of P_{XY} thus obtained is again superregular. But the deletion of these vertices may destroy the regularity of the link graphs of some further vertices in Z. (Roughly, these link graphs will be the ones which we call impure later on). When we delete the latter vertices in Z some of the rich edges in P_{XY} may now turn into poor ones. However, we can show that this does not turn into an iterative process if we choose Z' carefully.

Recall that the (co-)link graphs were introduced in Definition 5.18. Given a vertex $z \in Z$, we will still write L_z for the intersection of the original link graph of z with P_{XY}. We will now call this L_z the link graph of z. We denote by $X_z \subseteq X$ and $Y_z \subseteq Y$ the vertex classes of L_z. Thus $X_z = N_{P_{XZ}}(z)$ and $Y_z = N_{P_{YZ}}(z)$. The vertex classes of $L_{zz'}$ are denoted by $X_{zz'}$ and $Y_{zz'}$ respectively. We proceed similarly for the (co-)link graphs of vertices in Y and Z.

Definition 7.3 (typical link and colink graphs) The *link graph* L_z *of* $z \in Z$ *is typical* if it is $(\alpha d, \xi)$-regular, $(1-\varepsilon')d|X| \leq |X_z| \leq (1+\varepsilon')d|X|$ and $(1-\varepsilon')d|Y| \leq |Y_z| \leq (1+\varepsilon')d|Y|$. Given distinct $z, z' \in Z$, we say that the *colink graph* $L_{zz'}$ *is typical* if it is $(\alpha^2 d, \xi)$-regular, $(1-\varepsilon')^2 d^2|X| \leq |X_{zz'}| \leq (1+\varepsilon')^2 d^2|X|$ and $(1-\varepsilon')^2 d^2|Y| \leq |Y_{zz'}| \leq (1+\varepsilon')^2 d^2|Y|$. We adopt analogous definitions for link graphs and colink graphs of vertices in X and Y.

Using Remark 5.20, the fact that P is a subtriad of some $(d, \varepsilon'/4)$-superregular triad of the form $P(A_i B_i C_j)$, the definition of a superregular graph, Definition 5.5 and the fact that X, Y and Z were obtained by considering random partitions (c.f. Lemma 4.1), it is not hard to see that we may assume that P satisfies the following properties:

(P1) Each of P_{XY}, P_{YZ} and P_{XZ} is a (d, ε')-superregular graph.

(P2) P is (δ, r)-regular with respect to $\mathcal{H}$ and $d_{\mathcal{H}}(P) =: \alpha \geq \alpha_*/2$.

(P3) All but at most $\xi|Z|$ vertices $z \in Z$ have a typical link graph and the analogous statements are true for X and Y.

(P4) All but at most $\xi|Z|^2$ pairs of vertices $z, z' \in Z$ have a typical colink graph and the analogous statements are true for X and Y.

Note that α in (P2) depends on the triple (X, Y, Z).

Definition 7.4 (poor edges of P_{XY}) An edge of P_{XY} is *poor* if it lies in at most $(1-\eta^3)\alpha d^2|Z|$ hyperedges of $\mathcal{H}_P$. We let F_{poor} denote the set of poor edges in P_{XY}.

Recall that η was fixed in (4). We will now show that only a few edges can be poor. To do this, we first need the following observation.

Proposition 7.5 *Let* $F \subseteq E(P_{XY})$ *be a set of at least* $\beta d|X||Y|$ *edges where* $1 \gg \beta, d \gg \varepsilon' \geq 0$. *Then at least* $\beta(1-\beta^2)t(P)$ *triangles in* P *contain an edge from* F.

Proof. We say that a vertex $x \in X$ is *regular* if it is incident to at least $\varepsilon'|Y|$ edges of F. By (P1) we have $|N_{P_{XZ}}(x)| \geq (1-\varepsilon')d|Z|$ for every $x \in X$, which together with (P1) in turn implies that every regular vertex lies in at least $(1-\varepsilon')^2 d^2 d_F(x)|Z|$ triangles containing an edge of F. Thus every vertex x lies in at least $(1-\varepsilon')^2 d^2(d_F(x) - \varepsilon'|Y|)|Z|$ triangles containing an edge of F. Thus the number of triangles in P which contain an edge from F is at least $(1-\varepsilon')^2 d^2(\beta d - \varepsilon')|X||Y||Z|$. Together with Proposition 5.3 this gives the desired bound. □

Proposition 7.6 *The number of poor edges is less than $2\delta d|X||Y|$.*

Proof. Suppose that there are at least $2\delta d|X||Y|$ such edges. Consider the subgraph Q of P which consists of P_{XZ}, P_{YZ} together with any $P'_{XY} \subseteq P_{XY}$ consisting of $2\delta d|X||Y|$ of the poor edges. Propositions 5.3 and 7.5 together imply that $t(Q) \geq 2\delta(1-4\delta^2)t(P) \geq 2\delta(1-\delta)d^3|X||Y||Z|$. The first inequality implies that we may make use of the δ-regularity of P with respect to $\mathcal{H}$ (c.f. (P2)), which means that we have $d_{\mathcal{H}}(Q) \geq d_{\mathcal{H}}(P) - \delta \geq \alpha - \delta$. But on the other hand, we have

$$d_{\mathcal{H}}(Q) = \frac{|E(\mathcal{H}) \cap T(Q)|}{t(Q)} \leq \frac{|E(\mathcal{H}) \cap T(Q)|}{2\delta(1-\delta)d^3|X||Y||Z|}.$$

Thus

$$\begin{aligned} |E(\mathcal{H}) \cap T(Q)| &\geq 2\delta(1-\delta)d^3|X||Y||Z|(\alpha-\delta) \\ &\geq (1-\delta-\delta/\alpha)2\delta\alpha d^3|X||Y||Z| \\ &\overset{(4)}{\geq} (1-\eta^3/2)e(P'_{XY})\alpha d^2|Z|. \end{aligned}$$

So one of the edges of P'_{XY} lies in at least $(1-\eta^3/2)\alpha d^2|Z|$ hyperedges of $\mathcal{H}$, a contradiction to the definition of a poor edge. $\square$

Definition 7.7 (poor link graph) We say that the link graph L_z of a vertex $z \in Z$ is *poor* if at least $\alpha\xi^3 d^3|X||Y|$ of its edges are poor.

The next two results show that firstly any typical link graph which is not poor remains sufficiently regular if we delete the poor edges and secondly that only a few typical link graphs are poor. Note that this is not unexpected in view of the fact that the number of edges of L_z is typically about $\alpha d^3|X||Y|$.

Proposition 7.8 *Suppose that the link graph L_z of a vertex $z \in Z$ is typical and not poor. Then the subgraph of L_z obtained by deleting the poor edges is $(\alpha d, 2\xi)$-regular.*

Proof. Recall that X_z and Y_z denote the vertex classes of L_z. Let L_z^* denote the subgraph obtained from L_z by deleting the poor edges. Thus we have to show that L_z^* is $(\alpha d, 2\xi)$-regular. So consider sets $X_z^* \subseteq X_z$ and $Y_z^* \subseteq Y_z$ with $|X_z^*| \geq 2\xi|X_z|$ and $|Y_z^*| \geq 2\xi|Y_z|$. Since L_z is typical, L_z is $(\alpha d, \xi)$-regular, $|X_z| \geq (1-\varepsilon')d|X|$ and $|Y_z| \geq (1-\varepsilon')d|Y|$. Together with Definition 7.7 this implies that the number of edges in the subgraph of L_z^* induced by X_z^* and Y_z^* is at least

$$\begin{aligned} \alpha d(1-\xi)|X_z^*||Y_z^*| - \alpha\xi^3 d^3|X||Y| &\geq \alpha d(1-\xi)|X_z^*||Y_z^*| - \frac{\alpha\xi^3 d^3|X_z^*||Y_z^*|}{(1-\varepsilon')^2 d^2 \cdot 4\xi^2} \\ &\geq \alpha d(1-2\xi)|X_z^*||Y_z^*|, \end{aligned}$$

and the result follows. $\square$

Proposition 7.9 *At most $2\delta|Z|/\xi^3$ link graphs are poor.*

Proof. By definition, every poor edge lies in at most $\alpha d^2|Z|$ hyperedges of $\mathcal{H}_P$. In other words, it lies in at most $\alpha d^2|Z|$ link graphs. Thus the number of pairs (e, L_z) such that e is a poor edge in L_z is at most $\alpha d^2|Z||F_{poor}|$, where F_{poor} was the set of poor edges. By Proposition 7.6, this is at most $2\delta\alpha d^3|X||Y||Z|$. On the other hand, the number of pairs (e, L_z) as above is clearly at least $\alpha\xi^3 d^3|X||Y|\#_{poor}$, where $\#_{poor}$ denotes the number of poor link graphs. Thus $\#_{poor} \le 2\delta|Z|/\xi^3$, as required. □

Recall from Section 5.3 that

$$\delta_0 := \delta^{1/4}.$$

Proposition 7.10 *The subgraph of P_{XY} which is obtained by deleting the set F_{poor} of poor edges is (d, δ_0)-regular. In particular, it can be made into a $(d, 4\delta_0)$-superregular graph by deleting a set $\widetilde{X} \subseteq X$ of at most $3\delta_0|X|$ vertices in X and a set $\widetilde{Y} \subseteq Y$ of at most $3\delta_0|Y|$ vertices in Y.*

Proof. Consider sets $U \subseteq X$ and $W \subseteq Y$ with $|U| \ge \delta_0|X|$ and $|W| \ge \delta_0|Y|$. Let P'_{XY} be the subgraph obtained from P_{XY} by deleting all the poor edges. Since P_{XY} is (d, ε')-superregular by (P1) and since Proposition 7.6 stated that $|F_{poor}| \le 2\delta d|X||Y|$, it follows that

$$\begin{aligned} e_{P'_{XY}}(U, W) &\ge (1-\varepsilon')d|U||W| - 2\delta d|X||Y| \\ &\ge (1-\varepsilon')d|U||W| - 2\delta d|U||W|/\delta_0^2 \ge (1-\delta_0)d|U||W|. \end{aligned}$$

The corresponding upper bound $e_{P'_{XY}}(U, W) \le (1+\delta_0)|U||W|$ follows immediately from the superregularity of P_{XY}. Thus P'_{XY} is (d, δ_0)-regular.

It remains to show that P'_{XY} can be made $(d, 4\delta_0)$-superregular by deleting a small fraction of vertices. Note that since $\delta_0 \gg d$ by (4) we cannot apply Proposition 5.2 to achieve this. Let $X^* \subseteq X$ denote the set of all those vertices which in the graph P'_{XY} have either less than $(1-\delta_0)d|Y|$ neighbours in Y or more than $(1+\delta_0)d|Y|$ neighbours in Y. Define $Y^* \subseteq Y$ similarly. Since P'_{XY} is (d, δ_0)-regular we have $|X^*| \le 2\delta_0|X|$ and $|Y^*| \le 2\delta_0|Y|$. If $|Y^*| \ge \varepsilon'|Y|$, then the (d, ε')-superregularity of P_{XY} implies that all but at most $\varepsilon'|X|$ vertices in X have at most

$$(1+\varepsilon')d|Y^*| \le 2\delta_0(1+\varepsilon')d|Y| \le 5\delta_0 d|Y|/2$$

neighbours in Y^*. Since $5\delta_0 d/2 \ge \varepsilon'$, this also holds if $|Y^*| < \varepsilon'|Y|$. Similarly, all but at most $\varepsilon'|Y|$ vertices in Y have at most $5\delta_0 d|X|/2$ neighbours in X^*. Now delete X^*, Y^* and these two sets of at most $\varepsilon'|X|$ (respectively $\le \varepsilon'|Y|$) vertices. It is easy to check that the resulting subgraph of P_{XY} is $(d, 4\delta_0)$-superregular. □

Recall from Section 5.3 that

$$\delta_1 := 24\delta_0/\delta^{1/5}.$$

Definition 7.11 (impure link graph) Given a typical link graph L_z, we say that L_z is *impure* if at least $\delta_1|X_z|$ vertices in L_z are contained in the set $\widetilde{X}$ defined in Proposition 7.10 or if at least $\delta_1|Y_z|$ vertices in L_z are contained in the set $\widetilde{Y}$.

Proposition 7.12 *At most $\delta^{1/5}|Z|$ typical link graphs are impure.*

Proof. We double count the number of tuples (x, L_z) such that $x \in \widetilde{X} \cap X_z$ and such that L_z is typical. Since P_{XZ} is (d, ε')-superregular by (P1), each vertex $x \in \widetilde{X}$ is contained in at most $2d|Z|$ link graphs. Thus the number of tuples (x, L_z) as above is at most $|\widetilde{X}| \cdot 2d|Z| \leq 6\delta_0 d|X||Z|$. On the other hand, let N denote the number of typical link graphs L_z for which the vertex class X_z contains at least $\delta_1|X_z|$ vertices of $\widetilde{X}$. Note that $|X_z| \geq (1-\varepsilon')d|X|$ since L_z is typical and thus $\delta_1|X_z| \geq \delta_1 d|X|/2$. So the number of tuples (x, L_z) is at least $N \cdot \delta_1 d|X|/2$. Hence $N\delta_1 d|X|/2 \leq 6\delta_0 d|X||Z|$ and therefore $N \leq 12\delta_0|Z|/\delta_1 = \delta^{1/5}|Z|/2$. Argue similarly for Y_z to obtain the desired bound. □

Proposition 7.13 *Suppose that $z \in Z$ is a vertex such that the link graph L_z is typical, not poor and not impure. Then the subgraph obtained from L_z by deleting all its poor edges and its vertices in $\widetilde{X} \cup \widetilde{Y}$ is still $(\alpha d, 3\xi)$-regular and its vertex classes $X_z \setminus \widetilde{X}$ and $Y_z \setminus \widetilde{Y}$ satisfy*

- $(1 - 3\delta_1/2)d|X \setminus \widetilde{X}| \leq |X_z \setminus \widetilde{X}| \leq (1 + 3\delta_1/2)d|X \setminus \widetilde{X}|$ *and*
- $(1 - 3\delta_1/2)d|Y \setminus \widetilde{Y}| \leq |Y_z \setminus \widetilde{Y}| \leq (1 + 3\delta_1/2)d|Y \setminus \widetilde{Y}|$.

Proof. Recall that by Proposition 7.8, the subgraph obtained from L_z by deleting the poor edges is $(\alpha d, 2\xi)$-regular. Since L_z is not impure, only a small fraction of the vertices in each of its vertex classes lies in $\widetilde{X} \cup \widetilde{Y}$, which immediately implies the $(\alpha d, 3\xi)$-regularity. The lower bounds on $|X_z \setminus \widetilde{X}|$ and $|Y_z \setminus \widetilde{Y}|$ follow since L_z is typical and not impure. To see e.g. the upper bound on $|X_z \setminus \widetilde{X}|$ note that

$$|X_z \setminus \widetilde{X}| \leq (1+\varepsilon')d|X| \leq \frac{1+\varepsilon'}{1-3\delta_0} d|X \setminus \widetilde{X}| \leq (1 + 3\delta_1/2)d|X \setminus \widetilde{X}|,$$

as required. (Apply Proposition 7.10 to verify the second inequality.) □

Now we delete all vertices $z \in Z$ for which the link graph L_z is not typical or poor or impure. We add all these vertices to the exceptional set V_0. Let $Z' \subseteq Z$ be the subset of Z thus obtained. Then Propositions 7.9 and 7.12 together with (P3) and (4) imply that

$$|Z \setminus Z'| \leq \xi|Z| + 2\delta|Z|/\xi^3 + \delta^{1/5}|Z| \leq 2\xi|Z|. \tag{10}$$

Definition 7.14 (unhappy edges) Call an edge $xy \in P_{XY}$ *unhappy* if at least one of the following is true:

(U1) xy is not poor but forms a hyperedge in $\mathcal{H}_P$ with at most $(1-\eta^2)\alpha d^2|Z|$ of the vertices in Z'.

(U2) $|N_{P_{XZ}}(x) \cap N_{P_{YZ}}(y)| \geq (1+\varepsilon')^2 d^2 |Z|$ or $|N_{P_{XZ}}(x) \cap N_{P_{YZ}}(y)| \leq (1-\varepsilon')^2 d^2 |Z|$.

Proposition 7.15 *The following properties are satisfied:*

(i) *At most $4\varepsilon'|X||Y|$ edges are unhappy.*

(ii) *The set $\overline{X} \subseteq X$ of all those vertices which are incident to more than $3\varepsilon'|X|$ unhappy edges has size at most $\varepsilon'|X|$. Similarly, the set $\overline{Y} \subseteq Y$ of all those vertices which are incident to more than $3\varepsilon'|Y|$ unhappy edges has size at most $\varepsilon'|Y|$.*

(iii) *There exists sets $\widehat{X} \subseteq X$ and $\widehat{Y} \subseteq Y$ such that $|\widehat{X}| \leq 2\varepsilon'|X|$, $|\widehat{Y}| \leq 2\varepsilon'|Y|$ and such that the subgraph $P_{X'Y'}$ of P_{XY} induced by the sets $X' := X \setminus (\widetilde{X} \cup \overline{X} \cup \widehat{X})$ and $Y' := Y \setminus (\widetilde{Y} \cup \overline{Y} \cup \widehat{Y})$ is $(d, \sqrt{\varepsilon'})$-superregular.*

Proof. Suppose that an edge xy is unhappy because of (U1). Then by definition, in the graph $P_{XZ} \cup P_{YZ}$, x and y must have at least

$$(\eta^2 - \eta^3)\alpha d^2 |Z| \geq \eta^2 \alpha d^2 |Z|/2 \tag{11}$$

common neighbours in $Z \setminus Z' =: Z^*$. On the other hand, the regularity of P_{XZ} implies that all but at most $\varepsilon'|X|$ vertices in X have at most $2d|Z^*|$ neighbours in Z^*. Let X^* be the set of these vertices in X. (Thus $|X^*| \geq (1-\varepsilon')|X|$.) Fix $x \in X^*$. Then as before, the regularity of P_{YZ} implies that all but at most $\varepsilon'|Y|$ vertices in Y have at most $3d^2|Z^*|$ neighbours in $N_{P_{XY}}(x) \cap Z^*$. Let $Y^*(x)$ denote the set of all these vertices in Y. Then no edge xy with $y \in Y^*(x)$ is unhappy because of (U1) since such x and y have at most

$$3d^2|Z^*| \overset{(10)}{\leq} 6\xi d^2 |Z| \overset{(4)}{<} \frac{\eta^2 \alpha d^2 |Z|}{2}$$

common neighbours in Z^*. Together with (11), this shows that each $x \in X^*$ is incident to at most $\varepsilon'|Y|$ edges which are unhappy because of (U1). So altogether, the total number of edges which are unhappy because of (U1) is at most $|X^*| \cdot \varepsilon'|Y| + \varepsilon'|X||Y| \leq 2\varepsilon'|X||Y|$.

Now consider (U2). The superregularity of P_{XZ} and the regularity of P_{YZ} together imply that for all vertices $x \in X$ there are at most $2\varepsilon'|Y|$ edges incident to x which satisfy (U2). Thus there are at most $2\varepsilon'|X||Y|$ edges $xy \in P_{XY}$ which satisfy (U2). Altogether this shows that at most $4\varepsilon'|X||Y|$ edges are unhappy. Moreover, the proof also shows that at most $\varepsilon'|X|$ vertices in X are incident to more than $3\varepsilon'|Y|$ unhappy edges (indeed, this can only happen for the vertices $x \notin X^*$). A similar argument gives the analogous statement for Y. Thus we have proved (i) and (ii).

