UNIVERSITYᴼᶠ BIRMINGHAM

Research at Birmingham

A Multipartite Hajnal-Szemerédi Theorem

Keevash, Peter; Mycroft, Richard

Link to publication on Research at Birmingham portal

A MULTIPARTITE HAJNAL-SZEMERÉDI THEOREM

PETER KEEVASH AND RICHARD MYCROFT

ABSTRACT. The celebrated Hajnal-Szemerédi theorem gives the precise minimum degree threshold that forces a graph to contain a perfect K_k-packing. Fischer's conjecture states that the analogous result holds for all multipartite graphs except for those formed by a single construction. Recently, we deduced an approximate version of this conjecture from new results on perfect matchings in hypergraphs. In this paper, we apply a stability analysis to the extremal cases of this argument, thus showing that the exact conjecture holds for any sufficiently large graph.

1. INTRODUCTION

A fundamental result of Extremal Graph Theory is the Hajnal-Szemerédi theorem, which states that if k divides n then any graph G on n vertices with minimum degree $\delta(G) \geq (k-1)n/k$ contains a perfect K_k-packing, i.e. a spanning collection of vertex-disjoint k-cliques. This paper considers a conjecture of Fischer [4] on a multipartite analogue of this theorem. Suppose $V_1, \ldots, V_k$ are disjoint sets of n vertices each, and G is a k-partite graph on vertex classes $V_1, \ldots, V_k$ (that is, G is a graph on the vertex set $V_1 \cup \cdots \cup V_k$ such that no edge of G has both endvertices in the same V_j). We define the *partite minimum degree* of G, denoted $\delta^*(G)$, to be the largest m such that every vertex has at least m neighbours in each part other than its own, i.e.

$$\delta^*(G) := \min_{i \in [k]} \min_{v \in V_i} \min_{j \in [k] \setminus \{i\}} |N(v) \cap V_j|,$$

where $N(v)$ denotes the neighbourhood of v. Fischer conjectured that if $\delta^*(G) \geq (k-1)n/k$ then G has a perfect K_k-packing. This conjecture is straightforward for $k = 2$, as it is not hard to see that any maximal matching must be perfect in this case, but for odd $k \geq 3$ the conjecture does not hold, as can be seen from constructions provided by Catlin [1] (these counterexamples are presented in Construction 1.2 as $\Gamma_{n,k,k}$). For $k = 3$, Magyar and Martin [9] proved that, for large n, Catlin's construction is in fact the only counterexample to this conjecture. More precisely, they showed that if n is sufficiently large, G is a 3-partite graph with vertex classes each of size n and $\delta^*(G) \geq 2n/3$, then either G contains a perfect K_3-packing, or n is odd and divisible by 3, and G is isomorphic to the graph $\Gamma_{n,3,3}$ defined in Construction 1.2.

The implicit conjecture behind this result (stated explicitly by Kühn and Osthus [7]) is that the only counterexamples to Fischer's original conjecture are those constructed by Catlin, that is, the graphs $\Gamma_{n,k,k}$ defined in Construction 1.2 when n is odd and divisible by k. We refer to this as the modified Fischer conjecture. If k is even then n cannot be both odd and divisible by k, so the modified Fischer conjecture is the same as the original conjecture in this case. Martin and Szemerédi [10] proved that (the modified) Fischer's conjecture holds for $k = 4$. Another partial result was obtained by Csaba and Mydlarz [2], who gave a function $f(k)$ with $f(k) \to 0$ as $k \to \infty$ such that the conjecture holds for large n if one strengthens the degree assumption to $\delta^*(G) \geq (k-1)n/k + f(k)n$. Recently, an approximate version of the conjecture assuming the degree condition $\delta^*(G) \geq (k-1)n/k + o(n)$ was proved

independently and simultaneously by Keevash and Mycroft [6], and by Lo and Markström [8]. The proof in [6] was a quick application of the geometric theory of hypergraph matchings developed in the same paper; this will be formally introduced in the next section. By a careful analysis of the extremal cases of this result, we will obtain the following theorem, the case $r = k$ of which shows that (the modified) Fischer's conjecture holds for any sufficiently large graph. Note that the graph $\Gamma_{n,r,k}$ in the statement is defined in Construction 1.2.

Theorem 1.1. *For any $r \geq k$ there exists n_0 such that for any $n \geq n_0$ with $k \mid rn$ the following statement holds. Let G be an r-partite graph whose vertex classes each have size n such that $\delta^*(G) \geq (k-1)n/k$. Then G contains a perfect K_k-packing, unless rn/k is odd, $k \mid n$, and $G \cong \Gamma_{n,r,k}$.*

We now give the generalised version of the construction of Catlin [1] which shows Fischer's original conjecture to be false.

Construction 1.2. *Suppose rn/k is odd and k divides n. Let V be a vertex set partitioned into parts $V_1, \ldots, V_r$ of size n. Partition each V_i, $i \in [r]$ into subparts V_i^j, $j \in [k]$ of size n/k. Define a graph $\Gamma_{n,r,k}$, where for each $i, i' \in [r]$ with $i \neq i'$ and $j \in [k]$, if $j \geq 3$ then any vertex in V_i^j is adjacent to all vertices in $V_{i'}^{j'}$ with $j' \in [k] \setminus \{j\}$, and if $j = 1$ or $j = 2$ then any vertex in V_i^j is adjacent to all vertices in $V_{i'}^{j'}$ with $j' \in [k] \setminus \{3 - j\}$.*
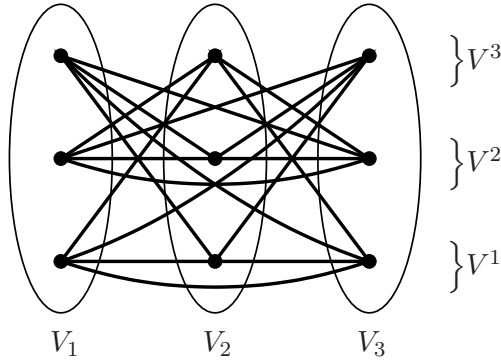
FIGURE 1. Construction 1.2 for the case $k = r = 3$.

Figure 1 shows Construction 1.2 for the case $k = r = 3$. For $n = k$ this is the exact graph of the construction; for larger n we 'blow up' the graph above, replacing each vertex by a set of size n/k, and each edge by a complete bipartite graph between the corresponding sets. In general, it is helpful to picture the construction as an r by k grid, with columns corresponding to parts V_i, $i \in [r]$ and rows $V^j = \bigcup_{i \in [r]} V_i^j$, $j \in [k]$ corresponding to subparts of the same superscript. Vertices have neighbours in other rows and columns to their own, except in rows V^1 and V^2, where vertices have neighbours in other columns in their own row and other rows besides rows V^1 and V^2. Thus $\delta^*(G) = (k-1)n/k$. We claim that there is no perfect K_k-packing. For any K_k has at most one vertex in any V^j with $j \geq 3$, so at most $k - 2$ vertices in $\bigcup_{j \geq 3} V^j$. Also $\left| \bigcup_{j \geq 3} V^j \right| = (k-2)rn/k$, and there are rn/k copies of K_k in a perfect packing. Thus each K_k must have $k - 2$ vertices in $\bigcup_{j \geq 3} V^j$, and so 2 vertices in $V^1 \cup V^2$, which must either both lie in V^1 or both lie in V^2. However, $|V^1| = rn/k$ is odd, so V^1 cannot be perfectly covered by pairs. Thus G contains no perfect K_k-packing.

This paper is organised as follows. In the next section we introduce ideas and results from [6] on perfect matchings in k-graphs. Section 3 gives an outline of the proof of Theorem 1.1. In Sections 4 to 7 we prove several preliminary lemmas, before combining these lemmas in Section 8 to prove Theorem 1.1.

Notation. The following notation is used throughout the paper: $[k] = \{1, \ldots, k\}$; if X is a set then $\binom{X}{k}$ is the set of subsets of X of size k; $x \ll y$ means that for every $y > 0$ there exists some $x_0 > 0$ such that the subsequent statement holds for any $x < x_0$ (such statements with more variables are defined similarly); if x is a vertex in a graph then $N(x)$ is the neighbourhood of x.

2. PERFECT MATCHINGS IN HYPERGRAPHS

In this section we describe the parts of the geometric theory of perfect matchings in hypergraphs from [6] that we will use in the proof of Theorem 1.1. We start with some definitions. A *hypergraph* G consists of a vertex set V and an edge set E, where each edge $e \in E$ is a subset of V. We say that G is a *k-graph* if every edge has size k. A *matching* M in G is a set of vertex-disjoint edges in G. We call M *perfect* if it covers all of V. We identify a hypergraph H with its edge set, writing $e \in H$ for $e \in E(H)$, and $|H|$ for $|E(H)|$. A *k-system* is a hypergraph J in which every edge of J has size at most k and $\emptyset \in J$. We refer to the edges of size r in J as *r-edges of J*, and write J_r to denote the r-graph on $V(J)$ formed by these edges. A *k-complex* J is a k-system whose edge set is closed under inclusion, i.e. if $e \in J$ and $e' \subseteq e$ then $e' \in J$. For any non-empty k-graph G, we may generate a k-complex $G^{\leq}$ whose edges are any $e \subseteq V(G)$ such that $e \subseteq e'$ for some edge $e' \in G$.

Let V be a set of vertices, and let $\mathcal{P}$ partition V into parts $V_1, \ldots, V_r$ of size n. Then we say that a hypergraph G with vertex set V is *$\mathcal{P}$-partite* if $|e \cap V_i| \leq 1$ for every $i \in [r]$ and $e \in G$. We say that G is *r-partite* if it is $\mathcal{P}$-partite for some partition $\mathcal{P}$ of V into r parts.