To prove (iii), let us first define $\widehat{X}$ and $\widehat{Y}$. Recall that $\widetilde{X}$ and $\widetilde{Y}$ were defined in Proposition 7.10. If $|\widetilde{Y} \cup \overline{Y}| \leq \varepsilon'|Y|$, we simply set $\widehat{X} = \emptyset$. If this is not the case, then since P_{XY} is (d, ε')-superregular, at most $2\varepsilon'|X|$ vertices in X have either less than $(1-\varepsilon')d|\widetilde{Y} \cup \overline{Y}|$ neighbours in $\widetilde{Y} \cup \overline{Y}$ or else more than $(1+\varepsilon')d|\widetilde{Y} \cup \overline{Y}|$ neighbours in $\widetilde{Y} \cup \overline{Y}$. Let $\widehat{X}$ denote the set of all these at most $2\varepsilon'|X|$ vertices in X. Define $\widehat{Y}$ similarly. It is now easy to check that $P_{X'Y'}$ is $(d, \sqrt{\varepsilon'})$-superregular. □

For later reference, we now summarize the properties of those graphs which we will need later on.

Proposition 7.16 *Let X', Y' and Z' be as defined above. Then the following holds.*

(i) $|X \setminus X'| \le \delta_1 |X|$, $|Y \setminus Y'| \le \delta_1 |Y|$ *and* $|Z \setminus Z'| \le 2\xi |Z|$,

(ii) $P_{X'Y'}$ *is* $(d, \sqrt{\varepsilon'})$*-superregular.*

(iii) *For every vertex $z \in Z'$ the link graph L_z is typical.*

(iv) *Given a vertex $z \in Z'$, let L_z^{rich} denote the graph obtained from $L_z[X' \cup Y']$ by deleting all the poor and the unhappy edges. Let X'_z and Y'_z denote the vertex classes of L_z^{rich}. Then for every vertex $z \in Z'$ the graph L_z^{rich} is $(\alpha d, 4\xi)$-regular, $(1 - 2\delta_1)d|X'| \le |X'_z| \le (1 + 2\delta_1)d|X'|$ and $(1 - 2\delta_1)d|Y'| \le |Y'_z| \le (1 + 2\delta_1)d|Y'|$.*

(v) *Given a vertex $x \in X'$, let $L'_x := L_x[Y' \cup Z']$. Let Y'_x and Z'_x denote the vertex classes of L'_x. Then for all but at most $2\xi|X'|$ vertices $x \in X'$ the graph L'_x is $(\alpha d, 4\xi)$-regular, $(1 - \varepsilon')d|Y'| \le |Y'_x| \le (1 + \varepsilon')d|Y'|$ and $(1 - \varepsilon')d|Z'| \le |Z'_x| \le (1 + \varepsilon')d|Z'|$. The same is true for all but at most $2\xi|Y'|$ vertices $y \in Y'$ and for all but at most $2\xi|Z'|$ vertices $z \in Z'$.*

(vi) *Given a pair of distinct vertices $z, z' \in Z'$, put $L'_{zz'} := L_{zz'}[X' \cup Y']$ and let $X'_{zz'}$ and $Y'_{zz'}$ denote the vertex classes of $L'_{zz'}$. Then for all but at most $2\xi|Z'|^2$ pairs of vertices $z, z' \in Z'$ the graph $L'_{zz'}$ is $(\alpha^2 d, 4\xi)$-regular, $(1 - \varepsilon')^2 d^2 |X'| \le |X'_{zz'}| \le (1 + \varepsilon')^2 d^2 |X'|$ and $(1 - \varepsilon')^2 d^2 |Y'| \le |Y'_{zz'}| \le (1 + \varepsilon')^2 d^2 |Y'|$. The same is true for all but at most $2\xi|X'|^2$ pairs $x, x' \in X'$ and for all but at most $2\xi|Y'|^2$ pairs $y, y' \in Y'$.*

(vii) *The subgraph $P_{X'Y'}^{rich}$, obtained from $P_{X'Y'}$ by deleting all its poor and unhappy edges is $(d, 5\delta_0)$-superregular.*

Proof. Propositions 7.10 and 7.15 together with inequality (10) imply (i). Property (ii) follows from Proposition 7.15 and property (iii) follows immediately from the definition of Z'.

To prove (iv), let $z \in Z'$ and let L_z^* be the subgraph obtained from L_z by deleting all poor edges and all vertices in $\widetilde{X} \cup \widetilde{Y}$. By Proposition 7.13, L_z^* is $(\alpha d, 3\xi)$-regular and the sizes of its vertex classes are as described there. But L_z^{rich} is obtained from L_z^* by deleting all the at most $3\varepsilon'|X|$ vertices lying in $\overline{X} \cup \widehat{X}$, all the at most $3\varepsilon'|Y|$ vertices lying in $\overline{Y} \cup \widehat{Y}$ as well as all the unhappy edges in L_z^* (cf. Proposition 7.15(ii) and (iii)). But since each vertex in $X'_z \subseteq X \setminus \overline{X}$ is incident to at most $3\varepsilon'|X|$ unhappy edges and since the vertices in Y'_z satisfy the analogous property, it follows that L_z^{rich} is still $(\alpha d, 4\xi)$-regular and that the sizes of its vertex classes are as desired in (iv).

Properties (v) and (vi) follow (with room to spare) from (P3) and (P4) respectively by using (P1). Property (vii) follows from Proposition 7.10 since X' was obtained from $X \setminus \widetilde{X}$ by deleting at most $3\varepsilon'|X|$ further vertices (and since the analogue holds for Y'). □

Finally, we single out those vertices which will be particularly useful in Section 9 because of their regularity properties.

Definition 7.17 ($X^\sharp$, $Y^\sharp$ and $Z^\sharp$) Let $X^\sharp \subseteq X'$ be the set of all those vertices $x \in X'$ which satisfy the following two properties.

- The link graph L_x of x is typical.
- L'_x is $(\alpha d, 4\xi)$-regular, $(1-\varepsilon')d|Y'| \leq |Y'_x| \leq (1+\varepsilon')d|Y'|$ and $(1-\varepsilon')d|Z'| \leq |Z'_x| \leq (1+\varepsilon')d|Z'|$. (Thus x does not belong to the at most $2\xi|X'|$. vertices described in Proposition 7.16(v).)

Define $Y^\sharp$ and $Z^\sharp$ similarly.

Property (P3) and Proposition 7.16(v) together imply that $|X' \setminus X^\sharp| \leq 4\xi|X'|$, $|Y' \setminus Y^\sharp| \leq 4\xi|Y'|$ and $|Z' \setminus Z^\sharp| \leq 4\xi|Z'|$.

Definition 7.18 (useful vertices in $X' \cup Y' \cup Z'$) We call a vertex $x \in X^\sharp$ *useful* if $(1-\varepsilon')d|Y^\sharp| \leq |N_{P_{XY}}(x) \cap Y^\sharp| \leq (1+\varepsilon')d|Y^\sharp|$ and $(1-\varepsilon')d|Z^\sharp| \leq |N_{P_{XZ}}(x) \cap Z^\sharp| \leq (1+\varepsilon')d|Z^\sharp|$. Similarly, we define the useful vertices in $Y^\sharp$ and in $Z^\sharp$.

Note that all but at most $5\xi|X'|$ (respectively $5\xi|Y'|$, $5\xi|Z'|$) vertices in X' (respectively Y', Z') are useful.

Recall that (X, Y, Z) was just one of the triples (X_k, Y_k, Z_k) obtained at the beginning of this section. We proceed in the same way with each of these triples (X_k, Y_k, Z_k) to obtain subsets X'_k, Y'_k and Z'_k.

8 Incorporating the exceptional vertices and choosing the bridges

8.1 Incorporating the exceptional vertices

Recall that when constructing the subsets $X'_k \subseteq X_k$, $Y'_k \subseteq Y_k$ and $Z'_k \subseteq Z_k$ in Section 7, we deleted at most $2\xi n$ vertices of $\mathcal{H}$ (cf. Proposition 7.15(i)). All these vertices are also added to V_0. Thus

$$|V_0| \overset{(8)}{\leq} \xi n + 2\xi n = 3\xi n.$$

Recall also that $|X_k| = |Y_k| = m/3$ and $|Z_k| = 2m/3$. Put

$$m_1 := m/100. \tag{12}$$

Choose sets $X''_k \subseteq X'_k$, $Y''_k \subseteq Y'_k$ and $Z''_k \subseteq Z'_k$ with $|X''_k| = |Y''_k| = m_1$ and $|Z''_k| = 2m_1$ among all such sets uniformly at random. In the remainder of this section and in the next section we will take care not to alter these sets X''_k, Y''_k and Z''_k. This will have the advantage that each of the triples (X''_k, Y''_k, Z''_k) is sufficiently regular so that in Section 11 we can find a loose path which contains *all* of the vertices in (X''_k, Y''_k, Z''_k) (and also a comparatively small number of vertices which were 'left over' in Section 9).

Put $X^*_k := X'_k \setminus X''_k$, $Y^*_k := Y'_k \setminus Y''_k$ and $Z^*_k := Z'_k \setminus Z''_k$. Remove at most $10\xi m/3 = 5\xi|Z_k|$ more vertices from the Z^*_k to ensure that

$$(2 - 10\xi)(m/3 - m_1) = |Z^*_k| \tag{13}$$

for each k. To see that Z^*_k is not smaller than this already, recall that $|Z_k \setminus Z'_k| \leq 2\xi|Z_k|$ by Proposition 7.16(i). (The fact that $|Z^*_k|$ is a little less than $|X^*_k| + |Y^*_k|$ will be useful in Lemma 9.1.) Add all these vertices to the exceptional set V_0. Thus

$$|V_0| \leq 8\xi n.$$

Next, we will find a loose path $\mathcal{L}$ which contains all the vertices of the exceptional set V_0. Let $v_1, \dots, v_r$ be an enumeration of the vertices in V_0. Let v_0 and v_{r+1} by any two vertices in X^*_1. We claim that for all the pairs $v_i v_{i+1}$ $(0 \leq i \leq r)$ we can find distinct vertices w_i such that $v_i v_{i+1} w_i$ is a hyperedge of $\mathcal{H}$, such that w_i is contained in one of the sets Z^*_k and such that none of the Z^*_k contains more than $\sqrt{\xi} m$ of the vertices w_i. It is easy to see that this is possible: Indeed, by Proposition 7.1, for each pair $v_i v_{i+1}$ there are at least $\sigma N/8$ sets Z_k such that $v_i v_{i+1}$ forms a hyperedge with at least $\sigma|Z_k|/8$ vertices in Z_k. (Recall that N denotes the number of triples (X_k, Y_k, Z_k).) Now by Lemma 4.1 and Proposition 7.16(i), we may assume that this is almost inherited by $Z'_k \setminus Z''_k$, i.e. there are at least $\sigma N/8$ sets Z_k such that $v_i v_{i+1}$ forms a hyperedge with at least $\sigma|Z'_k \setminus Z''_k|/9$ vertices in $Z'_k \setminus Z''_k$. Together with the fact that Z^*_k was obtained by deleting at most $5\xi|Z_k|$ further vertices, this implies that there are at least $\sigma N/8$ sets Z^*_k such that $v_i v_{i+1}$ forms a hyperedge with at least $\sigma|Z^*_k|/10$ vertices in Z^*_k. On the other hand, the total number of vertices w_i which we need is at most

$$|V_0| \leq 8\xi n \leq 12\xi N m \leq (\sqrt{\xi} m)(\sigma N/8)/2. \tag{14}$$

The second inequality follows since each triple (X_k, Y_k, Z_k) contains $4m/3$ vertices and we covered almost all vertices of $\mathcal{H}$ by such triples. Inequality (14) shows that we can choose the w_i greedily (as it shows that for all i there will always be one Z^*_k in which we have chosen less than $\sqrt{\xi} m$ vertices so far and which still contains a vertex w_i that forms a hyperedge together with v_i and v_{i+1}). This proves the claim.

Thus we have a loose path $\mathcal{L}$ which joins two vertices in X^*_1. We will remove all the vertices of this path from the sets Z^*_k. (So we do not remove v_0 and v_{r+1} from X^*_1.) We still denote the resulting sets by Z^*_k.

Definition 8.1 (still useful vertices in $X' \cup Y'$) We call a useful vertex $x \in X^*_k$ *still useful* if it satisfies the following three conditions:

(a) the link graph $L^*_x := L_x[Y^*_k \cup Z^*_k]$ spanned by Y^*_k and Z^*_k is still $(\alpha d, 5\xi)$-regular,

(b) the sizes of the vertex classes $Y^*_x \subseteq Y^*_k$ and $Z^*_x \subseteq Z^*_k$ of L^*_x satisfy $(1 - \varepsilon')d|Y^*_k| \leq |Y^*_x| \leq (1 + \varepsilon')d|Y^*_k|$ and $(1 - \varepsilon')d|Z^*_k| \leq |Z^*_x| \leq (1 + \varepsilon')d|Z^*_k|$.

(c) $(1-\varepsilon')d|Y_k^* \cap Y_k^\sharp| \le |P_{X_kY_k}(x) \cap Y_k^* \cap Y_k^\sharp| \le (1+\varepsilon')d|Y_k^* \cap Y_k^\sharp|$ and $(1-\varepsilon')d|Z_k^* \cap Z_k^\sharp| \le |P_{X_kZ_k}(x) \cap Z_k^* \cap Z_k^\sharp| \le (1+\varepsilon')d|Z_k^* \cap Z_k^\sharp|$.

The still useful vertices in Y_k^* are defined similarly.

Note that (a) holds for all but at most $2\varepsilon'|X_k|$ useful vertices in X_k^*. Indeed, recall that for every useful vertex x the link graph $L'_x = L_x[Y'_k \cup Z'_k]$ is $(\alpha d, 4\xi)$-regular. Moreover, the (d,ε')-superregularity of the bipartite graphs forming the triad P_k implies that for all but at most $2\varepsilon'|X_k|$ vertices x the sizes of the vertex classes of L_x^* are at least $9/10$ of those of L'_x. This implies that (a) holds for all but at most $2\varepsilon'|X_k|$ useful vertices in X_k^*. Similarly it follows that both (b) and (c) are true for all but at most $8\varepsilon'|X_k|$ vertices in X_k^*. Thus the number of useful vertices in X_k^* which are not still useful is at most $10\varepsilon'|X_k| \le 11\varepsilon'|X_k^*|$. Similarly, the number of those useful vertices in Y_k^* which are not still useful is at most $11\varepsilon'|Y_k^*|$.

8.2 Choosing the bridges

Recall that N denotes the number of triples (X_k, Y_k, Z_k). Write $(X_{N+1}, Y_{N+1}, Z_{N+1}) := (X_1, Y_1, Z_1)$. For all $2 \le k \le N$ we will glue the triples (X_k, Y_k, Z_k) and $(X_{k+1}, Y_{k+1}, Z_{k+1})$ together by choosing a single hyperedge connecting them. This is done as follows. We first choose vertices $x_k \in X_k^*$ and $y_{k+1} \in Y_{k+1}^*$ such that both x_k and y_{k+1} are still useful. For all these pairs x_k, y_{k+1} we then choose distinct vertices a_k such that $a_k x_k y_{k+1}$ is a hyperedge of $\mathcal{H}$ and such that each a_k lies in some Z_i^*. (This can be done by Proposition 7.1 since so far, we have removed at most $\sqrt{\xi}m$ vertices from each Z_k^*.) We will call the hyperedge $a_k x_k y_{k+1}$ the *kth bridge*.

To connect the triples (X_1, Y_1, Z_1) and (X_2, Y_2, Z_2) we proceed a little differently in order to incorporate the loose path $\mathcal{L}$ which contains the exceptional vertices: Recall that $v_0, v_{r+1} \in X_1^*$ are endvertices of $\mathcal{L}$. Choose vertices $x_1 \in X_1^*$ and $y_2 \in Y_2^*$ which are still useful. In the case when $|\mathcal{H}|$ is even we choose vertices a_0 and a_1 such that both $a_1 x_1 v_0$ and $a_0 v_{r+1} y_2$ are hyperedges of $\mathcal{H}$. (Each of a_0 and a_1 will again lie in some Z_i^* and these two vertices will be distinct from all the other a_k.) We will call the loose path which starts with the hyperedge $a_1 x_1 v_0$, continues with $\mathcal{L}$ and ends with the hyperedge $a_0 v_{r+1} y_2$ the *first bridge*. Let us now consider the case when $|\mathcal{H}|$ is odd. In this case, we make the first bridge even longer (it will contain 2 hyperedges which have exactly two vertices in common). We first choose a hyperedge $a_1 x_1 v_0$ as before. Next we pick a hyperedge $v_{r+1} b_0 b_1$ containing v_{r+1}. Then we pick any hyperedge $b_0 b_1 b_2$ containing both b_0 and b_1. Now we pick any vertex b_3 such that $b_2 b_3 y_2$ is a hyperedge of $\mathcal{H}$. All these vertices b_0, b_1, b_2, b_3 are chosen in such a way that they lie in some Z_i^*. Moreover, a_1 and all the b_i are chosen to be disjoint from each other as well as disjoint from $\mathcal{L}$ and all the bridges.) We will call the loose path which starts with the hyperedge $a_1 x_1 v_0$, continues with $\mathcal{L}$, $v_{r+1} b_0 b_1$, $b_0 b_1 b_2$ and ends with $b_2 b_3 y_2$ the *first bridge*.

We remove all the vertices a_i from the sets Z_k^* (as well as all the b_i if $|\mathcal{H}|$ is odd). We still denote the resulting sets by Z_k^*. Moreover, we remove v_0 and v_{r+1} from X_1^* and still denote the resulting set by X_1^*.

The aim now is to find for each k a loose path $\mathcal{P}_k$ such that $\mathcal{P}_k$ contains all the vertices in $X_k^* \cup Y_k^* \cup Z_k^*$ as well as all the vertices in $X_k'' \cup Y_k'' \cup Z_k''$ and such that the bridge vertices x_k and y_k are endvertices of $\mathcal{P}_k$. Then all these paths would form a loose Hamilton cycle together with all the bridges. But clearly, a necessary condition for this is that $X_k^* \cup Y_k^* \cup Z_k^* \cup X_k'' \cup Y_k'' \cup Z_k''$ has an odd number of vertices. Since $2|X_k''| = 2|Y_k''| = |Z_k''|$ and thus $|X_k'' \cup Y_k'' \cup Z_k''|$ is even, we need that

$$W_k^* := X_k^* \cup Y_k^* \cup Z_k^*$$

contains an odd number of vertices. We will now alter the sets W_k^* a little (and make the first bridge a little longer) to achieve this. We will call every vertex in W_k^* apart from x_k and y_k a non-bridge vertex. For every set W_k^* which has an even number of vertices we remove one non-bridge vertex from Z_k^*. Let W denote the set consisting of all these removed vertices. Let us next show that $|W|$ is even. We first consider the case when $|\mathcal{H}|$ is even. Then for each k for which $|W_k^*|$ was even, the set obtained from W_k^* by adding all the vertices in the kth bridge except for y_{k+1} has an odd number of vertices whereas the latter set has an even number of vertices if $|W_k^*|$ was odd. Hence $|W|$ is even if $|\mathcal{H}|$ is even. A similar argument works if $|\mathcal{H}|$ is odd since in this case the first bridge contains 2 consecutive hyperedges which have 2 vertices in common. Using Proposition 7.1 and the fact that $|X_1^*|$ is sufficiently large and almost all of its vertices are still useful, it is easy to see that for each $w \in W$ there are two still useful non-bridge vertices $s(w), s'(w) \in X_1^*$ such that for some k we have that both $|N_{\mathcal{H}}(w, s(w)) \cap Z_k^*| \geq \sigma |Z_k^*|/10$ and $|N_{\mathcal{H}}(w, s'(w)) \cap Z_k^*| \geq \sigma |Z_k^*|/10$ and such that all the pairs $s(w), s'(w)$ are disjoint for distinct vertices $w \in W$. Thus we can find two non-bridge vertices $a(w), a'(w) \in Z_k^*$ such that both $ws(w)a(w)$ and $ws'(w)a'(w)$ are hyperedges of $\mathcal{H}$ and such that all the pairs $a(w), a'(w)$ are disjoint for distinct vertices $w \in W$. Let $\mathcal{P}(w)$ denote the loose path consisting of the two hyperedges $ws(w)a(w)$ and $ws'(w)a'(w)$.

Let $w_1, \dots, w_{|W|}$ be any enumeration of the vertices in W. Lemma 8.2 below implies that we can glue $s'(w_i)$ to $s(w_{i+1})$ using a loose path $\mathcal{P}^*(w_{i+1})$ of length 5 such that all vertices of $\mathcal{P}^*(w_{i+1})$ lie in $X_1^* \cup Y_1^* \cup Z_1^* = W_1^*$ (see Fig. 3). Moreover, in a similar way we will glue the bridge vertex x_1 to $s(w_1)$ by a loose

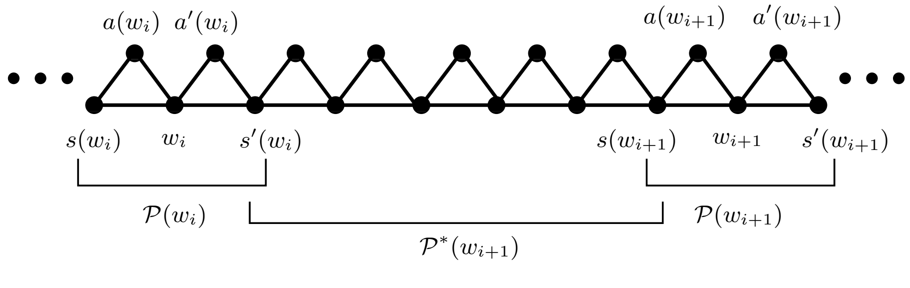

Figure 3: Extending the first bridge by incorporating the vertices $w_i \in W$ in order to make the leftover sets W_k^* even

path $\mathcal{P}^*(w_1)$. By Lemma 8.2, each of these paths $\mathcal{P}^*(w_i)$ can be chosen to be disjoint from the others and from all the vertices of the form $a(w)$, $a'(w)$. Since each $\mathcal{P}^*(w_i)$ has length 5 (i.e. it consists of 5 hyperedges), it uses 11 vertices of W_1. As $|W|$ is even, the number of vertices in the union of all the $\mathcal{P}^*(w_i)$ is even. Now enlarge the first bridge by adding all the $\mathcal{P}^*(w_i)$ and all the $\mathcal{P}(w_i)$ to obtain a new first bridge ending in $s'(w_{|W|})$. We write x_1 for this new endvertex $s'(w_{|W|})$. Thus the 2 endvertices x_1 and y_1 of the (modified) first bridge are both still useful. We delete all the vertices of this first bridge—apart from x_1 and y_1—from the sets W_k^*. Note that, for each k, the subset of W_k^* obtained in this way has an odd number of vertices. We still denote these subsets by W_k^* and the sets obtained from X_k^*, Y_k^*, Z_k^* by X_k^*, Y_k^*, Z_k^*. Since altogether we only removed a bounded number of vertices from the sets X_k^* and Y_k^* and at most $2\sqrt{\xi}|Z_k^*| \leq 4\sqrt{\xi}m/3$ vertices from each Z_k^*, Propositions 7.16(i) and inequality (13) together imply that (with room to spare in the error terms)

$$\begin{aligned} m/3 - m_1 - 2\delta_1 m \quad &\leq |X_k^*| \quad \leq m/3 - m_1, \\ m/3 - m_1 - 2\delta_1 m \quad &\leq |Y_k^*| \quad \leq m/3 - m_1, \\ (2 - 15\sqrt{\xi})(m/3 - m_1) \quad &\leq |Z_k^*| \quad \leq (2 - 10\xi)(m/3 - m_1). \end{aligned} \tag{15}$$

The following lemma is an analogue of the fact that in an ε-regular bipartite graph most pairs of vertices in one of the vertex classes are joined by many paths of length 2.

Lemma 8.2 *Any two distinct vertices $x, x' \in X_1^*$ which are still useful can be connected by a loose path of length* 5 *whose other vertices lie in $X_1^* \cup Y_1^* \cup Z_1^*$. In fact, there are even $|\mathcal{R}|$ such paths which meet only in x and x'.*

Proof. For simplicity, we write X^* for X_1^*, Y^* for Y_1^* etc. Roughly speaking, the idea of the proof is to show that a large fraction of the vertices in Y^* can play the role of an endpoint of a loose path of length two starting in x. Similarly, a large fraction of the vertices in X^* can play the role of an endpoint of a loose path of length two starting in x'. The regularity of the triad P_1 with respect to $\mathcal{H}$ then implies that there are many hyperedges of $\mathcal{H}$ which connect two such paths into a loose path of length 5. The precise argument is given below.

Recall that after the definition of those vertices in X^* and Y^* which are still useful (Definition 8.1), we removed only a bounded number of further vertices from Z^*. (We removed the vertices in the loose path $\mathcal{L}$ containing all the exceptional vertices before Definition 8.1.) Thus for vertices $v \in X^*$, $w \in Y^*$ which are still useful the properties of the link graphs L_v^* and L_w^* are not changed significantly. In particular, both L_x^* and $L_{x'}^*$ are still $(\alpha d, 6\xi)$-regular. Thus in the graph L_x^* at least $|Z_x^*|/2$ vertices in Z_x^* send at least $\alpha d|Y_x^*|/2$ edges to Y_x^*. Let Z_x^{**} denote the set of all these vertices in Z_x^*.

Recall that X^* and Y^* were obtained from X' and Y' by splitting off randomly chosen sets X'' and Y'' and then deleting a bounded number of further vertices (to incorporate the exceptional vertices and to construct the bridges). Hence by Propositions 7.16(iv) and Lemma 4.1 we may assume that for each $z \in Z_x^{**}$ the graph $L_z^* := L_z^{rich}[X^* \cup Y^*]$ is still $(\alpha d, 5\xi)$-regular and that its vertex class $Y_z^* \subseteq Y^*$ satisfies

$$3d|Y^*|/4 \leq |Y_z^*| \leq 5d|Y^*|/4. \tag{16}$$

Let Y_z^{**} denote the set of all those vertices in Y_z^* which in the graph L_z^* send at least $\alpha d|X_z^*|/2$ edges to X_z^*. Then $|Y_z^{**}| \geq 2|Y_z^*|/3 \geq d|Y^*|/2$. Put $Y^{**} := \bigcup_{z \in Z_x^{**}} Y_z^{**}$. Thus all vertices in $y^* \in Y^{**}$ can serve as endvertices of loose paths of length two that start in x.

We claim that $|Y^{**}| \geq |Y^*|/32$. Indeed, given a set $z_1, \dots, z_i$ of vertices in Z_x^{**} with $i \leq s$, where $s := \lceil 1/(8d) \rceil$, at least half of the remaining vertices $z \in Z_x^{**}$ satisfy

$$|Y_z^{**} \cap Y_{z_j}^{**}| \leq 2d|Y_z^{**}| \leq 2d|Y_z^*| \overset{(16)}{\leq} 3d^2|Y^*| \tag{17}$$

for all $j \leq i$. (The first inequality follows since by (P1) the triad P_1 is (d, ε')-superregular.) As long as $i \leq s$, define z_{i+1} to be such a vertex. Then

$$|Y^{**}| \geq |\bigcup_{i=1}^{s} Y_{z_i}^{**}| \overset{(17)}{\geq} s(d|Y^*|/2 - (s-1)3d^2|Y^*|/2) \geq |Y^*|/32.$$

Similarly to the above, we define $Z_{x'}^{**}$ to be the set of all those vertices $z \in Z_{x'}^*$ which send at least $\alpha d|Y_{x'}^*|/2$ edges to $Y_{x'}^*$. Moreover, given $z \in Z_{x'}^{**}$, we let X_z^{**} denote the set of all those vertices in X_z^* which in the graph L_z^* send at least $\alpha d|Y_z^*|/2$ edges to Y_z^*. As before, $|X_z^{**}| \geq d|X^*|/2$ and $|X^{**}| \geq |X^*|/32$ where $X^{**} := \bigcup_{z \in Z_{x'}^{**}} X_z^{**}$. (Thus all vertices in $x^* \in X^{**}$ can serve as endvertices of loose paths of length two that start in x'.)

Now recall from (P2) that the triad P_1 whose vertex classes are $X = X_1$, $Y = Y_1$ and $Z = Z_1$ is $(\delta, 1)$-regular with respect to $\mathcal{H}$ and $d_{\mathcal{H}}(P_1) = \alpha \geq \alpha^*/2$. Moreover, $|X^{**}| \geq |X|/80$, $|Y^{**}| \geq |Y|/80$ and $|Z^*| \geq |Z|/2$ with room to spare. Let Q denote the subtriad of P_1 induced by these sets. Thus Proposition 5.3 implies that $t(Q) \geq \delta \cdot t(P_1)$. Together with the regularity of P_1 with respect to $\mathcal{H}$ this shows that $d_{\mathcal{H}}(Q) \geq \alpha - \delta > \alpha/2$. Since the number $t(Q)$ of triangles spanned by Q is at least $d^3|X^{**}||Y^{**}||Z^*|/2$ (by Proposition 5.3 again) this implies that Q spans at least $\alpha t(Q)/2 \geq \alpha d^3|X^{**}||Y^{**}||Z^*|/4$ hyperedges of $\mathcal{H}$. This immediately implies that there are vertices $x^* \in X^{**} \setminus \{x, x'\}$ and $y^* \in Y^{**}$ so that the pair x^*y^* lies in at least $\alpha d^3|Z^*|/5 \geq 3$ hyperedges of $\mathcal{H}$.

These two vertices will lie in the middle hyperedge of our loose path between x and x' (Fig. 4).

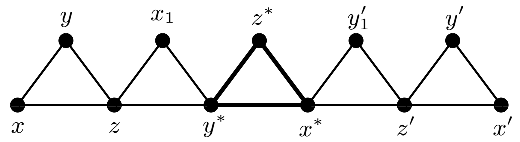

Figure 4: The loose path $\mathcal{P}$ which joins x to x'

We now choose the remaining vertices of the path: choose any vertex $z \in Z_x^{**}$ so that $y^* \in Y_z^{**}$ and choose any vertex $z' \in Z_{x'}^{**}$ so that $x^* \in X_{z'}^{**}$. Now we choose a vertex z^* so that $x^*y^*z^*$ forms a hyperedge of $\mathcal{H}$. Since there are at least three of these, we can assume that $z^* \neq z, z'$. We now choose two more vertices y and x_1 which together with z give us a loose path of length two between x and y^*. This can be done as follows. Take y to be any vertex in $Y_x^* \setminus \{y^*\}$ which is a neighbour of z in L_x^*. (Such a vertex exists since $z \in Z_x^{**}$ and thus z has at least $\alpha d|Y_x^*|/2$ neighbours in Y_x^*.) Then xyz is a hyperedge of $\mathcal{H}$. Take x_1 to be any vertex in $X_z^* \setminus \{x, x', x^*\}$ which is a neighbour of y^* in L_z^*. (Again, there are many candidates for x_1 since $y^* \in Y_z^{**}$.) Thus $x_1 y^* z$ is also a hyperedge of $\mathcal{H}$.

In the same way, we can also find two vertices $y', y_1' \neq y$ so that $x'y'z'$ and $z'y_1'x^*$ are hyperedges of $\mathcal{H}$. Altogether, this gives us a loose path of length 5 joining x and x'.

Remove the vertices of this path from $\mathcal{H}$ (except for x and x'). Since these are only boundedly many vertices, the regularity of the triad and the link graphs is not significantly affected and so we can we repeat the argument $|\mathcal{R}|$ times to obtain $|\mathcal{R}|$ disjoint loose paths of length 5, as required. □

9 Finding the equalizing paths and augmenting the bridges

In this section, we will find (for each k) our 'equalizing' path $\mathcal{Q}_k$ which contains almost all of the vertices in each of X_k^*, Y_k^* and Z_k^*.

Lemma 9.1 *For each triple (X_k, Y_k, Z_k) the induced hypergraph $\mathcal{H}_{P_k}$ contains a loose path $\mathcal{Q}_k$ with the following three properties:*

- *$\mathcal{Q}_k$ starts with the bridge vertex x_k and ends with some vertex $x_k^* \in X_k^*$ for which the link graph L_x was typical,*
- *$\mathcal{Q}_k$ contains only vertices in $X_k^* \cup Y_k^* \cup Z_k^*$ and avoids the bridge vertex y_k,*
- *the sets $X_k^{**} := X_k^* \setminus V(\mathcal{Q}_k)$, $Y_k^{**} := Y_k^* \setminus (V(\mathcal{Q}_k) \cup \{y_k\})$ and $Z_k^{**} := Z_k^* \setminus V(\mathcal{Q}_k)$ satisfy*

$$|X_k^{**}| = |Y_k^{**}| = 2\sqrt{\xi}m \quad \text{and} \quad |Z_k^{**}| = 2|X_k^{**}| + 1. \tag{18}$$

Since the sets X_k^*, Y_k^* and Z_k^* satisfy (15), the following lemma implies that such a path $\mathcal{Q}_k$ could be found if instead of the subhypergraph of $\mathcal{H}_{P_k}$ induced by X_k^*, Y_k^* and Z_k^* we would consider the complete 3-uniform hypergraph with these vertex sets. From successive applications of Lemma 9.3 it will then follow that any such path $\mathcal{Q}_k$ in the complete hypergraph can in fact also be found in subhypergraph of $\mathcal{H}_{P_k}$ induced by X_k^*, Y_k^* and Z_k^*.

Lemma 9.2 *Given $0 < \delta_1 \ll \xi \ll 1$, suppose that $\mathcal{G}$ is the complete 3-uniform 3-partite hypergraph with vertex classes A, B and C such that $||A| - |B|| \leq 10\delta_1|A|$, $(2 - 16\sqrt{\xi})|A| \leq |C| \leq (2 - 9\xi)|A|$ and such that $|A \cup B \cup C|$ is even and sufficiently large compared to $1/\delta_1$. Let $a \leq |A|/2$ be any natural number. Then $\mathcal{G}$ contains a loose path $\mathcal{P}$ whose first and last vertex lies in A and such that $|A \setminus V(\mathcal{P})| = |B \setminus V(\mathcal{P})| = a$ and $|C \setminus V(\mathcal{P})| = 2a + 1$.*

Proof. Given a directed loose path $\mathcal{P} = e_1 \dots e_r$ in $\mathcal{G}$ (where $e_i \in E(\mathcal{G})$), we say that a hyperedge e_i ($i \geq 2$) is of *type* A if the unique vertex that is contained in both e_i and e_{i-1} belongs to A. We say that e_1 is of *type* A if A contains one of the two vertices in e_1 that do not lie in e_2. Similarly we define hyperedges of *type* B and *type* C.

We will first consider the case when $|A| \geq |B|$. In this case the path $\mathcal{P}$ is constructed in 5 consecutive 'phases' as follows (Fig. 5). The first phase

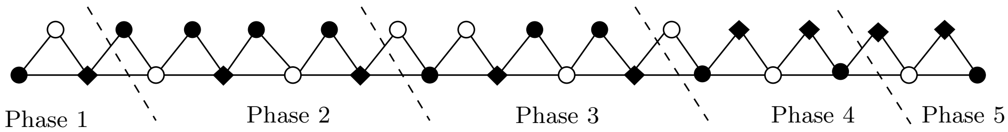

Figure 5: An example illustrating the case $|A| \geq |B|$. The black bullets are vertices in A, the white bullets lie in B and the squares in C. The Phase 4 is non-empty.

consists of a single hyperedge of type A. In the second phase we then continue by successively adding path segments consisting 2 hyperedges at a time: each segment consists of a hyperedge of type C followed by a hyperedge of type B. We claim that eventually we will obtain leftover subsets A^1, B^1, C^1 of A, B, C which satisfy $|A^1| = |B^1|$. Indeed, each path segment contains two new vertices in A and one new vertex in both B and C (here we do not count the first vertex in C as a new vertex as it already lies in the previously constructed part of $\mathcal{P}$). Thus after each such segment the difference between the sizes of the remaining subsets of A and B is reduced by 1, which proves the claim. Moreover, $|A^1| = |A| - 1 - 2(|A| - |B|) = |B^1|$ and $|C^1| = |C| - 1 - (|A| - |B|)$. Since $|B| \geq (1 - 10\delta_1)|A|$, it follows that

$$(2 - 17\sqrt{\xi})|A^1| \leq |C^1| \leq (2 - 8\xi)|A^1|. \tag{19}$$

In the third phase we successively add loose path segments of length 4, each of which consists of a hyperedge of type C, followed by a hyperedge of type A, followed by a hyperedge of type C, followed by a hyperedge of type B. We claim that in this way we can achieve that the number of leftover vertices in C is either exactly one more than twice the number of leftover vertices in A (and that the latter is still precisely as large as the number of leftover vertices in B) or exactly one less than twice the number of leftover vertices in A. Indeed, define k by $|C^1| = 2|A^1| - k$. So (19) implies that $8\xi|A^1| \leq k \leq 17\sqrt{\xi}|A^1|$. But each of the segments in Phase 3 uses 3 new vertices in A, 3 new vertices in B and only 2 new vertices in C (again, the first vertex from C in such a path is not counted). Thus after extending the path $\mathcal{P}$ by one such segment, we obtain leftover sets A^2, B^2 and C^2 for which $|A^2| = |B^2|$ and $|C^2| = 2|A^2| - k + 4$. Since $|A \cup B \cup C|$ is even and the current path $\mathcal{P}$ has an odd number of vertices, $|A^2 \cup B^2 \cup C^2|$ is odd. Thus k is odd. So after $\lfloor (k+1)/4 \rfloor$ steps in Phase 3 we obtain leftover sets A^3, B^3 and C^3 as claimed. Moreover, note that $|A^3| = |A^1| - 3\lfloor (k+1)/4 \rfloor \geq (1 - 13\sqrt{\xi})|A| \geq a + 2$.

If $2|A^3| - 1 = |C^3|$, then in the fourth phase we add 3 more hyperedges: one of type C, followed by one of type A, followed by one of type B. Thus we have 2 new vertices in each of A, B and C. So after this fourth phase we obtain leftover sets A^4, B^4 and C^4 for which $|A^4| = |B^4|$ and $|C^4| = 2|A^2| + 1$. If already $2|A^3| + 1 = |C^3|$, then the fourth phase is empty.

Finally, in the fifth phase we enlarge the path by alternately adding hyperedges of type A and B (ending with an hyperedge of type B) until we obtain a path $\mathcal{P}$ which misses the desired number of vertices in A, B and C.