Let J be a $\mathcal{P}$-partite k-system on V. For each $0 \leq j \leq k - 1$ we define the *partite minimum j-degree* $\delta_j^*(J)$ as the largest m such that any j-edge e has at least m extensions to a $(j + 1)$-edge in any part not intersected by e, i.e.

$$\delta_j^*(J) := \min_{e \in J_j} \min_{i: e \cap V_i = \emptyset} |\{v \in V_i : e \cup \{v\} \in J\}|.$$

The *partite degree sequence* is $\delta^*(J) = (\delta_0^*(J), \ldots, \delta_{k-1}^*(J))$. Note that we suppress the dependence on $\mathcal{P}$ in our notation: this will be clear from the context. Note also that this is *not* the standard notion of degree used in k-graphs, in which the degree of a set is the number of edges containing it. Our minimum degree assumptions will always be of the form $\delta(J) \geq \mathbf{a}$ pointwise for some vector $\mathbf{a} = (a_0, \ldots, a_{k-1})$, i.e. $\delta_i(J) \geq a_i$ for $0 \leq i \leq k - 1$. It is helpful to interpret this 'dynamically' as follows: when constructing an edge of J_k by greedily choosing one vertex at a time, there are at least a_i choices for the $(i + 1)$st vertex (this is the reason for the requirement that $\emptyset \in J$, which we need for the first choice in the process).

The following key definition relates our theorems on hypergraphs to graphs. Fix $r \geq k$ and a partition $\mathcal{P}$ of a vertex set V into r parts $V_1, \ldots, V_r$ of size n. Let G be a $\mathcal{P}$-partite graph on V. Then the *clique k-complex* $J(G)$ of G is the k-complex whose edges of size i are precisely the copies of K_i in G for $0 \leq i \leq k$. Note that $J(G)$ must be $\mathcal{P}$-partite. Furthermore, if $\delta^*(G) \geq (k-1)n/k - \alpha n$ and $0 \leq i \leq k - 1$, then the vertices of any copy of K_i in G have at least $n - in/k - i\alpha n$ common neighbours in each V_j which they do not

intersect. That is, if G satisfies $\delta^*(G) \geq (k-1)n/k - \alpha n$, then the clique k-complex $J(G)$ satisfies

$$(1) \qquad \delta^*(J(G)) \geq \left(n, \left(\frac{k-1}{k} - \alpha\right)n, \left(\frac{k-2}{k} - 2\alpha\right)n, \ldots, \left(\frac{1}{k} - (k-1)\alpha\right)n\right).$$

Note also that any perfect matching in the k-graph $J(G)_k$ corresponds to a perfect K_k-packing in G. So if we could prove that any $\mathcal{P}$-partite k-complex J on V which satisfies (1) must have a perfect matching in the k-graph J_k, then we would have already proved Theorem 1.1! Along these lines, Theorem 2.4 in [6] shows that any such J must have a matching in J_k which covers all but a small proportion of the vertices of J. (Here we assume $1/n \ll \alpha \ll 1/r, 1/k$.) However, two different families of constructions show that this condition does not guarantee a perfect matching in J_k; we refer to these as *space barriers* and *divisibility barriers*. We will describe these families in some detail, since the results of [6] show that these are essentially the only k-complexes J on V which satisfy (1) but do not have a perfect matching in J_k. Firstly, space barriers are characterised by a bound on the size of the intersection of every edge with some fixed set $S \subseteq V(J)$. If S is too large, then J_k cannot contain a perfect matching. The following construction gives the precise formulation.

Construction 2.1. *(Space barriers) Suppose $\mathcal{P}$ partitions a set V into r parts $V_1, \ldots, V_r$ of size n. Fix $j \in [k-1]$ and a set $S \subseteq V$ containing $s = \lfloor(j/k + \alpha)n\rfloor$ vertices in each part V_j. Then we denote by $J = J_r(S, j)$ the k-complex in which J_i (for $0 \leq i \leq k$) consists of all $\mathcal{P}$-partite sets $e \subseteq V$ of size i that contain at most j vertices of S. Observe that $\delta_i^*(J) = n$ for $0 \leq i \leq j-1$ and $\delta_i^*(J) = n - s$ for $j \leq i \leq k-1$, so (1) is satisfied. However, any matching in J_k has size at most $\left\lfloor\frac{|V \setminus S|}{k-j}\right\rfloor$ and so leaves at least $r(\alpha n - k)$ vertices uncovered.*

Having described the general form of space barriers, we now turn our attention to divisibility barriers. These are characterised by every edge satisfying an arithmetic condition with respect to some partition $\mathcal{Q}$ of V. To be more precise, we need the following definition. Fix any partition $\mathcal{Q}$ of a vertex set V into d parts $V_1, \ldots, V_d$. For any $\mathcal{Q}$-partite set $S \subseteq V$ (that is, S has at most one vertex in each part of $\mathcal{Q}$), the *index set of S with respect to $\mathcal{Q}$* is $i_\mathcal{Q}(S) := \{i \in [d] : |S \cap V_i| = 1\}$. For general sets $S \subseteq V$, we have the similar notion of the *index vector of S with respect to $\mathcal{Q}$*; this is the vector $\mathbf{i}_\mathcal{Q}(S) := (|S \cap V_1|, \ldots, |S \cap V_d|)$ in $\mathbb{Z}^d$. So $\mathbf{i}_\mathcal{Q}(S)$ records how many vertices of S are in each part of $\mathcal{Q}$. Observe that if S is $\mathcal{Q}$-partite then $\mathbf{i}(S)$ is the characteristic vector of the index set $i(S)$. When $\mathcal{Q}$ is clear from the context, we write simply $i(S)$ and $\mathbf{i}(S)$ for $i_\mathcal{Q}(S)$ and $\mathbf{i}_\mathcal{Q}(S)$ respectively, and refer to $i(S)$ simply as the *index* of S. We will consider the partition $\mathcal{Q}$ to define the order of its parts so that $\mathbf{i}_\mathcal{Q}(S)$ is well-defined.