The case when $|A| \leq |B|$ is similar. We start again with a hyperedge of type A. But then we continue alternatingly with hyperedges of type C and type A (ending with a hyperedge of type A) until the subsets A^1, B^1, C^1 of A, B, C thus obtained satisfy $|A^1| = |B^1|$. Now we proceed precisely as before. □

Lemma 9.3 *Suppose that $X^\diamond \subseteq X^* \cap X^\sharp$, $Y^\diamond \subseteq Y^* \cap Y^\sharp$, $Z^\diamond \subseteq Z^* \cap Z^\sharp$ are sets of size at least $\sqrt{\xi}m$. Moreover, suppose that $x^\diamond \in X^*$ satisfies $|N_{P_{XY}}(x^\diamond) \cap Y^\diamond| \geq (1 - 2\varepsilon')d|Y^\diamond|$ and $|N_{P_{XZ}}(x^\diamond) \cap Z^\diamond| \geq (1 - 2\varepsilon')d|Z^\diamond|$. Then there are vertices $y^\diamond \in Y^\diamond$ and $z^\diamond \in Z^\diamond$ such that $x^\diamond y^\diamond z^\diamond$ forms a hyperedge of $\mathcal{H}$ and such that*

(i) $|N_{P_{XY}}(y^\diamond) \cap X^\diamond| \geq (1 - \varepsilon')d|X^\diamond|$ *and* $|N_{P_{YZ}}(y^\diamond) \cap Z^\diamond| \geq (1 - \varepsilon')d|Z^\diamond|$,

(ii) $|N_{P_{XZ}}(z^\diamond) \cap X^\diamond| \geq (1 - \varepsilon')d|X^\diamond|$ *and* $|N_{P_{YZ}}(z^\diamond) \cap Y^\diamond| \geq (1 - \varepsilon')d|Y^\diamond|$.

The analogous statement where we are given a vertex $y^\diamond \in Y^\diamond$ and seek vertices $x^\diamond$ and $z^\diamond$ (or where we are given a vertex $z^\diamond$) also holds.

Proof. Let $Y^\diamond_{x^\diamond} := N_{P_{XY}}(x^\diamond) \cap Y^\diamond$ and $Z^\diamond_{x^\diamond} := N_{P_{XZ}}(x^\diamond) \cap Z^\diamond$. We first prove that $L_{x^\diamond}[Y^\diamond \cup Z^\diamond] =: L^\diamond_{x^\diamond}$ is still $(\alpha d, 2\sqrt{\xi})$-regular. Indeed, since $x^\diamond \in X^\sharp$, the link graph $L_{x^\diamond}$ is typical (cf. Definition 7.17) and thus by Definition 7.3 in particular $(\alpha d, \xi)$-regular. So we only need to show that $2\sqrt{\xi}|Y^\diamond_{x^\diamond}| \geq \xi|Y_{x^\diamond}|$ and that the analogue holds for $Z^\diamond_{x^\diamond}$. But this is true since

$$|Y^\diamond_{x^\diamond}| \geq (1 - 2\varepsilon')d|Y^\diamond| \geq (1 - 2\varepsilon')\sqrt{\xi}dm \geq \frac{(1 - 2\varepsilon')\sqrt{\xi}}{1 + \varepsilon'}|Y_{x^\diamond}| \geq \sqrt{\xi}|Y_{x^\diamond}|/2.$$

Here we used that $|Y_{x^\diamond}| \leq (1 + \varepsilon')d|Y| < (1 + \varepsilon')dm$ since $L_{x^\diamond}$ is typical (cf. Definition 7.3). A similar argument works for $Z^\diamond_{x^\diamond}$. This proves the claim.

Since by (P1) each of P_{XY}, P_{YZ} and P_{XZ} is (d, ε')-superregular, at most $2\varepsilon'|Y| \leq |Y^\diamond_{x^\diamond}|/2$ vertices in $Y^\diamond$ violate (i) and at most $2\varepsilon'|Z| \leq |Z^\diamond_{x^\diamond}|/2$ vertices in $Z^\diamond$ violate (ii). The $(\alpha d, 2\sqrt{\xi})$-regularity of $L^\diamond_{x^\diamond}$ implies that it contains an edge $y^\diamond z^\diamond$ missing these vertices. But $y^\diamond$ and $z^\diamond$ are as required in the lemma.

The proof for the analogous statements when we are given a vertex $y^\diamond$ or $z^\diamond$ is identical. □

Proof of Lemma 9.1. Put

$$a := 2\sqrt{\xi}m = 6\sqrt{\xi}|X|.$$

Thus our desired loose path $\mathcal{Q}_k$ should contain all but a vertices in each of X_k^* and $Y_k^* \setminus \{y_k\}$ and all but $2a+1$ vertices in Z_k^*. An application of Lemma 9.2 with $a = 2\sqrt{\xi}m$ as above, $A := X_k^*$, $B := Y_k^* \setminus \{y_k\}$ and $C := Z_k^*$ guarantees the existence of some loose path as desired in the *complete* 3-partite 3-uniform hypergraph with vertex classes A, B and C (and not yet in the subhypergraph of $\mathcal{H}_{P_k}$ induced by these sets). Indeed, we can apply Lemma 9.2 since by (15) the sizes of A, B and C satisfy the assumptions of this lemma. Moreover, we have made $|X_k^* \cup Y_k^* \cup Z_k^*|$ odd in Section 8.2 and so $|A \cup B \cup C|$ is even. Clearly, we may assume that x_k is an endvertex of the path $\mathcal{P}$ guaranteed by Lemma 9.2.

Our aim now is to show that there exists a loose path $\mathcal{Q}_k$ in $\mathcal{H}_{P_k}$ which meets the sets X_k^*, $Y_k^* \setminus \{y_k\}$ and Z_k^* in the same way as the path $\mathcal{P}$ meets the sets A, B and C (in the sense that if e.g. the 14th and the 15th hyperedge of $\mathcal{P}$ meet in some vertex in $A = X_k^*$, then so do the 14th and the 15th hyperedge of $\mathcal{Q}_k$). Moreover, the path $\mathcal{Q}_k$ will use only vertices in $X_k^\diamond := X_k^* \cap X_k^\sharp$, $Y_k^\diamond := (Y_k^* \cap Y_k^\sharp) \setminus \{y_k\}$ and $Z_k^\diamond := Z_k^* \cap Z_k^\sharp$. Thus it will also satisfy the first condition of Lemma 9.1. The hyperedges of $\mathcal{Q}_k$ will be chosen successively—the existence of the next hyperedge will be guaranteed by Lemma 9.3 at each step. Indeed, let us first show that we can apply Lemma 9.3 with $x^\diamond := x_k$ to find the first hyperedge of $\mathcal{Q}_k$. Since x_k was chosen to be still useful, it follows from the definition of the still useful vertices (Definition 8.1(c)) and the fact that we only removed a bounded number of vertices from Y_k^* and Z_k^* after defining the still useful vertices that we have

$$|N_{P_{X_kY_k}}(x_k) \cap Y_k^\diamond| \geq (1-2\varepsilon')d|Y_k^\diamond| \quad \text{and} \quad |N_{P_{X_kZ_k}}(x_k) \cap Z_k^\diamond| \geq (1-2\varepsilon')d|Z_k^\diamond|.$$

Thus $x^\diamond = x_k$ satisfies the assumptions of Lemma 9.3 and so we can apply this lemma to find the first hyperedge of our path $\mathcal{Q}_k$.

Let us now show that Lemma 9.3 can in fact be applied in each step. The sets $X^\diamond$, $Y^\diamond$ and $Z^\diamond$ in Lemma 9.3 will be those subsets of $X_k^\diamond$, $Y_k^\diamond$ and $Z_k^\diamond$ which avoid the path segment of $\mathcal{Q}_k$ constructed so far (except for the vertex to which we want to attach the next hyperedge). We have to guarantee that each of these 3 sets has size at least $\sqrt{\xi}m$. But this holds for the first two of these sets: on the one hand, the path segment of $\mathcal{Q}_k$ constructed so far avoids at least $a = 2\sqrt{\xi}m$ vertices in each of X_k^* and Y_k^*. On the other hand, $|X_k^* \setminus X_k^\sharp|, |Y_k^* \setminus Y_k^\sharp| \leq \sqrt{\xi}m$. (Indeed, to see this, note that after Definition 7.17 we showed that $|X_k' \setminus X_k^\sharp| \leq 4\xi|X_k'| \leq 4\xi m$.) So let us now show that the subset of $Z^\diamond$ which avoids the path segment of $\mathcal{Q}_k$ constructed so far also has size at least $\sqrt{\xi}m$. Again, this holds since one the one hand, this path segment avoids at least $2a = 4\sqrt{\xi}m$ vertices in Z_k^*, while on the other hand $|Z_k^* \setminus Z_k^\sharp| \leq |Z_k' \setminus Z_k^\sharp| \leq 4\xi|Z_k'| \leq 4\xi m$. It is easily seen that by Lemma 9.3(i) and (ii) also the vertex in the current $\mathcal{Q}_k$ to which the next hyperedge is to be attached satisfies the assumptions of Lemma 9.3. Thus we can find our path $\mathcal{Q}_k$ by applying Lemma 9.3 successively. □

After we have found $\mathcal{Q}_k$, we will extend the kth bridge joining the triple (X_k, Y_k, Z_k) to the triple $(X_{k+1}, Y_{k+1}, Z_{k+1})$ by adding $\mathcal{Q}_k$. Thus the kth bridge has a new starting point $x_k^* \in X_k^*$ and still ends with $y_{k+1} \in Y_{k+1}^*$. We set $y_{k+1}^* := y_{k+1}$.

10 Perfect matchings in superregular graphs

10.1 Random perfect matchings

In this subsection, we collect several results about (random) perfect matchings in bipartite superregular graphs which are all proven in [17]. The main result is Theorem 10.3. Given a superregular graph G and a subgraph H of G, it gives precise bounds on the likely number of edges of H which are contained in a random perfect matching M of G.

The next lemma implies that if we are given a (super-)regular graph G and a 'bad' subgraph F of G which is comparatively sparse, then a random perfect matching of G will probably only contain a few bad edges.

Lemma 10.1 *For all positive constants ε and d with $d \leq 1$ and $\varepsilon \leq 1/6$ there exists an integer $n_0 = n_0(\varepsilon, d)$ such that the following holds. Let G be a (d, ε)-superregular graph whose vertex classes A and B satisfy $|A| = |B| =: n \geq n_0$. Let M be a perfect matching chosen uniformly at random from the set of all perfect matchings of G. Let F be a subgraph of G such that all but at most $\Delta' n$ vertices in F have degree most $\Delta' dn$ in F, where $1/2 \geq \Delta' \geq 18\varepsilon$. Then the probability that M contains at least $9\Delta' n$ edges of F is at most $\mathrm{e}^{-\varepsilon n}$. Moreover, the assertion also holds if we assume that G is dn-regular (where dn is an integer).*

The following lemma shows that a randomly chosen 2-factor in a (super-)regular graph G will typically only contain few cycles. A similar observation was also used in Frieze and Krivelevich [8].

Lemma 10.2 *For all positive constants $\varepsilon < 1/64$ and $d \leq 1$ there exists an integer $n_0 = n_0(\varepsilon, d)$ such that the following holds. Let G be a (d, ε)-superregular graph whose vertex classes A and B satisfy $|A| = |B| =: n \geq n_0$. Let M_1 be any perfect matching in G. Let M_2 be a perfect matching chosen uniformly at random from the set of all perfect matchings in $G - M_1$. Let $R = M_1 \cup M_2$ be the resulting 2-factor. Then the probability that R contains more than $n/(\log n)^{1/5}$ cycles is at most e^{-n}. Moreover, the statement also holds if we assume that G is dn-regular (where dn is an integer) and that G and M_1 are disjoint.*

We now come to the main result of this section. We will apply this later on with the link graphs of the vertices in $Z''_k \cup Z^{**}_k$ playing the role of H (and similarly for the colink graphs). The special case of Theorem 10.3 where H is a sufficiently large induced subgraph of G is already due to Rödl and Ruciński [20].

Theorem 10.3 *For all positive constants $d, \nu_0, \eta \leq 1$ there is a positive $\varepsilon = \varepsilon(d, \nu_0, \eta)$ and an integer $N_0 = N_0(d, \nu_0, \eta)$ such that the following holds for all $n \geq N_0$ and all $\nu \geq \nu_0$. Let $G = (A, B)$ be a (d, ε)-superregular bipartite graph whose vertex classes both have size n and let H be a subgraph of G with $e(H) = \nu e(G)$. Choose a perfect matching M uniformly at random in G. Then with probability at least $1 - \mathrm{e}^{-\varepsilon n}$ we have*

$$(1 - \eta)\nu n \leq |M \cap E(H)| \leq (1 + \eta)\nu n.$$

The intuition behind this result is the following: If the inclusion of the edges of G into the random perfect matching M would be mutually independent and equally likely, then the probability that a given edge e is contained in M would be close to $n/e(G)$. Thus the expected value of $|M \cap E(H)|$ would be close to $ne(H)/e(G) = \nu n$. The above result would thus immediately follow by an application of some large deviation bound on the tail of the binomial distribution. Note that Lemma 10.1 does not follow from Theorem 10.3 (not even the first part follows, as the graph F there can be much sparser than the graph H in Theorem 10.3).

10.2 Perfect matchings in subgraphs of superregular graphs

Recall that a k-factor in a graph G is a spanning subgraph of G in which every vertex has degree k. The following lemma (or more precisely Corollary 10.5 following it) guarantees the existence of a k-factor in a graph G which in turn is a spanning subgraph of high minimum degree of a superregular graph G'. We will apply Corollary 10.5 twice in the proof of Lemma 11.4. As we will see later on, working with the k-factor has the advantage that while we may not be able to apply Lemmas 10.1 and 10.2 to the graph G under consideration directly, we can apply them (or more precisely the 'moreover' parts) to this k-factor.

Lemma 10.4 *Let d, ε be constants such that $0 < \varepsilon < 1/3$ and $2\varepsilon \leq d \leq 1$. Let G' be a (d, ε)-regular bipartite graph with vertex classes A and B, where $|A| = |B| = \overline{m}$. Let G be a spanning subgraph of G' with minimum degree $\delta(G) \geq 2d\overline{m}/3$. Then G has a perfect matching.*

Proof. Consider any set $I \subseteq A$. We will show that $|N_G(I)| \geq |I|$. Then G contains a perfect matching by Hall's theorem. Clearly, we may assume that $|I|, |N_G(I)| \geq \delta(G) \geq \varepsilon\overline{m}$. Also, the number $e_G(I, N_G(I))$ of edges in G between I and $N_G(I)$ satisfies

$$e_G(I, N_G(I)) \geq \delta(G)|I| \geq \frac{2}{3}d\overline{m}|I|$$

On the other hand, the (d, ε)-regularity of G' implies that

$$e_G(I, N_G(I)) \leq (1 + \varepsilon)d|I||N_G(I)| < \frac{4}{3}d|I||N_G(I)|.$$

Combining these inequalities shows that $|N_G(I)| > \overline{m}/2$. Thus we may now assume that $|I| > \overline{m}/2$. Put $\widetilde{I} := B \setminus N_G(I)$. Clearly, we may assume that $\widetilde{I} \neq \emptyset$. Thus $|N_G(\widetilde{I})| \geq \delta(G) \geq \varepsilon\overline{m}$ and so we may assume that $|I| \leq (1 - \varepsilon)\overline{m}$. This in turn shows that we may assume that $|\widetilde{I}| \geq \varepsilon\overline{m}$. Thus the above double counting argument applied to $\widetilde{I}$ instead of I shows that $|N_G(\widetilde{I})| > \overline{m}/2$. But this is a contradiction, as $N_G(\widetilde{I}) \subseteq A \setminus I$. □

The following corollary follows immediately by repeated applications of Lemma 10.4.

Corollary 10.5 *Let d, ε be constants such that $0 < \varepsilon < 1/3$ and $2\varepsilon \leq d \leq 1$. Let G' be a (d, ε)-regular bipartite graph with vertex classes A and B, where $|A| = |B| = \overline{m}$. Let G be a spanning subgraph of G' with minimum degree $\delta(G) \geq \theta d\overline{m}$ where $\theta \geq 2/3$. Then G contains a $\lfloor(\theta - 2/3)d\overline{m}\rfloor$-factor.*

11 Finding a loose Hamilton path in the remainder of each triple

Our aim now is to find for each k a loose path $\mathcal{Q}_k^*$ which contains all the vertices in

$$W_k := X_k^{**} \cup Y_k^{**} \cup Z_k^{**} \cup X_k'' \cup Y_k'' \cup Z_k'',$$

starts in x_k^*, ends in $y_k = y_k^*$ and avoids all vertices outside W_k except for x_k^* and y_k^*. (Recall that x_k^* was an endpoint of the kth bridge and $y_k = y_k^*$ was an endpoint of the $(k-1)$th bridge. Moreover, recall that the sets X_k^{**}, Y_k^{**} and Z_k^{**} were the 'leftover vertices' defined in Lemma 9.1 and that the sets X_k'', Y_k'' and Z_k'' were the random sets set aside in Section 8.1.) The union of all these paths $\mathcal{Q}_k$ and all the bridges will form the desired loose Hamilton cycle in $\mathcal{H}$.

So let us consider any k. In what follows, we will write W for W_k, $\mathcal{Q}^*$ for $\mathcal{Q}_k^*$, X^{**} for X_k^{**} etc. The existence of $\mathcal{Q}^*$ will be proved in two steps. In Section 11.1 we first find a Hamilton path R^* in $G^{rich} := P_{X'Y'}^{rich}[X^{**} \cup X'' \cup Y^{**} \cup Y'']$ which starts in some neighbour x'' of y^* in X'' and ends in some neighbour y'' of x^* in Y''. In Section 11.2 we then show that we have chosen R^* in such a way that $x^*R^*y^*$ can be extended to the desired loose path $\mathcal{Q}^*$ as follows: there will be a bijection between $Z'' \cup Z^{**}$ and the edges of $x^*R^*y^*$ such that if $z \in Z'' \cup Z^{**}$ is sent to $xy \in x^*R^*y^*$ then xyz is a hyperedge of $\mathcal{H}$. For this to work we need that $|Z^{**} \cup Z''| = e(x^*R^*y^*) = 2|X^{**} \cup X''| + 1$. But this holds since we adjusted the sizes of the triples by our 'equalizing path' in Lemma 9.1.

To show the existence of the desired bijection we will consider an auxiliary bipartite graph H^* whose first vertex class consists of the set of edges of R^* and whose other vertex class consists of $Z^{**} \cup Z''$. We put an edge between $xy \in R^*$ and z if xyz forms a hyperedge of $\mathcal{H}_P$. We then show that H^* contains a perfect matching using Hall's theorem.

11.1 Finding the Hamilton path R^*

Recall from Section 5.3 that

$$\varepsilon = 50\sqrt{\varepsilon'}, \quad \nu = 10^5\sqrt{\xi}. \tag{20}$$

Put $\bar{n} := |X^{**} \cup X''|$. Then by (18) and (12)

$$|X''| + 2\sqrt{\xi}m = \bar{n} \geq |X''| = m_1 = m/100. \tag{21}$$

Choose random subsets X^r and Y^r of X'' and Y'' respectively such that

$$|X^r| = |Y^r| = \nu|X''| = \nu|Y''| \leq \nu\bar{n}. \tag{22}$$

Recall that $P_{X'Y'}$ is $(d, \sqrt{\varepsilon'})$-superregular by Proposition 7.16(ii). Moreover, in Section 8.1 we obtained X'' from a random partition of X' where $|X''| = m/100 \geq 3|X'|/100$. Together with Lemma 4.1 this implies that we may assume that $P_{X'Y'}[X'' \cup Y'']$ is $(d, 40\sqrt{\varepsilon'})$-superregular. By a similar argument we may assume that

$$G_0 := P_{X'Y'}[(X'' \setminus X^r) \cup (Y'' \setminus Y^r)]$$

is $(d, 50\sqrt{\varepsilon'})$-superregular, i.e. (d, ε)-superregular. Moreover, we have that

$$|X^r| = |Y^r| = \frac{\nu m}{100} = 10^3 \sqrt{\xi} m \geq 3 \cdot 10^3 \sqrt{\xi} |X'|. \tag{23}$$

The Hamilton path R^* will be found in several steps. In Lemma 11.3 we choose two disjoint random perfect matchings M_1 and M_2 in G_0. The resulting 2-factor $M_1 \cup M_2$ will consist of only a few cycles. We then delete the poor and unhappy edges from this 2-factor and at most one edge from each cycle to obtain a small number of paths. We call the collection of these paths $\mathcal{P}$. In Lemma 11.4 we will extend these paths into a 2-factor of the larger graph $P^{rich}_{X'Y'}[X^{**} \cup X'' \cup Y^{**} \cup Y'']$. This 2-factor is then transformed into the desired Hamilton path R^* by altering only a few edges. Theorem 10.3 implies that $M_1 \cup M_2$ contains a significant number of vertices of the link graph L^{rich}_z for all $z \in Z'' \cup Z^{**}$ and thus this is also true for R^*. In other words, the auxiliary bipartite graph H^* has high minimum degree (see Lemma 11.10), which is one of the properties we need in order to ensure that H^* contains a perfect matching. We cannot apply Theorem 10.3 directly to the graph $P^{rich}_{X'Y'}[X^{**} \cup X'' \cup Y^{**} \cup Y'']$ as this is not sufficiently regular, which is the reason why we find a 2-factor in G_0 first and then use this to find a 2-factor of $P^{rich}_{X'Y'}[X^{**} \cup X'' \cup Y^{**} \cup Y'']$.

Definition 11.1 (N_e and F_e) Given an edge $e = xy \in P_{X'Y'}$, we write $N_e := N_{P_{XZ}}(x) \cap N_{P_{YZ}}(y) \cap Z''$. Furthermore, we denote by F_e the set of all those edges $e' = x'y' \in P_{X'Y'}$ which are disjoint from e and satisfy $|N_e \cap N_{e'}| \geq (1 + 6\xi) d^4 |Z''|$.

The following proposition will enable us to choose our Hamilton path R^* so that for all of its edges e, R^* does not contain too many edges from F_e. This will be one of the properties we need to ensure that the auxiliary graph H^* contains a perfect matching.

Proposition 11.2 *Suppose that $x \in X'$ and $y \in Y'$ are such that the edge $e := xy$ is not unhappy. Then all but at most $\varepsilon^2|X|$ vertices $x' \in X'$ are incident to at most $\varepsilon^2|X|$ edges in F_e.*

Proof. Since $e = xy$ is not unhappy, we know that

$$|N_{P_{XZ}}(x) \cap N_{P_{YZ}}(y) \cap Z'| \leq (1 + \varepsilon')^2 d^2 |Z| \leq (1 + 3\xi) d^2 |Z'|$$

(use Definition 7.14 and the fact that $|Z| \leq |Z'|/(1-2\xi)$ by Proposition 7.16(i)). Since Z'' was obtained by considering a random subset of Z', Lemma 4.1 shows that we may assume that $|N_e| \leq (1 + 4\xi) d^2 |Z''|$. But since P_{XZ} is (d, ε')-superregular by (P1), this implies that there are at most $\varepsilon'|X| \leq \varepsilon^2|X|$ vertices $x' \in X'$ for which $|N_e \cap N_{P_{XZ}}(x')| \geq (1 + 5\xi) d^3 |Z''|$. We claim that any vertex $x' \in X'$ with $|N_e \cap N_{P_{XZ}}(x')| \leq (1 + 5\xi) d^3 |Z''|$ is incident to at most $\varepsilon'|Y| = \varepsilon'|X| \leq \varepsilon^2|X|$ edges in F_e. This is true since by (P1) the graph P_{YZ} is (d, ε')-superregular and thus there are at most $\varepsilon'|Y|$ vertices $y' \in Y'$ for which $|N_e \cap N_{P_{XZ}}(x') \cap N_{P_{YZ}}(y')| \geq (1 + 6\xi) d^4 |Z''|$. □

We now apply the following lemma to obtain two perfect matchings M_1 and M_2 in the graph G_0. Recall from Definition 7.2 that $\mathcal{H}_P$ denotes the subhypergraph of $\mathcal{H}$ whose vertex set is $X \cup Y \cup Z$ and whose hyperedges are those hyperedges of $\mathcal{H}$ which are triangles of the triad P.

Lemma 11.3 *G_0 contains two edge-disjoint perfect matchings M_1 and M_2 such that each of them satisfies the following properties:*

(i) *For every vertex $z \in Z'' \cup Z^{**}$ the matching M_i meets the link graph L_z^{rich} in at least $(1-\eta_1)\alpha d^2 \bar{n}$ edges.*

(ii) *For at most $\sqrt{\xi}\bar{n}^2$ pairs $z, z' \in Z'' \cup Z^{**}$ there are more than $(1+\eta_1)\alpha^2 d^4 \bar{n}$ edges $xy \in M_i$ for which both xyz and xyz' are hyperedges of $\mathcal{H}_P$.*

(iii) *For every vertex $z \in Z'' \cup Z^{**}$ there are at most $\varepsilon \bar{n}$ other vertices $z' \in Z'' \cup Z^{**}$ such that M_i contains more than $(1+\eta_1)d^4\bar{n}$ edges xy for which both xyz and xyz' are hyperedges of $\mathcal{H}_P$.*

(iv) *$|M_i \cap F_e| \leq 170\varepsilon\bar{n}$ for every edge $e \in P_{X'Y'}$ which is not unhappy.*

(v) *M_i contains at most $\xi\bar{n}$ edges which are poor or unhappy.*

(vi) *The 2-factor $M_1 \cup M_2$ contains at most $\bar{n}/(\log \bar{n})^{1/5}$ cycles.*

Proof. Let M_1 be a perfect matching in G_0 which is chosen uniformly at random. Let M_2 be a perfect matching in $G_0 - M_1$ chosen uniformly at random. We will show that with high probability M_1 and M_2 have the desired properties.

Let us first show that the probability that (i) fails for M_1 is exponentially small in $|M_1|$. Recall that X'' was a random subset of X' of size $m/100 \geq 3|X'|/100$ and that X^r was a random subset of X'' of size $\nu|X''|$. Moreover, recall that by Proposition 7.16(iv) the graph L_z^{rich} is $(\alpha d, 4\xi)$-regular and the sizes of its vertex classes are as described there. Together with Lemma 4.1 this implies that we may assume that $L_z^{rich} \cap G_0 = L_z^{rich}[(X'' \setminus X^r) \cup (Y'' \setminus Y^r)]$ is still $(\alpha d, 150\xi)$-regular and that both its vertex classes

$$X_z'' := N_{P_{XZ}}(z) \cap (X'' \setminus X^r) \quad \text{and} \quad Y_z'' := N_{P_{YZ}}(z) \cap (Y'' \setminus Y^r)$$

satisfy

$$(1-3\delta_1)d|X'' \setminus X^r| \leq |X_z''| \leq (1+3\delta_1)d|X'' \setminus X^r| \tag{24}$$

and

$$(1-3\delta_1)d|Y'' \setminus Y^r| \leq |Y_z''| \leq (1+3\delta_1)d|Y'' \setminus Y^r|. \tag{25}$$

Thus Theorem 10.3 applied with $G := G_0$, $H := L_z^{rich} \cap G_0$ and η_1^3 playing the role of η (and so with ν close to αd^2) gives that for every vertex $z \in Z'' \cup Z^{**}$ the probability that M_1 meets less than $(1-\eta_1^2)\alpha d^2 |M_1|$ edges of L_z^{rich} is at most $\mathrm{e}^{-\varepsilon|M_1|}$. Since

$$\begin{aligned} \bar{n} &= |M_1| + |X^{**}| + |X^r| \overset{(18),(22)}{\leq} |M_1| + 2\sqrt{\xi}m + \nu\bar{n} \overset{(20)}{\leq} |M_1| + 10^6\sqrt{\xi}\bar{n} \\ &\overset{(4)}{\leq} |M_1| + \eta_1\bar{n}/2, \end{aligned} \tag{26}$$

we have $(1-\eta_1^2)\alpha d^2|M_1| \geq (1-\eta_1)\alpha d^2\bar{n}$. Thus the probability that (i) fails for M_1 is at most $|Z|\mathrm{e}^{-\varepsilon|M_1|} \leq \mathrm{e}^{-\varepsilon|M_1|/2}$.

Similarly, Proposition 7.16(vi), Lemma 4.1 and Theorem 10.3 (applied with $H := L'_{zz'} \cap G_0$ and so with ν close to $\alpha^2 d^4$) together imply that we may assume that for all but at most $2\xi|Z'|^2 \leq \sqrt{\xi}\bar{n}^2$ pairs $z, z' \in Z'' \cup Z^{**}$ the probability that M_1 meets $L'_{zz'}$ in more than $(1+\eta_1)\alpha^2 d^4\bar{n}$ edges is at most $\mathrm{e}^{-\varepsilon|M_1|}$. This means that with probability at most $|Z|^2\mathrm{e}^{-\varepsilon|M_1|} \leq \mathrm{e}^{-\varepsilon|M_1|/2}$ condition (ii) fails for M_1.

Let us now show that with exponentially small probability (iii) fails for M_1. Fix any $z \in Z'' \cup Z^{**}$. Since by (P1) both P_{XZ} and P_{YZ} are (d,ε')-superregular, (24) and (25) together imply that there are at most $4\varepsilon'|Z| \leq \varepsilon\bar{n}$ vertices $z' \in Z'' \cup Z^{**}$ for which either $(1-\varepsilon')d|X''_z| \leq |X''_z \cap X''_{z'}| \leq (1+\varepsilon')d|X''_z|$ does not hold or for which $(1-\varepsilon')d|Y''_z| \leq |Y''_z \cap Y''_{z'}| \leq (1+\varepsilon')d|Y''_z|$ does not hold. Let $\overline{Z} \subseteq Z'' \cup Z^{**}$ denote the set of all the other vertices in $Z'' \cup Z^{**}$. Then for each vertex $z' \in \overline{Z}$ the subgraph $H_{zz'}$ of G_0 induced by $X''_z \cap X''_{z'}$ and $Y''_z \cap Y''_{z'}$ is still $(d, \sqrt{\varepsilon})$-regular (this holds since G_0 is (d,ε)-superregular). It suffices to show that for each $z' \in \overline{Z}$ the probability that M_1 contains at least $(1+\eta_1)d^4\bar{n}$ edges $xy \in M_1$ for which both xyz and xyz' are hyperedges of $\mathcal{H}_P$ is at most $\mathrm{e}^{-\varepsilon|M_1|}$. But this holds since Theorem 10.3 applied with $G := G_0$ and $H := H_{zz'}$ implies that the probability that M_1 meets $H_{zz'}$ in more than $(1+\eta_1)d^4|M_1|$ edges is at most $\mathrm{e}^{-\varepsilon|M_1|}$. But $(1+\eta_1)d^4|M_1| \leq (1+\eta_1)d^4\bar{n}$.

Proposition 11.2 and Lemma 10.1 together imply that (iv) fails for M_1 with probability at most $\bar{n}^2\mathrm{e}^{-\varepsilon|M_1|} \leq \mathrm{e}^{-\varepsilon|M_1|/2}$. (Apply Lemma 10.1 with $\Delta' := 18\varepsilon$, $G := G_0$ and with the graph consisting of the edges in F_e playing the role of F.)

Let us now consider (v). Proposition 7.6 states that the number of poor edges is at most $2\delta d|X||Y|$. Thus all but at most $2\sqrt{\delta}|X| \leq \delta^{1/3}|X'' \setminus X^r|$ vertices in X'' are incident to at most $\sqrt{\delta}d|Y| \leq \delta^{1/3}d|Y'' \setminus Y^r|$ poor edges. Similarly, by Proposition 7.15(i) all but at most $4\sqrt{\varepsilon'}|X| \leq (\varepsilon')^{1/3}|X'' \setminus X^r|$ vertices in X are incident to at most $\sqrt{\varepsilon'}|Y| \leq (\varepsilon')^{1/3}|Y'' \setminus Y^r|$ unhappy edges. Altogether this implies that at most $2\delta^{1/3}|X'' \setminus X^r|$ vertices in $|X'' \setminus X^r|$ send more that $2\delta^{1/3}d|Y'' \setminus Y^r|$ edges to $Y'' \setminus Y^r$ which are either poor or unhappy. Thus we can apply Lemma 10.1 (where the role of F is played by the subgraph of G_0 consisting of the poor and unhappy edges) to conclude that (v) fails for M_1 with probability at most $\mathrm{e}^{-\varepsilon|M_1|}$. (Recall that $\delta \ll \xi$ by (4).)

Combining everything we proved so far shows that the probability that one of (i)–(v) fails for M_1 is at most $5\mathrm{e}^{-\varepsilon|M_1|/2} < 1$. Thus there is an outcome for M_1 which satisfies (i)–(v). Fix any such M_1 and consider a matching M_2 chosen uniformly at random in $G_0 - M_1$. Since deleting M_1 hardly changes the graph G_0, it follows similarly as before that each of (i)–(v) fails for M_2 with exponentially small probability. Moreover, Lemma 10.2 implies that the probability that (vi) fails is also exponentially small. Thus there exists an outcome for M_2 which satisfies (i)–(vi). □

We now consider the 2-factor $M_1 \cup M_2$. Delete all the poor and all the unhappy edges contained in $M_1 \cup M_2$. Now delete one edge from each remaining cycle in $M_1 \cup M_2$. Lemma 11.3(v)&(vi) together imply that we deleted at most $2\xi\bar{n} + \bar{n}/(\log\bar{n})^{1/5} \leq 3\xi\bar{n}$ edges from $M_1 \cup M_2$. Let $\mathcal{P}$ be the set of all the paths in $M_1 \cup M_2$ obtained in this way together with all the trivial paths consisting of the vertices in $X^{**} \cup Y^{**}$. Note that

$$|\mathcal{P}| \leq |X^{**}| + |Y^{**}| + 3\xi\bar{n} \overset{(18)}{\leq} 5\sqrt{\xi}m. \tag{27}$$

The following lemma shows that we can add a set E' of edges to the paths in $\mathcal{P}$ to obtain a 2-factor K with vertex set $X'' \cup X^{**} \cup Y'' \cup Y^{**}$ which consists of few cycles. (Our desired Hamilton path R^* will be obtained from K by changing a few of its edges.) Moreover, no edge in E' will be poor or unhappy and $E' \cap F_e$ will be small for each edge $e \in P_{X'Y'}$ which is not unhappy.

Lemma 11.4 *There exists a set $E' \subseteq E(P_{X'Y'})$ of edges which satisfies the following properties:*

(i) *Together with the edges in E' the paths in $\mathcal{P}$ form a 2-factor K in $P_{X'Y'}[X'' \cup X^{**} \cup Y'' \cup Y^{**}]$. Thus the vertex classes of K are $X'' \cup X^{**}$ and $Y'' \cup Y^{**}$.*

(ii) *K consists of at most $\bar{n}/(\log \bar{n})^{1/5}$ cycles.*

(iii) *E' avoids all the poor and all the unhappy edges.*

(iv) *$|E' \cap F_e| \leq \varepsilon\bar{n}$ for every edge $e \in P_{X'Y'}$ which is not unhappy.*

(v) *$|E'| \leq 4\nu\bar{n}$.*

The proof of the lemma consists of 2 steps: First we find a set M_{P^r} of independent edges which form a set of paths of odd length (called $\mathcal{P}'$) together with $\mathcal{P}$. The set of vertices lying in some path from $\mathcal{P}'$ will be $X'' \cup X^{**} \cup Y'' \cup Y^{**}$. In the second step we find a perfect matching M' between the sets $S' \subseteq X'' \cup X^{**}$ and $T' \subseteq Y'' \cup Y^{**}$ of endvertices of the paths from $\mathcal{P}'$. Both matchings M_{P^r} and M' will be chosen randomly. We will then set $E' := M_{P^r} \cup M'$.

Proof. Let S be the set of all those endvertices of the paths from $\mathcal{P}$ which lie in $X^{**} \cup (X'' \setminus X^r)$. Similarly, let T be the set of endvertices of the paths from $\mathcal{P}$ in $Y^{**} \cup (Y'' \setminus Y^r)$. Fix a set $\widetilde{S} \subseteq S$ which contains all the vertices in S that are trivial paths in $\mathcal{P}$ as well as one endvertex of every nontrivial path in $\mathcal{P}$ which starts and ends in S (and which therefore has an even number of edges). Define $\widetilde{T} \subseteq T$ similarly. It is easy to see that $|\widetilde{S}| = |\widetilde{T}|$. Also, note that

$$|\widetilde{S}| \leq |\mathcal{P}| \overset{(27)}{\leq} 5\sqrt{\xi}m. \tag{28}$$

Recall that $P^{rich}_{X'Y'}$ was defined in Proposition 7.16(vii) to be the subgraph obtained from $P_{X'Y'}$ by deleting all the poor and unhappy edges. We proved that $P^{rich}_{X'Y'}$ is $(d, 5\delta_0)$-superregular. For convenience, let $P^r := P^{rich}_{X'Y'}[X^r \cup \widetilde{S} \cup Y^r \cup \widetilde{T}]$. Let us first show that the minimum degree of P^r is large. Recall that both X^r and Y^r were obtained by considering a random partition of X'' and Y'' and that in turn X'' and Y'' were obtained by considering a random partition of

X' and Y'. Thus Lemma 4.1 implies that we may assume that the minimum degree of P^r satisfies

$$\begin{aligned}\delta(P^r) &\geq (1-\sqrt{\delta_0})d|X^r| = (1-\sqrt{\delta_0})d|X^r \cup \widetilde{S}| \cdot \frac{|X^r|}{|X^r \cup \widetilde{S}|} \\ &\overset{(23),(28)}{\geq} (1-\sqrt{\delta_0})d|X^r \cup \widetilde{S}| \cdot \frac{10^3\sqrt{\xi}m}{10^3\sqrt{\xi}m + 5\sqrt{\xi}m} \\ &\geq \frac{9}{10}d|X^r \cup \widetilde{S}|.\end{aligned}$$

The reason we set aside the random sets X^r and Y^r at the beginning of this section is precisely that we can guarantee that the graph P^r (as well as the graph P' defined later on) has large minimum degree. This would not necessarily be true if $X^r = Y^r = \emptyset$.

Since $P_{X'Y'}$ is $(d, \sqrt{\varepsilon'})$-superregular by Proposition 7.16(ii) and $|Y^r \cup \widetilde{T}| = |X^r \cup \widetilde{S}| \geq |X^r| \geq 10^3\sqrt{\xi}|X'|$ by (23), its subgraph $P_{X'Y'}[X^r \cup \widetilde{S} \cup Y^r \cup \widetilde{T}]$ is still $(d, \sqrt{\varepsilon'}/(10^3\sqrt{\xi}))$-regular and thus $(d, \sqrt{\varepsilon})$-regular (use that $\sqrt{\varepsilon'}/(10^3\sqrt{\xi}) \leq \sqrt{\varepsilon}$). Since P^r is a subgraph of $P_{X'Y'}[X^r \cup \widetilde{S} \cup Y^r \cup \widetilde{T}]$ we may apply Corollary 10.5 to obtain a $d|X^r \cup \widetilde{S}|/5$-factor in P^r. Denote it by G^r.