Construction 2.2. *(Divisibility barriers) Suppose $\mathcal{Q}$ partitions a set V into d parts, and L is a lattice in $\mathbb{Z}^d$ (i.e. an additive subgroup) with $\mathbf{i}(V) \notin L$. Fix any $k \geq 2$, and let G be the k-graph on V whose edges are all k-tuples e with $\mathbf{i}(e) \in L$. For any matching M in G with vertex set $S = \bigcup_{e \in M} e$ we have $\mathbf{i}(S) = \sum_{e \in M} \mathbf{i}(e) \in L$. Since we assumed that $\mathbf{i}(V) \notin L$ it follows that G does not have a perfect matching.*

For the simplest example of a divisibility barrier take $d = 2$ and $L = \langle(-2, 2), (0, 1)\rangle$. So $(x, y) \in L$ precisely when x is even. Then the construction described has $|V_1|$ odd, and the edges of G are all k-tuples $e \subseteq V$ such that $|e \cap V_1|$ is even. If $|V| = n$ and $|V_1| \sim n/2$, then any set of $k-1$ vertices of G is contained in around $n/2$ edges of G, but G contains no perfect matching.

We now consider the multipartite setting. Let $\mathcal{P}$ partition a vertex set V into parts $V_1, \ldots, V_r$ of size n, and let $\mathcal{Q}$ be a partition of V into d parts $U_1, \ldots, U_d$ which refines $\mathcal{P}$. Then we say that a lattice $L \subseteq \mathbb{Z}^d$ is *complete with respect to* $\mathcal{P}$ if L contains every difference of basis vectors $\mathbf{u}_i - \mathbf{u}_j$ for which U_i and U_j are contained in the same part V_ℓ of $\mathcal{P}$, otherwise we say that L is *incomplete with respect to* $\mathcal{P}$. The idea behind this definition is that if L is incomplete with respect to $\mathcal{P}$, then it is possible that $\mathbf{i}_\mathcal{Q}(V) \notin L$, in which case we would have a divisibility barrier to a perfect matching, whilst if L is complete with respect to $\mathcal{P}$ then this is not possible.

There is a natural notion of minimality for an incomplete lattice L with respect to $\mathcal{P}$. We say that $\mathcal{Q}$ is *transferral-free* if L does not contain any difference of basis vectors $\mathbf{u}_i - \mathbf{u}_j$ for which U_i, U_j are contained in the same part V_ℓ of $\mathcal{P}$. For suppose L does contain some such difference $\mathbf{u}_i - \mathbf{u}_j$ and form a partition $\mathcal{Q}'$ from $\mathcal{Q}$ by merging parts U_i and U_j of $\mathcal{Q}$. Let $L' \subseteq \mathbb{Z}^{d-1}$ be the lattice formed by this merging (that is, by replacing the ith and jth coordinates with a single coordinate equal to their sum). Then L' is also incomplete with respect to $\mathcal{P}$, so we have a smaller divisibility barrier.

Let J be an r-partite k-complex whose vertex classes $V_1, \ldots, V_r$ each have size n. The next theorem, Theorem 2.9 from [6], states that if J satisfies (1) and J_k is not 'close' to either a space barrier or a divisibility barrier, then J_k contains a perfect matching. Moreover, we can find a perfect matching in J_k which has roughly the same number of edges of each index. More precisely, for a perfect matching M in J_k and a set $A \in \binom{[r]}{k}$ let $N_A(M)$ be the number of edges $e \in M$ with index $i(e) = A$. We say that M is *balanced* if $N_A(M)$ is constant over all $A \in \binom{[r]}{k}$, that is, if there are equally many edges of each index. Similarly, we say that M is *γ-balanced* if $N_A(M) = (1 \pm \gamma)N_B(M)$ for any $A, B \in \binom{[r]}{k}$. Finally, we formalise the notion of closeness to a space or divisibility barrier as follows. Let G and H be k-graphs on a common vertex set V of size n. We say G is *β-contained* in H if all but at most βn^k edges of G are edges of H. Also, given a partition $\mathcal{P}$ of V into d parts, we define the *μ-robust edge lattice* $L_\mathcal{P}^\mu(G) \subseteq \mathbb{Z}^d$ to be the lattice generated by all vectors $\mathbf{v} \in \mathbb{Z}^d$ such that there are at least μn^k edges $e \in G$ with $\mathbf{i}_\mathcal{P}(e) = \mathbf{v}$.