For each edge $e \in P_{X'Y'}$ which is not unhappy, we now apply Lemma 10.1 with $\Delta' := 18\varepsilon$ and where G^r plays the role of G and the subgraph of G^r consisting of the edges in $F_e \cap E(G^r)$ plays the role of F. (Recall that F_e was introduced in Definition 11.1.) Together with Proposition 11.2 this guarantees a perfect matching M_{P^r} in P^r such that

$$|M_{P^r} \cap F_e| \leq 170\varepsilon|M_{P^r}| \leq \varepsilon\bar{n}/2 \tag{29}$$

for every edge $e \in P_{X'Y'}$ which is not unhappy. Indeed, the last inequality follows from the fact that

$$|M_{P^r}| = |X^r| + |\widetilde{S}| \overset{(22),(28)}{\leq} \nu\bar{n} + 5\sqrt{\xi}m \overset{(20),(21)}{\leq} 2\nu\bar{n}. \tag{30}$$

We add all the edges in M_{P^r} to the paths in $\mathcal{P}$ and denote by $\mathcal{P}'$ the set of paths thus obtained. Hence all the paths in $\mathcal{P}'$ have odd length. In particular, none of the paths is trival. Let $S' \subseteq X'' \cup X^{**}$ and $T' \subseteq Y'' \cup Y^{**}$ be the sets of endvertices of the paths in $\mathcal{P}'$ (Fig. 6). Note that $X^r \subseteq S'$, $Y^r \subseteq T'$ and

$$|S' \setminus X^r| = |T' \setminus Y^r| \leq |\mathcal{P}| \overset{(27)}{\leq} 5\sqrt{\xi}m. \tag{31}$$

(Indeed, $S' \setminus X^r$ is obtained from S by deleting those vertices in $\widetilde{S}$ which are endpoints of paths in $\mathcal{P}$ having positive even length.) Next we will construct an auxiliary matching M_1' between S' and T' as follows. Label the paths in $\mathcal{P}'$ from 1 to $|\mathcal{P}'|$. This induces a labelling of the vertices in S' from 1 to $|\mathcal{P}'|$. Since all the paths in $\mathcal{P}'$ have (positive) odd length, every label occurs precisely once in S'. Similarly, we have a labelling of T'. This gives us a perfect matching M_1' between S' and T' (we match the vertices with the same label). Thus each edge in M_1' joins the two endvertices of some path in M_1' and vice versa.

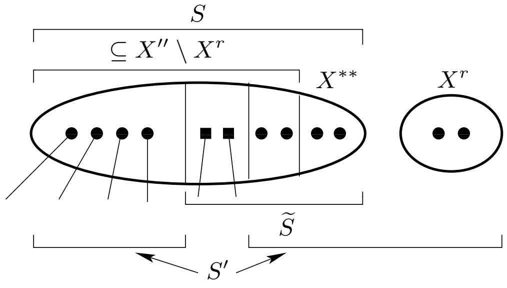

Figure 6: The vertex class S' of the matching M' and the class $X^r \cup \widetilde{S}$ of the matching M_{P^r}. The 'square' vertices are endpoints of paths of positive even length in $\mathcal{P}$ and are not contained in S'. The edges which are indicated all lie in paths from $\mathcal{P}$.

Our aim now is to find a perfect matching M' between S' and T' which avoids both the paths in $\mathcal{P}'$ and the matching M'_1 and which is such that the 2-factor obtained by adding the edges in M' to the paths in $\mathcal{P}'$ has the properties required in the lemma. (The set E' in the lemma will then be $M_{P^r} \cup M'$.) The matching M' will be chosen at random inside a regular spanning subgraph of $P' := P^{rich}_{X'Y'}[S' \cup T'] - (M'_1 \cup E(\bigcup \mathcal{P}'))$. To find this regular spanning subgraph, we will apply Corollary 10.5 again. Thus we first need to show that P' contains a spanning regular subgraph whose degree is sufficiently large. But this follows similarly as the analogous assertion for P^r which we proved at the beginning of the proof of Lemma 11.4. Indeed, together with (31) a similar argument as there shows that P' contains a spanning $d|S'|/5$-factor P''.

Let M' be a perfect matching in P'' chosen uniformly at random. Let K denote the 2-factor obtained from $\mathcal{P}'$ by adding the edges in M' and put $E' := M_{P^r} \cup M'$. Then K satisfies (i) and E' satisfies (iii). The latter holds since both M_{P^r} and M' consist of edges in $P^{rich}_{X'Y'}$ and thus contain neither poor nor unhappy edges by the definition of $P^{rich}_{X'Y'}$ (cf. Proposition 7.16(vii)). Also (v) is satisfied since

$$|E'| = |M_{P^r}| + |M'| = |M_{P^r}| + |S'| \overset{(31)}{\leq} |M_{P^r}| + |X^r| + 5\sqrt{\xi}m \overset{(22),(30)}{\leq} 2\nu\bar{n} + \nu\bar{n} + 5\sqrt{\xi}m \overset{(20),(21)}{\leq} 4\nu\bar{n}.$$

Moreover, if we apply Lemma 10.1 in the same way as in the proof of (29), we see that (iv) holds with probability at least 3/4.

We now wish to prove that with large probability M' is such that (ii) holds. Lemma 10.2 with $G := P''$ applied to M'_1 and the random perfect matching M' shows that with probability at least 3/4 they form a 2-factor in the graph $P'' \cup M_1$ which has most $|S'|/(\log|S'|)^{1/5} \leq \bar{n}/(\log \bar{n})^{1/5}$ cycles. Since for each path in $\mathcal{P}'$ there is an edge in M'_1 which joins the two endvertices of that path, this immediately implies that with probability at least 3/4 the graph $K = (\bigcup \mathcal{P}') \cup M'$ contains at most $\bar{n}/(\log \bar{n})^{1/5}$ cycles. Thus altogether this shows that the probability that M' (and thus K and E') has the desired properties is greater than zero. □

Recall that

$$G^{rich} := P^{rich}_{X'Y'}[X'' \cup X^{**} \cup Y'' \cup Y^{**}].$$

So K is a 2-factor of G^{rich} as K avoids all the poor and unhappy edges. Since X'' and Y'' were obtained by a considering a random partition of X' and Y', Lemma 4.1 and 7.16(vii) together imply that we may assume that for every vertex of G^{rich}, the number of neighbours is at least $(1-6\delta_0)d|X''|$, which is at least $99d\bar{n}/100$ by (21). In particular

$$\delta(G^{rich}) \geq 99d\bar{n}/100. \tag{32}$$

Our aim now is to transform the 2-factor K guaranteed by Lemma 11.4 into a cycle on the same set of vertices which meets neither poor nor unhappy edges. Moreover, it will be important that this can be done by changing at most $O(\bar{n}/(\log \bar{n})^{1/5}$ edges. This would be quite simple if we knew that G^{rich} is (d, ε)-regular, say. Unfortunately, we cannot guarantee this and so we will proceed as follows. Choose any cycle of K, delete an edge on this cycle and let R denote the path thus obtained. If one of the endvertices of R, x say, has a G^{rich}-neighbour, z say, on some other cycle C in K, then we extend R by adding the edge xz as well as C and by deleting one of the two edges on C incident to z. Note that, since all the cycles in K have even length, the extension of R obtained in this way has odd length, just as R.

We continue in this fashion until we have obtained a path R whose endvertices $x \in X'' \cup X^{**}$ and $y \in Y'' \cup Y^{**}$ have all their G^{rich}-neighbours on R. We view R as being directed from x to y. Let S_x^- denote the set of predecessors of the G^{rich}-neighbours of x on R and let S_y^+ denote the set of successors of the G^{rich}-neighbours of y on R. If one of the vertices in $S_x^- \cup S_y^+$ has a G^{rich}-neighbour on some cycle C of $K - R$, then it is easy to verify that we may extend R further by incorporating the vertices of C. Thus we may assume that $N_{G^{rich}}(S_x^-), N_{G^{rich}}(S_y^+) \subseteq V(R)$. The following lemma implies that $|N_{G^{rich}}(S_x^-)| \geq \bar{n}/2$. (To see this, use that S_x^- lies in the vertex class $X'' \cup X^{**}$ of G^{rich} and $|S_x^-| \geq \delta(G^{rich}) \geq 99d\bar{n}/100$. Take I to be the entire other vertex class of G^{rich}.) Thus

$$|R| \geq \bar{n}.$$

Lemma 11.5 *Suppose that A and I are subsets of $V(G^{rich})$ which lie in different vertex classes of G^{rich}. Furthermore, suppose that $|A| \geq d\bar{n}/10^5$ and $|I| \geq \bar{n}/10^5$. Then $|N_{G^{rich}}(A) \cap I| \geq |I|/2$.*

Proof. It suffices to consider the case when $A \subseteq X'' \cup X^{**}$ and $I \subseteq Y'' \cup Y^{**}$. The argument for the other case is analogous. Recall that by Proposition 7.16(ii) the graph $P_{X'Y'}$ is $(d, \sqrt{\varepsilon'})$-superregular and thus (d, ε)-superregular by (20). Hence in the graph $P_{X'Y'}$, we know that for all but at most $2\varepsilon|X'| \leq \sqrt{\varepsilon}|A|$ vertices $a \in A$ the number of neighbours of a in I lies between $(1-\varepsilon)d|I|$ and $(1+\varepsilon)d|I|$. Of course, some of these neighbours may not be neighbours of a in G^{rich}. However, since by Proposition 7.16(vii) the graph $P^{rich}_{X'Y'}$ is $(d, 5\delta_0)$-superregular and $P_{X'Y'}$ is (d, ε)-superregular, every vertex in A is incident to at most $(5\delta_0 + \varepsilon)d|Y'|$ edges in $E(P_{X'Y'}) \setminus E(P^{rich}_{X'Y'})$. Note that

$(5\delta_0+\varepsilon)d|Y'| \leq \sqrt{\delta_0}d|I|$ by (21). Thus for all but at most $\sqrt{\varepsilon}|A|$ vertices in A the number of G^{rich}-neighbours in I is at least $(1-2\sqrt{\delta_0})d|I|$. Hence the number $e_{rich}(A,I)$ of edges between A and I in G^{rich} satisfies

$$e_{rich}(A,I) \geq (1-\sqrt{\varepsilon})|A|(1-2\sqrt{\delta_0})d|I|.$$

On the other hand, the (d,ε)-regularity of $P_{X'Y'}$ implies that

$$e_{rich}(A,I) \leq |A||N_{G^{rich}}(A)\cap I|(1+\varepsilon)d.$$

Combining the above two inequalities gives the desired result. □

The next lemma implies that either our path R can be made into a cycle on the same vertex set by changing at most 9 edges or else R can be enlarged further by adding another cycle from the 2-factor K and changing at most 6 edges.

Lemma 11.6 *Suppose that R is a path of odd length in G^{rich} such that all the G^{rich}-neighbours of its endvertices x and y lie on R, such that $K-V(R)$ consists of cycles and such that $|R|\geq \bar{n}$. Then one of the following holds:*

(i) *G^{rich} contains a cycle C whose vertex set is $V(R)$ and for which $|E(R)\triangle E(C)|\leq 9$.*

(ii) *G^{rich} contains a path R' of odd length which is obtained from the union of R with some cycle $C\in K-V(R)$ by adding at most 3 edges in $E(G^{rich})\setminus (E(R)\cup E(C))$, deleting one edge on C and deleting at most 2 edges on R. In particular, $K-V(R)$ consists of cycles.*

Proof. Let J be a minimal initial segment of the path R when directed from x to y such that $|J\cap N_{G^{rich}}(x)|\geq |N_{G^{rich}}(x)|/2$ or $|J\cap N_{G^{rich}}(y)|\geq |N_{G^{rich}}(y)|/2$. Denote by $\overline{J}$ the segment of R which consists of all edges not in J.

We first consider the case when $|J|\geq|\overline{J}|$ and $|J\cap N_{G^{rich}}(x)|\geq |N_{G^{rich}}(x)|/2$. The minimality of J implies that $|\overline{J}\cap N_{G^{rich}}(y)|\geq \lfloor |N_{G^{rich}}(y)|/2\rfloor$. Let I be a minimal initial segment of J such that $|I\cap N_{G^{rich}}(x)|\geq |N_{G^{rich}}(x)|/4$ or $|I|\geq |J|/2$. Let $\overline{I}$ denote the segment of J which consists of all the edges not in I.

Let us first assume that $|I|\geq |J|/2$ and thus $|I|\geq |R|/4\geq \bar{n}/4$. Then the minimality of I implies that $|\overline{I}\cap N_{G^{rich}}(x)|\geq \lfloor |N_{G^{rich}}(x)|/4\rfloor$. Let $\overline{I}_x^-$ denote the set of all those vertices in $\overline{I}$ which are predecessors of vertices in $N_{G^{rich}}(x)$. Similarly, let Let $\overline{J}_y^+$ denote the set of all those vertices in $\overline{J}$ which are successors of vertices in $N_{G^{rich}}(y)$. We may assume that all the G^{rich}-neighbours of the vertices in $\overline{I}_x^-\cup\overline{J}_y^+$ lie on R since otherwise we could modify R to obtain a path R' as in (ii). (Indeed, if for example a vertex $v\in\overline{I}_x^-$ has a G^{rich}-neighbour w on the cycle $D\subseteq K-V(R)$, then we take for R' the path obtained from $R\cup D$ by adding the edge vw and the edge between x and the sucessor v^+ of v on R as well as by deleting the edge vv^+ and one of the edges on D incident with w.)

Partition I into two segments I_1 and I_2 whose size is as equal as possible such that I_1 is an inital segment of I. Lemma 11.5 now implies that

$$|N_{G^{rich}}(\overline{I}_x^-) \cap I_1| \geq \bar{n}/100$$

and

$$|N_{G^{rich}}(\overline{J}_y^+) \cap I_2| \geq \bar{n}/100.$$

Indeed, to see for example the first inequality, apply Lemma 11.5 with $\overline{I}_x^-$ playing the role of A and $I_1 \cap (Y'' \cup Y^{**})$ playing the role of I. This can be done since $|\overline{I}_x^-| \geq |N_{G^{rich}}(x) \cap \overline{I}| - 1 \geq |N_{G^{rich}}(x)|/8 \geq \delta(G^{rich})/8 \geq d\bar{n}/10$ by (32) and

$$|I_1 \cap (Y'' \cup Y^{**})| \geq \lfloor |I_1|/2 \rfloor \geq |I|/8 \geq \bar{n}/50.$$

Since $\bar{n}/10^2 \overset{(21)}{\geq} m/10^4 \geq 3|X'|/10^4$ and since $P^{rich}_{X'Y'}$ is $(d, 5\delta_0)$-regular, it follows that $P^{rich}_{X'Y'}$ contains an edge vw such that v is the predecessor of some vertex $v' \in N_{G^{rich}}(\overline{I}_x^-) \cap I_1$ on R and w is the successor of some vertex $w' \in N_{G^{rich}}(\overline{J}_y^+) \cap I_2$. Then $vw \in G^{rich}$ and it is easy to see that the path R can be modified to obtain a cycle C as in (i) (Fig. 7).

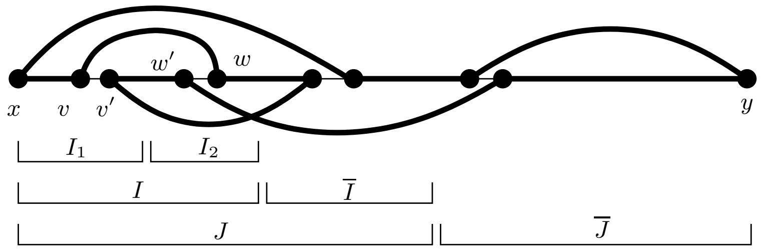

Figure 7: Modifying the path R to obtain a cycle C in the first case of the proof of Lemma 11.6.

The proofs of all the remaining cases are similar. For instance in the case when $|J \cap N_{G^{rich}}(y)| \geq |N_{G^{rich}}(y)|/2$ and $|J| \geq |\overline{J}|$, we now consider the successors of the G^{rich}-neighbours of x in $\overline{J}$ instead of the predecessors. Again, we can assume that all these successors have all their G^{rich}-neighbours on R since otherwise R could be modified into a path R' satisfying (ii). But now, this path R' would also contain a new edge incident to y (apart from 2 other new edges, see Figure 8).

Different types for the cycle C that can be obtained from a modification of R are shown in Figure 9. All remaining types can be obtained by exchanging x and y in Figures 7 and 9. (For example, the first situation in Figure 9 corresponds to the case when $|J \cap N_{G^{rich}}(x)| \geq |N_{G^{rich}}(x)|/2$, $|J| \geq |\overline{J}|$ and $|I| < |J|/2$.) $\square$

We apply Lemma 11.6 repeatedly until we are in Case (i) of the lemma. Thus we have obtained a cycle C in G^{rich} whose length is at least $|R| \geq \bar{n}$ and for which $K - V(C)$ is a union of cycles. The following lemma implies that if C is

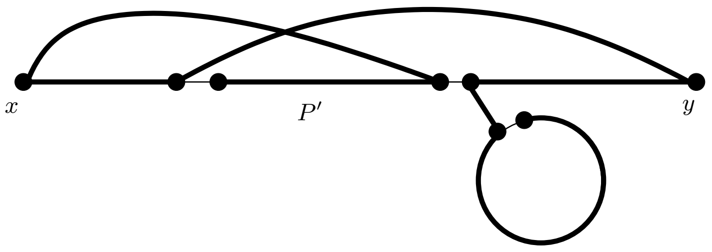

Figure 8: The path R' in the case when $|J \cap N_{G^{rich}}(y)| \geq |N_{G^{rich}}(y)|/2$ and $|J| \geq |\overline{J}|$.

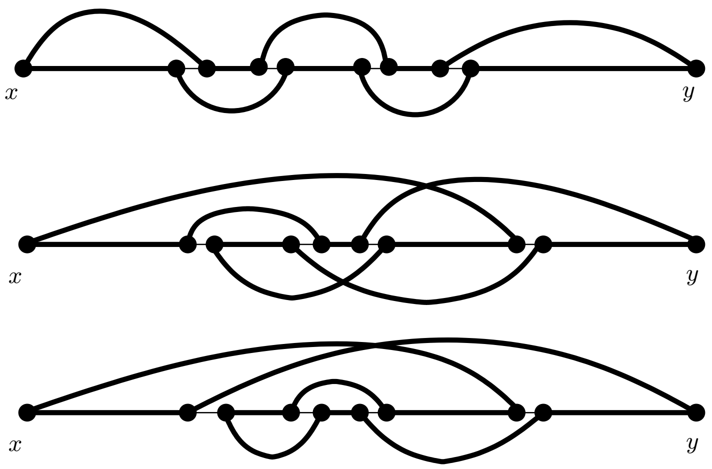

Figure 9: Different types of cycles C which can be obtained by a modification of R.

not yet a Hamilton cycle in G^{rich}, then some cycle D in $K - V(C)$ must send an edge to C.

Lemma 11.7 *Suppose that C is a cycle of length at least $\bar{n}$ in G^{rich} such that $K - V(C)$ is non-empty and consists of cycles. Then G^{rich} contains an edge between C and some cycle in $K - V(C)$.*

Proof. Suppose not and put $U := V(G^{rich}) \setminus V(C) = V(K) \setminus V(C)$. Then $|U| \geq \delta(G^{rich}) \geq 99d\bar{n}/100$ by (32). But U contains the same number of vertices in each vertex class of G^{rich}; and the same is true for C. Thus Lemma 11.5 applied with $A := U \cap (X'' \cup X^{**})$ and $I := C \cap (Y'' \cup Y^{**})$ implies the existence of an edge in G^{rich} between U and C, a contradiction. $\square$

If our cycle C is not yet a Hamilton cycle, we apply Lemma 11.7 to find an edge in G^{rich} between C and some cycle D in $K \setminus V(C)$. We then consider any path obtained from $C \cup D$ by adding this edge and deleting two suitable edges. We again call this path R. We then apply Lemma 11.6 repeatedly until we obtain a new cycle. If this cycle is not yet a Hamilton cycle in G^{rich} we apply

Lemma 11.7 again and continue. Continuing in this way we eventually obtain a Hamilton cycle of G^{rich}, which we again denote by C.

We will now modify C slightly to obtain a Hamilton path R^* in G^{rich} which will be as described at the beginning of Section 11. To do this, we need the following lemma. It will enable us to transform the Hamilton cycle C into a Hamilton path R^* which 'attaches' to the bridge vertices x^* and y^* (see Lemma 11.9(i)).

Lemma 11.8 *Let C be a Hamilton cycle in G^{rich} and let $U, W \subseteq V(G^{rich})$ be two sets of size at least $d\bar{n}/100$ which are contained in different vertex classes of G^{rich}. Then G^{rich} has a Hamilton path R which starts in U, ends in W and contains only 3 edges outside C.*

Proof. Without loss of generality we may assume that $U \subseteq Y'' \cup Y^{**}$ and $W \subseteq X'' \cup X^{**}$. It is easy to find disjoint sets $U_1 \subseteq U$, $W_1 \subseteq W$, $I \subseteq Y'' \cup Y^{**}$ and $J \subseteq X'' \cup X^{**}$ such that $|U_1|, |W_1| \geq d\bar{n}/400$ and $|I|, |J| \geq \bar{n}/16$ and such that, for a suitable orientation $\overrightarrow{C}$ of C, the vertices in U_1, W_1, I and J occur on $\overrightarrow{C}$ in the following order: first the vertices in U_1, then the vertices in W_1, then the vertices in J and then those in I.

Let U_1' (respectively W_1') be the set of all successors of vertices in U_1 (respectively W_1) on $\overrightarrow{C}$. Lemma 11.5 implies that

$$|N_{G^{rich}}(U_1') \cap I| \geq |I|/2 \geq \bar{n}/100 \overset{(21)}{\geq} m/10^4 \geq 3|Y'|/10^4$$

and similarly that

$$|N_{G^{rich}}(W_1') \cap J| \geq 3|X'|/10^4.$$

Together with the fact that $P^{rich}_{X'Y'}$ is $(d, 5\delta_0)$-regular, this implies that $P^{rich}_{X'Y'}$ contains an edge $u''w''$ such that u'' is the successor of some vertex $u' \in N_{G^{rich}}(U_1') \cap I$ and w'' is the predecessor of some vertex $w' \in N_{G^{rich}}(W_1') \cap J$. In particular, $u''w'' \in G^{rich}$. Let u^+ (respectively w^+) be a G^{rich}-neighbour of

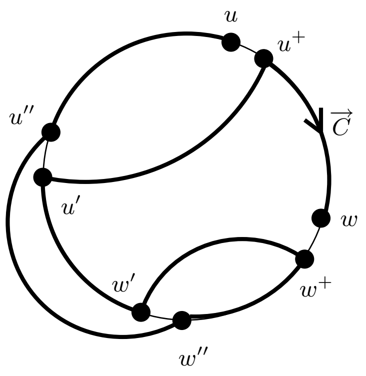

Figure 10: Modifying the Hamilton cycle C to obtain R^*.

u' (respectively w') in U_1' (respectively W_1'). Let u and w be the predecessors of u^+ and w^+. Then $P := u\overleftarrow{C}u''w''\overleftarrow{C}w^+w'\overrightarrow{C}u'u^+\overrightarrow{C}w$ is a Hamilton path as required (Fig. 10). $\square$

Recall that the bridge vertices x^* and y^* are (still) useful and that the link graph L'_{x^*} was defined in Lemma 7.16(v). Put $Y''_{x^*} := Y'' \cap V(L'_{x^*}) = Y'' \cap N_{P_{XY}}(x^*)$, $Z''_{x^*} := Z'' \cap V(L'_{x^*}) = Z'' \cap N_{P_{XZ}}(x^*)$ and let L''_{x^*} be the subgraph of L'_{x^*} induced by Y''_{x^*} and Z''_{x^*}. Define X''_{y^*}, Z''_{y^*} and L''_{y^*} similarly. Since x^* is a useful vertex, the graph L'_{x^*} is $(\alpha d, 4\xi)$-regular and its vertex class Y'_{x^*} satisfies $|Y'_{x^*}| \geq (1-\varepsilon')d|Y'|$ (cf. Definitions 7.17 and 7.18). Since Y'' was obtained by considering a random subset of Y' of size $m/100 \geq 3|Y'|/100$ and since Z'' was obtained by considering a random subset of Z' of size $2m/100 \geq 3|Z'|/100$, Lemma 4.1 implies that we may assume that L''_{x^*} is still $(\alpha d, 150\xi)$-regular and $|Y''_{x^*}| \geq (1-\varepsilon)d|Y''|$. Similarly, we may assume that L''_{y^*} is $(\alpha d, 150\xi)$-regular and $|X''_{y^*}| \geq (1-\varepsilon)d|X''|$.

Let $Y(x^*)$ be the set of all those vertices in Y''_{x^*} which have at least $\alpha d|Z''_{x^*}|/2$ neighbours in Z''_{x^*}. Define $X(y^*) \subseteq X''_{y^*}$ similarly. Since both L''_{x^*} and L''_{y^*} are $(\alpha d, 150\xi)$-regular we have

$$|Y(x^*)| \geq (1-150\xi)|Y''_{x^*}| \geq (1-150\xi)(1-\varepsilon)d|Y''| \geq d\bar{n}/2$$

and

$$|X(y^*)| \geq (1-150\xi)|X''_{y^*}| \geq (1-150\xi)(1-\varepsilon)d|X''| \geq d\bar{n}/2.$$

Lemma 11.8 now implies that our Hamilton cycle C of G^{rich} can be modified into a Hamilton path R^* of G^{rich} which starts in some vertex $y'' \in Y(x^*)$, ends in some vertex $x'' \in X(y^*)$ and contains only 3 edges outside C. By definition of $Y(x^*)$ and $X(y^*)$ we can find vertices $z^*_x \in Z''_{x^*}$ and $z^*_y \in Z''_{y^*}$ such that $z^*_x \neq z^*_y$, $y''z^*_x \in L''_{x^*}$ and $x''z^*_y \in L''_{y^*}$. But this means that both $x^*y''z^*_x$ and $y^*x''z^*_y$ are hyperedges of $\mathcal{H}_P \subseteq \mathcal{H}$. For later reference, we summarize the properties of R^* in the following lemma.

Lemma 11.9 *The path R^* has the following properties:*

(i) *R^* is a Hamilton path in G^{rich}. Thus the vertex set of R^* is $X'' \cup X^{**} \cup Y'' \cup Y^{**}$. The endvertices of R^* are $x'' \in X''$ and $y'' \in Y''$. There are two distinct vertices $z^*_x, z^*_y \in Z''$ such that both $x^*y''z^*_x$ and $y^*x''z^*_y$ are hyperedges of $\mathcal{H}$.*

(ii) *R^* avoids all the poor and all the unhappy edges.*

(iii) $|E(R^*) \setminus (M_1 \cup M_2)| \leq 10^6\sqrt{\xi}\bar{n}$.

(iv) *All but at most $20\bar{n}/(\log \bar{n})^{1/5}$ edges in $(M_1 \cup M_2) \setminus E(R^*)$ are poor or unhappy.*

(v) $|E(R^*) \cap F_e| \leq 350\varepsilon\bar{n}$ *for every edge $e \in F_e$ which is not unhappy.*

Proof. Properties (i) and (ii) are clear from the construction of R^*. To check (iii), recall that R^* was obtained from the union $M_1 \cup M_2$ of the random matchings M_1 and M_2 as follows. We first deleted some edges from $M_1 \cup M_2$ to obtain the set of paths $\mathcal{P}$. As noted in the paragraph before (27), in this step we deleted at most $2\xi\bar{n}$ edges which are poor or unhappy and at most $\bar{n}/(\log \bar{n})^{1/5}$ further edges. We then applied Lemma 11.4 to add a set E' of at most $4\nu\bar{n} = 4\cdot 10^5\sqrt{\xi}\bar{n}$

edges to obtain a 2-factor K with at most $\bar{n}/(\log \bar{n})^{1/5}$ cycles. Finally, we changed at most $19\bar{n}/(\log \bar{n})^{1/5}$ edges of K to obtain R^* (with room to spare in the constant). So altogether R^* contains at most $10^6\sqrt{\xi}\bar{n}$ edges not in $M_1 \cup M_2$, as required in (iii). Moreover, the above argument shows that (iv) holds too. To verify (v), use that $E(R^*)$ is obtained from a subset of $M_1 \cup M_2 \cup E'$ by adding at most $19\bar{n}/(\log \bar{n})^{1/5}$ edges. Thus (v) follows immediately from Lemmas 11.3(iv) and 11.4(iv). □

11.2 Finding the loose path $\mathcal{Q}^*$

The aim of this section is to show that there exists a loose path $\mathcal{Q}^*$ as described at the beginning of Section 11. We denote by Z''' the subset obtained from $Z'' \cup Z^{**}$ by deleting the vertices z_x^* and z_y^* guaranteed in Lemma 11.9(i). We consider the following bipartite auxiliary graph H^*. The vertex classes of H^* are Z''' and $M := E(R^*)$. Two vertices $z \in Z'''$ and $xy \in M$ are joined by an edge in H^* if $xyz \in \mathcal{H}_P$. (Recall that $\mathcal{H}_P$ was defined in Definition 7.2.) Put

$$n_* := 2\bar{n} - 1.$$

Note that

$$|Z'''| = |M| = n_*$$

since the definition of X'', Y'', Z'' in Section 8.1 together with Lemma 9.1 imply that

$$|Z'' \cup Z^{**}| = 2|X'' \cup X^{**}| + 1 = 2|Y'' \cup Y^{**}| + 1 = |E(R^*)| + 2 = |M| + 2.$$

To prove the existence of the desired loose path $\mathcal{Q}^*$, it suffices to show that H^* has a perfect matching. To show the latter, we will check that H^* satisfies Hall's condition. The intuition is that H^* satisfies Hall's condition because it looks like a random graph (and thus for instance most pairs of vertices do not have many neighbours in common). We put

$$M' := (M_1 \cup M_2) \cap M \quad \text{and} \quad \overline{Z}'' := Z'' \cap Z''' = Z'' \setminus \{z_x^*, z_y^*\}.$$

(Recall that M_1 and M_2 were matchings as in Lemma 11.3.)

Lemma 11.10 *The bipartite graph $H^* = (Z''', M)$ satisfies the following properties.*

(i) *For every vertex $v \in M$ we have $|N_{H^*}(v) \cap \overline{Z}''| \geq (1-\eta)\alpha d^2 n_*$. Similarly, for every vertex $z \in Z'''$ we have $|N_{H^*}(v) \cap M'| \geq (1-\eta)\alpha d^2 n_*$. In particular, the minimum degree of H^* is at least $(1-\eta)\alpha d^2 n_*$.*

(ii) *All but at most $\sqrt{\xi} n_*^2$ pairs $z, z' \in Z'''$ satisfy $|N_{H^*}(z) \cap N_{H^*}(z') \cap M'| \leq (1+\eta)\alpha^2 d^4 |M'|$.*

(iii) *For every $z \in Z'''$ there are at most $2\varepsilon n_*$ other vertices $z' \in Z'''$ for which $|N_{H^*}(z) \cap N_{H^*}(z') \cap M'| \geq (1+\eta)d^4|M'|$.*

(iv) *For every $v \in M$ there are at most $200\varepsilon n_*$ other vertices $v' \in M$ for which* $|N_{H^*}(v) \cap N_{H^*}(v') \cap \overline{Z}''| \geq (1+\eta)d^4|\overline{Z}''|$.

(v) $|M| \leq |M'| + 10^6\sqrt{\xi}n_*$.

Proof. To prove (i), recall that by Lemma 11.9(ii) no edge of R^* is poor or unhappy. Thus, by Definition 7.14, for every edge xy of R^* there are at least $(1-\eta^2)\alpha d^2|Z|$ vertices $z \in Z'$ such that xyz is a hyperedge of $\mathcal{H}_P$. Since Z'' was obtained by considering a random partition of Z', together with Lemma 4.1 this shows that we may assume that there are least $(1-\eta^{3/2})\alpha d^2|Z''|$ vertices $z \in Z''$ such that xyz is a hyperedge of $\mathcal{H}_P$. Since $|\overline{Z}''| = |Z''| - 2$, this implies in that H^* all the vertices in $M = E(R^*)$ have at least $(1-\eta^{3/2})\alpha d^2|Z''| - 2$ neighbours in $\overline{Z}''$. But $(1-\eta^{3/2})\alpha d^2|Z''| - 2 \geq (1-\eta)\alpha d^2 n_*$ since $|Z^{**}| = 4\sqrt{\xi}m + 1$ by Lemma 9.1 and thus

$$n_* = |Z'''| = |Z''| + |Z^{**}| - 2 \leq |Z''| + 4\sqrt{\xi}m = (1 + 200\sqrt{\xi})|Z''|$$

(use that $|Z''| = m/50$ and $\xi \ll \eta$ by (4)).

Now consider any vertex $z \in Z'''$. By Lemma 11.3(i), $M_1 \cup M_2 \supseteq M'$ meets the link graph L_z^{rich} in at least $2(1-\eta_1)\alpha d^2\bar{n}$ edges. Recall that by definition, L_z^{rich} does not contain poor or unhappy edges. Thus Lemma 11.9(iv) implies that all but at most $20\bar{n}/(\log \bar{n})^{1/5}$ edges in $E(L_z^{rich}) \cap (M_1 \cup M_2)$ belong to $M' \subseteq M = E(R^*)$. But every edge in $E(L_z^{rich}) \cap M'$ corresponds to a neighbour of z in H^*. Thus in H^* the vertex z has at least $2(1-\eta_1)\alpha d^2\bar{n} - 20\bar{n}/(\log \bar{n})^{1/5} \geq (1-\eta)\alpha d^2 n_*$ neighbours in M'. (Recall that $\eta_1 \ll \eta$ by (4).) This completes the proof of (i).

Properties (ii) and (iii) immediately follow from Lemma 11.3(ii) and (iii) respectively. To check (iv), recall that every vertex $v \in M$ stands for an edge $e \in P_{X'Y'}$ which is not unhappy. Lemma 11.9(v) implies that the set M meets F_e in at most $350\varepsilon\bar{n} \leq 200\varepsilon n_*$ edges. But if $e' := v' \in M$ is such that $|N_{H^*}(v) \cap N_{H^*}(v') \cap \overline{Z}''| \geq (1+\eta)d^4|\overline{Z}''|$ then either e' is one of the at most 2 edges in $E(R^*) = M$ incident to e or else $e' \in F_e$ by Definition 11.1 and the fact that $\xi \ll \eta$ by (4). Finally, (v) follows immediately from Lemma 11.9(iii). $\square$

Proposition 11.11 *Let $I \subseteq M$ be a set of size at least $(1-\eta)\alpha d^2 n_*$. Then $|N_{H^*}(I)| \geq \alpha^2 n_*/4$. Similarly, if $\widetilde{I} \subseteq Z'''$ is a set of size at least $(1-\eta)\alpha d^2 n_*$ then $|N_{H^*}(\widetilde{I})| \geq \alpha^2 n_*/4$.*

Proof. Put $t := \lceil \alpha/(2d^2) \rceil$. Lemma 11.10(iv) implies that I contains a set J of size t such that

$$|N_{H^*}(v) \cap N_{H^*}(v') \cap \overline{Z}''| \leq (1+\eta)d^4|\overline{Z}''| \leq (1+\eta)d^4 n_*$$

for all $v \neq v'$ in J. (Indeed, $t \cdot 200\varepsilon n_* \ll |I|$ since $\varepsilon \ll d \ll \alpha_*$ by (4) and since $\alpha \geq \alpha_*/2$ by (P2) in Section 7.2. Thus, using Lemma 11.10(iv), we can find such a set J greedily.) Together with Lemma 11.10(i) this shows that

$$
\begin{aligned}
|N_{H^*}(I)| &\geq |N_{H^*}(J) \cap \overline{Z}''| \geq t \cdot (1-\eta)\alpha d^2 n_* - \binom{t}{2}(1+\eta)d^4 n_* \\
&= td^2 n_* \left((1-\eta)\alpha - \frac{t-1}{2}(1+\eta)d^2 \right) \\
&\geq \frac{\alpha}{2} n_* \left((1-\eta)\alpha - \frac{\alpha}{4}(1+\eta) \right) \geq \frac{\alpha^2 n_*}{4},
\end{aligned}
$$

as required. The proof for $\widetilde{I}$ is similar except that we now apply Lemma 11.10(iii) instead of (iv). $\square$

Proposition 11.12 *Let $I \subseteq M$ be a set of size at least $\alpha^2 n_*/4$ such that $|N_{H^*}(I)| \leq (1-8\eta)n_*$. Then $|N_{H^*}(I)| \geq |I|$.*

Proof. Set $I' := I \cap M'$. Since $|M| \leq |M'| + 10^6\sqrt{\xi} n_*$ by Lemma 11.10(v), we have

$$
|I'| \geq |I| - 10^6\sqrt{\xi} n_* \geq |I| - 10^6\sqrt{\xi} \cdot 4|I|/\alpha^2 \geq (1 - 10^7\sqrt{\xi}/\alpha^2)|I| \overset{(4)}{\geq} (1-\eta)|I|. \tag{33}
$$

For the last inequality recall also that $\alpha \geq \alpha_*/2$ by (P2). Put $N := N_{H^*}(I')$. We now double count the paths of length two in H^* whose endvertices lie in N and whose midpoint lies in I'. Let f denote the number of these paths. Lemma 11.10(i) implies that

$$
f \geq \sum_{v \in I'} \binom{|N_{H^*}(v)|}{2} \geq |I'| \binom{(1-\eta)\alpha d^2 n_*}{2} \geq (1-2\eta)\alpha^2 d^4 n_*^2 |I'|/2.
$$

On the other hand, using Lemma 11.10(ii) and (iii) we get

$$
\begin{aligned}
f &\leq \sum_{z,z' \in N} |N_{H^*}(z) \cap N_{H^*}(z') \cap M'| \\
&\leq (1+\eta)\alpha^2 d^4 |M'| \binom{|N|}{2} + \sqrt{\xi} n_*^2 (1+\eta) d^4 |M'| + 2\varepsilon n_*^2 |M'| \\
&\leq (1+2\eta)\alpha^2 d^4 |M'| \frac{|N|^2}{2} \\
&\leq (1+2\eta)\alpha^2 d^4 n_* \frac{(1-8\eta) n_* |N|}{2}.
\end{aligned}
$$

(Indeed, to verify the 3rd line use that $\varepsilon \ll d \ll \xi \ll \eta \ll \alpha_*$ by (4), $\alpha_*/2 \leq \alpha$ by (P2) and $|N| \geq \alpha^2 n_*/4$ by Proposition 11.11. The last line follows since we assumed that $|N_{H^*}(I)| \leq (1-8\eta)n_*$.) The lower and upper bound for f together imply that

$$
|I| \overset{(33)}{\leq} \frac{1}{1-\eta}|I'| \leq \frac{(1+2\eta)(1-8\eta)}{(1-3\eta)}|N| \leq |N| \leq |N_{H^*}(I)|,
$$

as required. $\square$

We now can combine Lemma 11.10(i) with Propositions 11.11 and 11.12 to conclude that H^* has a perfect matching. As mentioned already at the beginning of Section 11.2, this implies the existence of our desired loose path $\mathcal{Q}^*$ and thus completes the proof of Theorem 1.1.