Theorem 2.3. *Suppose that $1/n \ll \gamma \ll \alpha \ll \mu, \beta \ll 1/r$, $r \geq k$ and $k \mid rn$. Let $\mathcal{P}'$ partition a set V into parts $V_1, \ldots, V_r$ each of size n. Suppose that J is a $\mathcal{P}'$-partite k-complex with*

$$\delta^*(J) \geq \left(n, \left(\frac{k-1}{k} - \alpha \right) n, \left(\frac{k-2}{k} - \alpha \right) n, \ldots, \left(\frac{1}{k} - \alpha \right) n \right).$$

Then J has at least one of the following properties.

  1 **(Matching):** *J_k contains a γ-balanced perfect matching.*
  2 **(Space barrier):** *J_k is β-contained in $J_r(S, j)_k$ for some $j \in [k-1]$ and $S \subseteq V$ with $\lfloor jn/k \rfloor$ vertices in each V_i, $i \in [r]$.*
  3 **(Divisibility barrier):** *There is some partition $\mathcal{P}$ of $V(J)$ into $d \leq kr$ parts of size at least $\delta_{k-1}^*(J) - \mu n$ such that $\mathcal{P}$ refines $\mathcal{P}'$ and $L_\mathcal{P}^\mu(J_k)$ is incomplete with respect to $\mathcal{P}'$.*

Note that the fact that the perfect matching in J_k is γ-balanced in the first property is not stated in the statement of the theorem in [6]. However, examining the short derivation of this theorem from Theorem 7.11 in [6] shows this to be the case.
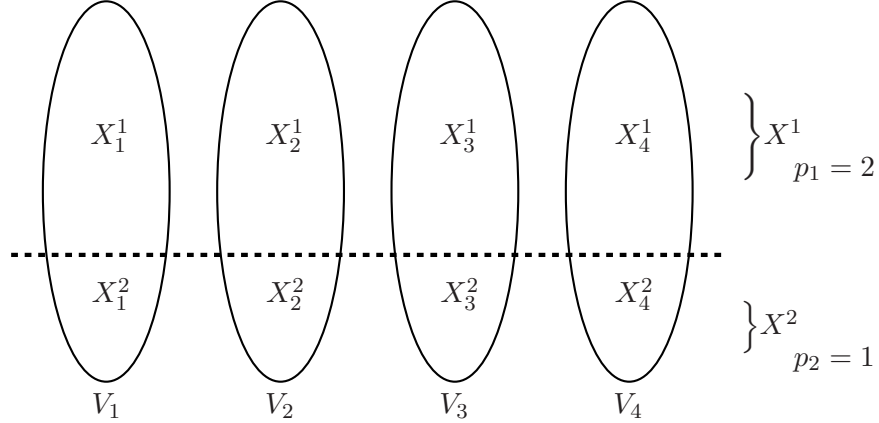
FIGURE 2. A row-decomposition of a 4-partite graph G into 2 rows.

3. OUTLINE OF THE PROOF

In this section we outline the proof of Theorem 1.1. For ease of explanation we restrict to the case when G is an r-partite graph whose vertex classes each have size kn and $\delta^*(G) \geq (k-1)n$. Our strategy consists of the following three steps:

(i) Impose a row structure on G.
(ii) Find balanced perfect clique-packings in each row.
(iii) Glue together the row clique-packings to form a K_k-packing of G.

For step (i) we partition $V(G)$ into *blocks* X_j^i, so that each vertex class V_j is partitioned into s blocks $X_j^1, \ldots, X_j^s$. This partition is best thought of as an $s \times r$ grid, with rows $X^i := \bigcup_{j \in [r]} X_j^i$ and columns the vertex classes $V_j = \bigcup_{i \in [s]} X_j^i$. We insist that all the blocks in a given row X^i have equal size $p_i n$, where $\sum_{i \in [s]} p_i = k$. We call a partition of $V(G)$ which satisfies these conditions an *s-row-decomposition* of G. We also require that G has density close to 1 between any two blocks which do not lie in the same row or column (we refer to the smallest such density as the *minimum diagonal density*). Figure 2 illustrates this structure. We begin with the trivial 1-row-decomposition of G with a single row (so the blocks are the vertex classes V_j). If it is possible to split this row into two rows to obtain a row-decomposition with minimum diagonal density at least $1 - d$ (where d will be small), then we say that G is *d-splittable*. If so, we partition G in this manner, and then examine in turn whether either of the two rows obtained is splittable (for some larger value of d). By repeating this process, we obtain a row-decomposition of G with high minimum diagonal density in which no row is splittable; this argument is formalised in Lemma 4.1.