Lemma 11.13 *The graph H^* contains a perfect matching.*

Proof. We show that H^* satisfies Hall's condition. Thus consider any $I \subseteq M$. We have to show that $|N_{H^*}(I)| \geq |I|$. Clearly, this holds if $|I| \leq \delta(H^*)$ or $|I| > n_* - \delta(H^*)$. Together with Lemma 11.10(i) this shows that we may assume that $(1-\eta)\alpha d^2 n_* \leq |I| \leq (1-(1-\eta)\alpha d^2)n_*$. But then Proposition 11.11 applied to I and to $\widetilde{I} := Z''' \setminus N_{H^*}(I)$ shows that Hall's condition holds if $|I| \leq \alpha^2 n_*/4$ or $|I| > (1-\alpha^2/4)n_*$. Thus we may assume that $|I| \geq \alpha^2 n_*/4$ and $|N_{H^*}(I)| \leq (1-\alpha^2/4)n_* \leq (1-8\eta)n_*$ (to see the latter inequality use that $\eta \ll \alpha_*$ by (4) and $\alpha_*/2 \leq \alpha$ by (P2)). But now Proposition 11.12 implies that Hall's condition also holds in this case. $\square$

12 Perfect packings

12.1 Proof of Theorem 1.3

In this subsection, we briefly describe the modifications to the proof of Theorem 1.1 that need to be made in order to obtain Theorem 1.3. For convenience, given a copy of $\mathcal{C}_4$, we say that the *base pair* of this copy consists of those two vertices which lie in both of its hyperedges.

The first point in the proof of Theorem 1.3 where we need to proceed differently than in the proof of Theorem 1.1 is Section 8 where we incorporated the set V_0 of exceptional vertices and chose the bridges which connect the triples (X'_k, Y'_k, Z'_k). Instead of choosing a loose path $\mathcal{L}$ which contains all the vertices in V_0 we now proceed as follows: First we make $|V_0|$ even by including an arbitrary additional vertex if necessary. Then we divide the vertices in V_0 arbitrarily into pairs v_i, w_i. Now we can argue as in the paragraph preceding (14) to find (for each i) two hyperedges $v_i w_i z$ and $v_i w_i z'$ of $\mathcal{H}$ such that the vertices z and z' lie in some Z^*_k and such that none of the sets Z^*_k is used too often in this step. For each i, these two hyperedges form a copy of $\mathcal{C}_4$. We delete all the vertices contained in these copies from the triples (X'_k, Y'_k, Z'_k).

Our aim now is to find a perfect $\mathcal{C}_4$-packing in the remainder of each triple. (Together with all the $\mathcal{C}_4$'s chosen before, this would form a perfect $\mathcal{C}_4$-packing in $\mathcal{H}$.) For this, a necessary condition is that the number of vertices in each triple is divisible by 4. So we remove at most 3 vertices from each triple to achieve this and let W denote the set of vertices which we removed in this step. So W contains a bounded number of vertices and $|W|$ is divisible by four. As with V_0, we now divide the vertices of W arbitrarily into pairs v_i, w_i. Since the minimum degree of $\mathcal{H}$ is at least $n/4$ and since at most $3n/100 + 2|V_0| + 3|\mathcal{R}|$ vertices of $\mathcal{H}$ do not lie in some triple (X^*_k, Y^*_k, Z^*_k), for all i we can find an index $k = k(i)$ so that the set $U(i) := X^*_k \cup Y^*_k \cup Z^*_k$ has the property that $v_i w_i$ forms a hyperedge with at least $|U(i)|/10$ of its vertices, say (here X^*_k, Y^*_k and Z^*_k are as defined in Section 8). Thus for all i there is a copy of $\mathcal{C}_4$ whose base pair is v_iw_i and which contains two further vertices $U(i)$. Clearly, we may assume that all these copies are disjoint from each other (though some of the $U(i)$ may well be identical). Remove these copies. (So we have now removed all vertices in W). For each pair $U(i)$, $U(i+1)$ where i is odd, we will now find two copies of $\mathcal{C}_4$ whose base pairs both contain one vertex in each of $U(i)$ and $U(i+1)$ and where the remaining 4 vertices all lie in the same triple. Since the total number of pairs v_i, w_i is even (and thus the total number of the $U(i)$ is even when counted with multiplicities), it follows that after removing the vertices in all these copies of $\mathcal{C}_4$, in each cluster the number of vertices will then be divisible by 4, as desired. To find these copies of $\mathcal{C}_4$, for each odd i we consider five pairs $a_{ij}b_{ij}$ of vertices where $a_{ij} \in U(i)$ and $b_{ij} \in U(i+1)$ for all j with $1 \leq j \leq 5$. Similarly as above, it follows that there must be indices j, j' and k such that a linear number of vertices in $X_k^* \cup Y_k^* \cup Z_k^*$ form a hyperedge of $\mathcal{H}$ with both $a_{ij}b_{ij}$ and $a_{ij'}b_{ij'}$. So for each odd i we can find 2 copies of $\mathcal{C}_4$ as described above (the base pairs are $a_{ij}b_{ij}$ and $a_{ij'}b_{ij'}$).

Instead of Lemma 9.1, which guaranteed the existence of an 'equalizing' path $\mathcal{Q}_k$ which contains almost all of the vertices in each of X_k^*, Y_k^* and Z_k^* we now need the following result:

Lemma 12.1 *For each triple (X_k, Y_k, Z_k) the induced hypergraph $\mathcal{H}_{P_k}$ contains a $\mathcal{C}_4$-packing $\mathcal{Q}_k$ with the following two properties:*

- *$\mathcal{Q}_k$ contains only vertices in $X_k^* \cup Y_k^* \cup Z_k^*$,*
- *the sets $X_k^{**} := X_k^* \setminus V(\mathcal{Q}_k)$, $Y_k^{**} := Y_k^* \setminus V(\mathcal{Q}_k)$ and $Z_k^{**} := Z_k^* \setminus V(\mathcal{Q}_k)$ satisfy*
$$|X_k^{**}| = |Y_k^{**}| = 2\sqrt{\xi}m \quad \text{and} \quad |Z_k^{**}| = 2|X_k^{**}|. \tag{34}$$

We will only argue that one can find $\mathcal{Q}_k$ in the complete 3-partite 3-uniform hypergraph spanned by X_k^*, Y_k^* and Z_k^*. The existence of $\mathcal{Q}_k$ in $\mathcal{H}_{P_k}$ then follows using easily using a greedy argument based on the regularity of P_k (similarly to Lemma 9.3).

To find $\mathcal{Q}_k$ in the complete hypergraph, we proceed in a step by step fashion as in the proof of Lemma 9.2. For simplicity, let $A := X_k^*$, $B := Y_k^*$ and $C := Z_k^*$. There are 3 phases (instead of 5 in Lemma 9.2). If $|A| > |B|$, then in the first phase, we take out copies of $\mathcal{C}_4$ whose base pairs consist of one vertex in B and one in C. In this way, we eventually achieve that the leftover sets $A_1 \subseteq A$, $B_1 \subseteq B$ and $C_1 \subseteq C$ satisfy $|A_1| = |B_1|$. If $|A| < |B|$, each base pair in the first phase will consist of one vertex in A and one in C. In the second phase, we successively take out pairs of copies of $\mathcal{C}_4$ where the base pair of the first copy has its vertices in A and C and in the second copy the base pair has its vertices in B and C (each such pair of copies of $\mathcal{C}_4$ is called a segment). Let $c := |C_1|$, $a := |A_1|$, $b := |B_1|$ and define s by $c = 2a - s$. Note that s is nonnegative since we know (as in the proof of Lemma 9.2) that the size of A and B is almost the same and that C is a little less than twice as large as A. We claim that s is divisible by four. Indeed, to prove the claim, note that $a + b + c$ is divisible by four by the argument preceding Lemma 12.1 above. If we also have that $a+b=2a$ is divisible by four, then clearly c and thus also s is divisible by four. If this is not the case, then $a+b=2a$ is still divisible by 2, i.e. $a+b=4i+2$ for some integer i. But since $a+b+c$ is divisible by four, this means that $c=4j+2$ for some integer j, which in turn implies that $s=2a-c=4i-4j$, as claimed. Now let a', b' c' be the sizes of the three sets after removing one segment as described above. Then $c'=c-2$, $a'=a-3$ and $b'=a'$. So $c'=c-2=2a-s-2=2(a'+3)-s-2=2a'-(s-4)$. Since s is divisible by four, this means that eventually the second phase must terminate with leftover sets $A_2 \subseteq A_1$, $B_2 \subseteq B_1$ and $C_2 \subseteq C_1$ whose sizes satisfy $2|A_2|=2|B_2|=|C_2|$. In the third phase we successively take out copies of $\mathcal{C}_4$ whose base pairs consist of one vertex in A and one in B until the sizes of the leftover sets thus obtained satisfy the assertion of the lemma.

Now let W_k denote the set of leftover vertices in each triple as in the beginning of Section 11. In what follows, we will usually write X'' instead of X''_k again (and similarly for the other sets). Instead of finding a Hamilton path R^* in the graph $G^{rich} = P^{rich}_{X'Y'}[X^{**} \cup X'' \cup Y^{**} \cup Y'']$ which satisfies the properties described in Lemma 11.9, we proceed as follows: first we obtain a perfect matching M_1 of G_0 as guaranteed by Lemma 11.3. Similarly as in the proof of Lemma 11.4 we modify M_1 into a perfect matching R^* of the graph G^{rich} which has the following properties:

Lemma 12.2 *G^{rich} contains a perfect matching R^* with the following properties:*

(i) *R^* avoids all the poor and all the unhappy edges.*

(ii) $|E(R^*) \setminus M_1| \leq 10^6\sqrt{\xi}\bar{n}$.

(iii) *All of the edges in $M_1 \setminus E(R^*)$ are poor or unhappy.*

(iv) *$|E(R^*) \cap F_e| \leq 350\varepsilon\bar{n}$ for every edge $e \in F_e$ which is not unhappy.*

Thus Lemma 12.2 is an analogue of Lemma 11.9. Finally, we consider a random partition of the vertices $Z'' \cup Z^{**}$ into two parts Z_1''' and Z_2''' of equal size (note that $|Z'' \cup Z^{**}|$ is even by (34) and the definition of Z'' in Section 8.1). Also, for $i=1,2$ we define a bipartite auxiliary graph H_i^* whose vertex sets are Z_i''' and $M := E(R^*)$ and where two vertices in H_i^* are joined by an edge if the corresponding three vertices in $\mathcal{H}$ form a hyperedge in $\mathcal{H}_P$. Note that $|M| = |Z_i'''|$. If we set $n_* := |M|$, $M' := M_1 \cap M$ and replace H^* by H_i^*, then our construction implies that Lemma 11.10 remains valid (and thus also Lemmas 11.11–11.13). Hence both H_1^* and H_2^* contain a perfect matching. Clearly, the union of these perfect matchings corresponds to a perfect $\mathcal{C}_4$-packing in the subhypergraph of $\mathcal{H}_{P_k}$ induced by the vertices in W_k.

12.2 Deriving Theorem 1.2 from Theorem 1.1

In this subsection, we use the large deviation bound in Section 4 (Lemma 4.2) to deduce Theorem 1.2 from Theorem 1.1. The idea is to randomly split the vertex set of the given hypergraph $\mathcal{H}$ into disjoint sets of k vertices each. Then for any i, the probability that the hypergraph $\mathcal{H}_i$ induced by the ith set of k vertices does not satisfy the requirements of Theorem 1.1 is exponentially small in k. However, in order to prove that with positive probability *all* of the $\mathcal{H}_i$ satisfy the requirements of Theorem 1.1 we have to proceed more carefully: we will obtain the $\mathcal{H}_i$ by a sequence of successive binary partitions.

Proposition 12.3 *There exists an integer n_1 such that the following holds for all $n \geq n_1$ and for all positive constants λ and γ with $1/3 \leq \lambda \leq 2/3$ and $0 < \gamma < 1$. Suppose we are given a hypergraph $\mathcal{G}$ with n vertices whose minimum degree is at least $n/4+\gamma n-n^{5/8}$ and where $\lambda n \in \mathbb{N}$. Then there is a partition of the vertex set of $\mathcal{G}$ into sets U and W such that $|U| = \lambda n =: u$, $|W| = (1-\lambda)n =: w$, and such that the induced subhypergraph $\mathcal{G}[U]$ has minimum degree at least $u/4 + \gamma u - u^{5/8}$ whereas $\mathcal{G}[W]$ has minimum degree at least $w/4 + \gamma w - w^{5/8}$.*

Proof. Consider a partition of the vertices of $\mathcal{G}$ into two sets U and W with $|U| = u$ chosen uniformly at random from the set of all such partitions. Let $\varepsilon := (1 - \lambda^{3/8})/u^{3/8}$. Consider a fixed pair x, y of vertices of $\mathcal{G}$ and define the random variable $X := |N_{\mathcal{G}}(x, y) \cap U|$. Note that $\varepsilon\mathbb{E}X \leq \varepsilon u = (1 - \lambda^{3/8})u^{5/8}$ and so

$$\begin{aligned}\mathbb{E}X &\geq \lambda(n/4 + \gamma n - n^{5/8}) = u/4 + \gamma u - \lambda^{3/8}u^{5/8}\\ &\geq u/4 + \gamma u - u^{5/8} + \varepsilon\mathbb{E}X.\end{aligned}$$

Then by Lemma 4.2 (and using the fact that $\mathbb{E}X \geq u/5$),

$$\begin{aligned}\mathbb{P}[X \leq u/4 + \gamma u - u^{5/8}] &\leq \mathbb{P}[X \leq (1 - \varepsilon)\mathbb{E}X]\\ &\leq 2\exp\left\{-\frac{1}{3}\frac{(1 - \lambda^{3/8})^2}{u^{3/4}}\frac{u}{5}\right\} \leq n^{-3}.\end{aligned}$$

The last inequality follows since we assumed n (and thus u) to be sufficiently large. Since the same result also holds with U replaced by W, it follows that the probability that there exists a pair of vertices x, y in $\mathcal{G}$ with $|N_{\mathcal{G}}(x, y)\cap U| < u/4 + \gamma u - u^{5/8}$ or $|N_{\mathcal{G}}(x, y) \cap W| < w/4 + \gamma w - w^{5/8}$ is strictly less than one. Hence the desired partition exists. □

Proof of Theorem 1.2. Let $\sigma = \gamma/2$. Set $k_0 = \max\{n_0(\sigma), n_1, (2/\gamma)^{8/3}\}$, where n_0 is the function defined in Theorem 1.1 and where n_1 is as defined as in Proposition 12.3. Suppose that $k \geq k_0$ and that k divides n. Consider a hypergraph $\mathcal{H}$ with n vertices and with minimum degree at least $n/4+\gamma n$. For simplicity, first assume that $n = 2^\ell k$, where ℓ is an integer. By applying Proposition 12.3 with $\lambda = 1/2$ first to $\mathcal{H}$ and then successively to the induced subhypergraphs obtained from it, we eventually obtain subhypergraphs $\mathcal{H}_1, \dots, \mathcal{H}_{2^\ell}$ where each of them has exactly k vertices and has minimum degree at least $k/4 + \gamma k - k^{5/8} \geq k/4 + \gamma k/2$. Thus by Theorem 1.1, each of the $\mathcal{H}_i$ contains a loose Hamilton cycle, as required.

If n cannot be written in the form $2^\ell k$, then we apply the same argument as above, but now we choose the sizes of the partition classes according to the binary expansion of n/k (e.g. if $n = 5k$, we first split the vertices into parts of size $2k$ and $3k$ respectively.) Note that we can always do this such that the ratio of the sizes of the parts lies between 1/3 and 2/3 (and thus Lemma 12.3 can be applied). □

References

[1] N. Alon and J. Spencer, *The Probabilistic Method* (2nd edition), Wiley-Interscience 2000.

[2] J.C. Bermond, A. Germa, M.C. Heydemann and D. Sotteau, Hypergraphes hamiltoniens, *Prob. Comb. Théorie Graph Orsay* **260** (1976) 39–43.

[3] B. Bollobás, *Extremal Graph Theory*, Academic Press, 1978.

[4] A. Czygrinow and B. Nagle, A note on codegreee problems for hypergraphs, *Bulletin of the ICA* **32** (2001), 63–69.

[5] Y. Dementieva, P.E. Haxell, B. Nagle and V. Rödl, On characterizing hypergraph regularity, *Random Structures & Algorithms* **21** (2002), 293–335.

[6] Y. Dementieva, P.E. Haxell, B. Nagle and V. Rödl, An algorithmic version of the hypergraph regularity lemma, in preparation.

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

[8] A. Frieze and M. Krivelevich, On packing Hamilton cycles in ε-regular graphs, preprint 2003.

[9] W.T. Gowers, Hypergraph regularity and the multidimensional Szemerédi Theorem, preprint 2004.

[10] S. Janson, T. Łuczak and A. Ruciński, *Random graphs*, Wiley-Interscience 2000.

[11] M. Jerrum, A. Sinclair and E. Vigoda, A polynomial time approximation algorithm for the permanent of a matrix with non-negative entries, *Proceedings of the 33rd ACM Symposium on Theory of Computing* (2001) 712–720.

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

[13] J. Komlós, G. N. Sárközy and E. Szemerédi, Blow-up lemma, *Combinatorica* **17** (1997), 109–123.

[14] J. Komlós, G. N. Sárközy and E. Szemerédi, Proof of the Alon-Yuster conjecture, *Discrete Math.* **235** (2001), 255–269.

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

[16] D. Kühn and D. Osthus, Matchings in hypergraphs of large minimum degree, submitted.

[17] D. Kühn and D. Osthus, Multicoloured Hamilton cycles and perfect matchings in pseudo-random graphs, submitted.

[18] L. Lovász and M.D. Plummer, *Matching theory*, North-Holland, 1986.

[19] B. Nagle and V. Rödl, Regularity properties for triple systems, *Random Structures & Algorithms* **23** (2003), 263–332.

[20] V. Rödl and A. Ruciński, Perfect matchings in ε-regular graphs and the Blow-up lemma, *Combinatorica* **19** (1999) 437–452.

[21] V. Rödl, A. Ruciński and E. Szemerédi, A Dirac-type theorem for 3-uniform hypergraphs, preprint 2004.

[22] V. Rödl and J. Skokan, Regularity lemma for k-uniform hypergraphs, *Random Structures & Algorithms* **25** (2004), 1–42.

Daniela Kühn & Deryk Osthus
School of Mathematics
Birmingham University
Edgbaston
Birmingham B15 2TT
UK
E-mail addresses: {kuehn,osthus}@maths.bham.ac.uk