For step (ii) we require a balanced perfect K_{p_i}-packing in each row X^i. We first use the results of Section 2 to obtain a near-balanced perfect K_{p_i}-packing in $G[X^i]$. Fix i and take J to be the clique p_i-complex of $G[X^i]$. So we regard the row X^i as an r-partite vertex set whose parts are the blocks $X_1^i, \ldots, X_r^i$, and the edges of J_j are the j-cliques in $G[X^i]$ for $j \leq p_i$. The assumption $\delta^*(G) \geq (k-1)n$ implies that

$$\delta^*(J) \geq (p_i n, (p_i - 1)n, (p_i - 2)n, \ldots, n).$$

Then Theorem 2.3 (with p_i playing the role of k) implies that J_{p_i} contains a near-balanced perfect matching, unless J_{p_i} is close to a space or divisibility barrier. In Section 4 we consider

a space barrier, showing in Lemma 4.2 that since $G[X^i]$ is not d-splittable, J_{p_i} cannot be close to a space barrier. We then consider a divisibility barrier in Section 5. For $p_i \geq 3$, Lemma 5.3 shows that since $G[X^i]$ is not d-splittable, J_{p_i} also cannot be close to a divisibility barrier. However, the analogous statement for $p_i = 2$ is false, for the following reason.

We say that $G[X^i]$ is 'pair-complete' if it has a structure close to that which appears in rows V^1 and V^2 of Construction 1.2. That is, there is a partition of X^i into 'halves' S and $X^i \setminus S$, such that each vertex class V_j is partitioned into two equal parts, and both $G[S]$ and $G[X^i \setminus S]$ are almost complete r-partite graphs. Such a row is not d-splittable if r is odd, but J_2 is close to a divisibility barrier. However, Lemma 5.2 shows that this is essentially the only such example, that is, that if $p_i = 2$ and $G[X^i]$ is neither d-splittable nor pair-complete then J_2 is not close to a divisibility barrier. So unless $p_i = 2$ and $G[X^i]$ is pair-complete, Theorem 2.3 implies that J_{p_i} contains a near-balanced perfect K_{p_i}-packing. In Section 6 we then show that we can actually obtain a balanced perfect matching in J_{p_i}. Indeed, in Lemma 6.2 we first delete some 'configurations' from $G[X^i]$; these are subgraphs of $G[X^i]$ that can be expressed as two disjoint copies of K_{p_i} in $G[X^i]$ in two different ways (with different index sets). After these deletions we proceed as just described to find a near-balanced perfect K_{p_i}-packing in $G[X^i]$. Then by carefully choosing which pair of disjoint edges to add to the matching from each 'configuration', we obtain a balanced perfect K_{p_i}-packing in $G[X^i]$, as required. This leaves only the case where $p_i = 2$ and $G[X^i]$ is pair-complete; in this case Lemma 6.4 gives a balanced perfect K_2-packing in $G[X^i]$, provided that each half has even size.

For step (iii), we construct auxiliary hypergraphs, perfect matchings in which describe how to glue together the perfect K_{p_i}-packings in the rows into a perfect K_k-packing of G. Recall that the row-decomposition of G was chosen to have large minimum diagonal density, so almost every vertex of any block X_j^i has few non-neighbours in any block $X_{j'}^{i'}$ in a different row and column. Assume for now that this row-decomposition of G has the stronger condition of large *minimum diagonal degree*, i.e. that we can delete 'almost' from the previous statement. For each row i, we partition its perfect K_{p_i}-packing into sets $E_{\sigma,i}$, one for each injective function $\sigma : [k] \to [r]$. For each σ we then form an auxiliary s-partite s-graph H_σ, where for each $i \in [s]$ the i-th vertex class of H_σ is the set $E_{\sigma,i}$ (so a copy of K_{p_i} in $G[X^i]$ is a vertex of H_σ). Edges in H_σ are those s-tuples of vertices for which the corresponding copies of K_{p_i} together form a copy of K_k in G. We defer the details of the partition to the final section of this paper; the crucial point is that the large minimum diagonal degree of G ensures that each H_σ has sufficiently large vertex degree to guarantee a perfect matching. Taking the copies of K_k in G corresponding to the union of these matchings gives a perfect K_k-packing in G, completing the proof.

The above sketch glosses over the use of the precise minimum degree condition in Theorem 1.1. Indeed, to replace our minimum diagonal density condition with a minimum diagonal degree condition, we must remove all 'bad' vertices, namely those which have many non-neighbours in some block in a different row and column. To achieve this, before step (ii) we delete some vertex-disjoint copies of K_k from G which cover all bad vertices. We must ensure that the number of vertices deleted from row X^i is a constant multiple of p_i for each i, so that we will be able to join together the K_{p_i}-packings of the undeleted vertices of each X^i to form a K_k-packing of G. We also need to ensure that each half has even size in pair-complete rows. This is accomplished in Section 7, which is the most lengthy and

technical part of the paper. After this, it is fairly quick to complete the proof as outlined above in Section 8.

4. ROW DECOMPOSITIONS AND SPACE BARRIERS

In this section we formalise our description of row-decompositions and the iterative process of splitting rows described in Section 3. We then show that the clique p-complex of any row obtained at the end of this process is not close to a space barrier. Note that many definitions and results of this section (and later sections) require that the size of each vertex class should be a multiple of k, which is not assumed in the statement of Theorem 1.1. However, our first step in the proof of Theorem 1.1 will be to remove vertices so that this condition is satisfied, allowing these definitions and results to be used.

4.1. **Row-decompositions.** Fix $r \geq 2$, and let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$ each of size kn. Suppose $s \in [k]$ and p_i, $i \in [s]$ are positive integers with $\sum_{i \in [s]} p_i = k$. Write $p = (p_i : i \in [s])$. An *s-row-decomposition* $X = (X_j^i)_{i \in [s], j \in [r]}$ of G (of *type* p), consists of subsets $X_j^i \subseteq V_j$ with $|X_j^i| = p_i n$ for each $i \in [s]$ and $j \in [r]$ such that each V_j is partitioned by the sets X_j^i with $i \in [s]$. We refer to the sets X_j^i as the *blocks*, and the sets $X^i := X_1^i \cup \cdots \cup X_r^i$ for $i \in [s]$ as the *rows*. We call the parts $X_j := V_j = X_j^1 \cup \cdots \cup X_j^s$ for $j \in [r]$ the *columns*, so G has s rows and r columns. Given subsets A, B of different vertex classes of G, let $G[A, B]$ denote the bipartite subgraph of G induced by $A \cup B$. We write $e_G(A, B) = |G[A, B]|$ and define the *density* of G between A and B as $d_G(A, B) = \frac{e_G(A,B)}{|A||B|}$. We usually write $e(A, B) = e_G(A, B)$ and $d(A, B) = d_G(A, B)$, as G is clear from the context. The *minimum diagonal density* of G is defined to be the minimum of $d(X_j^i, X_{j'}^{i'})$ over all $i \neq i'$ and $j \neq j'$. If G has only one row then for convenience we define the minimum diagonal density of G to be 1. Note that all this terminology depends on the choice of row-decomposition of G, but this will be clear from the context.

For any $i \in [s]$ with $p_i \geq 2$ we may obtain an $(s+1)$-row-decomposition of G by partitioning the row X^i of G. Indeed, choose positive integers y and z with $y + z = p_i$. For each $j \in [r]$ partition X_j^i into sets Y_j^i and Z_j^i with $|Y_j^i| = yn$ and $|Z_j^i| = zn$. Take $p_i' := y$, $p_{s+1}' := z$ and $p_\ell' := p_\ell$ for each $\ell \in [s] \setminus \{i\}$, and for each $j \in [r]$ let $\hat{X}_j^i := Y_j^i$, $\hat{X}_j^{s+1} := Z_j^i$ and $\hat{X}_j^\ell := X_j^\ell$ for each $\ell \in [s] \setminus \{i\}$. Then the blocks $\hat{X}_j^\ell$ form an $(s + 1)$-row-decomposition of G of type $p' = (p_\ell' : \ell \in [s + 1])$.

Bearing in mind the proof strategy sketched above, we are happy to split rows provided that we keep the minimum diagonal density close to 1. Thus we make the following definition. Let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$ each of size pn. We say G is *d-splittable* if for some $p' \in [p - 1]$ we may choose sets $S_i \subseteq V_i$, $i \in [r]$ with $|S_i| = p'n$, such that for any $i, i' \in [r]$ with $i \neq i'$ we have $d(S_i, V_{i'} \setminus S_{i'}) \geq 1 - d$. It is helpful to think of G as being a row-decomposition with just one row; then G is d-splittable if it is possible to partition this row into two rows as described above so that the minimum diagonal density is at least $1 - d$. Note that this definition depends on p, however this will always be clear from the context. Note also that G can never be d-splittable if $p = 1$. The next proposition shows that we can iteratively split G until we reach a row-decomposition which has high minimum diagonal density and does not have any splittable row.

Proposition 4.1. *Suppose that $1/n \ll d_0 \ll \cdots \ll d_k \ll 1/r$ and $r \geq 2$. Let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$ each of size kn. Then for some $s \in [k]$ there exists an s-row-decomposition X of G with minimum diagonal density at least $1 - k^2 d_{s-1}$ such that each row $G[X^i]$ of G is not d_s-splittable.*

Proof. Initially we take the trivial 1-row-decomposition of G with one row whose blocks are the vertex classes $V_1, \ldots, V_r$ of G. We now repeat the following step. Given an s-row-decomposition of G, if every row $G[X^i]$ is not d_s-splittable, then terminate. Alternatively, if $G[X^i]$ is d_s-splittable for some $i \in [s]$, according to some sets $S_j \subseteq X_j^i$, $j \in [r]$, then partition each block X_j^i into two blocks S_j and $X_j^i \setminus S_j$ to obtain an $(s+1)$-row-decomposition of G.

Since $G[X^i]$ can only be d_s-splittable if $p_i \geq 2$, this process must terminate with $s \leq k$. Then we have an s-row-decomposition of G all of whose rows are not d_s-splittable, so it remains only to show that G has minimum diagonal density at least $1 - k^2 d_{s-1}$. If $s = 1$ then this is true by definition, so we may assume $s \geq 2$. Consider any rows $i \neq i'$ and columns $j \neq j'$. Since X_j^i and $X_{j'}^{i'}$ do not lie in the same row of G, at some point in the process we must have partitioned blocks Y_j^ℓ and $Y_{j'}^\ell$ into $S_j, Y_j^\ell \setminus S_j$ and $S_{j'}, Y_{j'}^\ell \setminus S_{j'}$ with $X_j^i \subseteq S_j$ and $X_{j'}^{i'} \subseteq Y_{j'}^\ell \setminus S_{j'}$ respectively. Since $G[Y^\ell]$ was d_t-splittable for some $t \leq s - 1$, we have $d(S_j, Y_{j'}^\ell \setminus S_{j'}) \geq 1 - d_{s-1}$. Then, since $|X_j^i| \geq |S_j|/k$ and $|X_{j'}^{i'}| \geq |Y_{j'}^\ell|/k$, we have $d(X_j^i, X_{j'}^{i'}) \geq 1 - k^2 d_{s-1}$, as required. $\qquad\square$

4.2. **Avoiding space barriers.** Let G be an r-partite graph whose vertex classes have size pn with $\delta^*(G) \geq (p-1)n - \alpha n$, and let $J = J(G)$ be the clique p-complex of G. In this section we show that if G is not d-splittable then there is no space barrier to a perfect matching in J_p. We shall use this result in combination with the results of the next section to find a perfect clique packing in each row. We also prove that if $p < r$ then G contains many copies of K_{p+1}; this result will play an important role in the proof of Lemma 7.2.

Lemma 4.2. *Suppose that $1/n \ll \alpha \ll \beta \ll d \ll 1/r$ and $2 \leq p \leq r$. Let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$ each of size pn with $\delta^*(G) \geq (p-1)n - \alpha n$. Suppose that G is not d-splittable. Then*

(i) *for any $p' \in [p-1]$ and sets $S_i \subseteq V_i$, $i \in [r]$ of size $p'n$ there are at least βn^p copies of K_p in G with more than p' vertices in $G[S]$, where $S = \bigcup_{i \in [r]} S_i$, and*

(ii) *if $p < r$ then there are at least βn^{p+1} copies of K_{p+1} in G.*

Proof. For (i), since G is not d-splittable, we may suppose that $d(S_1, V_{p'+1} \setminus S_{p'+1}) < 1 - d$. Let A be the set of vertices in S_1 with fewer than $(1 - d/2)(p - p')n$ neighbours in $V_{p'+1} \setminus S_{p'+1}$. Write $|A| = ap'n$. Then $(1 - d/2)(1 - a) < d(S_1, V_{p'+1} \setminus S_{p'+1}) < 1 - d$, so $a > d/2$. We now greedily form a copy of $K_{p'+1}$ in $G[S]$ by choosing a vertex $v_i \in S_i$ for each $i \in [p'+1]$ in turn (in increasing order). We do this so that $v_1 \in A$ and $v_i \in N(v_j)$ for any $j < i$. There are $|A| \geq dn/2$ suitable choices for v_1. For each $i \in \{2, \ldots, p'\}$ we have chosen $i-1$ vertices prior to choosing v_i, so there are at least $|S_i| - (i-1)(pn - \delta^*(G)) \geq p'n - (p'-1)(n + \alpha n) \geq (1 - p'\alpha)n$ suitable choices for v_i. Finally, since $v_1 \in A$ has at least $(p - p')nd/2 \geq nd/2$ non-neighbours in $V_{p'+1} \setminus S_{p'+1}$, and at most $|V_{p'+1}| - \delta^*(G) \leq n + \alpha n$ non-neighbours in $V_{p'+1}$ in total, it has fewer than $(1 - d/2 + \alpha)n \leq (1 - d/3)n$ non-neighbours in $S_{p'+1}$. This means that there are at least $|S_{p'+1}| - (1 - d/3)n - (p'-1)(pn - \delta^*(G)) \geq dn/4$ suitable choices for $v_{p'+1}$. Together we conclude that there are at least $(dn/4)(dn/2)((1 - p'\alpha)n)^{p'-1} \geq 2\beta n^{p'+1}$

copies of $K_{p'+1}$ in $G[S]$. Each such copy can be extended to a copy of K_p in G with more than p' vertices in S by choosing $v_i \in V_i$ for each $p'+2 \leq i \leq p$ in turn, so that each v_i chosen is a neighbour of every v_j with $j \leq i$. For each $p'+2 \leq i \leq p$ there are at least $pn - (i-1)(pn - \delta^*(G)) \geq pn - (p-1)(n+\alpha n) \geq (1 - p\alpha)n$ suitable choices for v_i, so we obtain at least $2\beta n^{p'+1}((1-p\alpha)n)^{p-p'-1} \geq \beta n^p$ such copies of K_p.

For (ii), introduce new constants with $\beta \ll \gamma \ll \beta' \ll d_1 \ll d_2 \ll d$, and suppose for a contradiction that there are fewer than βn^{p+1} copies of K_{p+1} in G. Say that a vertex $x \in V(G)$ is *bad* if it lies in at least $\sqrt{\beta} n^p$ copies of K_{p+1} in G, and let X be the set of all bad vertices. Then $\sqrt{\beta} n^p |X| \leq r\beta n^{p+1}$, so $|X| \leq r\sqrt{\beta} n$. We now show that for any $i \in [r]$, any vertex $v \in V_i \setminus X$ has at most $(p-1)n + \gamma n$ neighbours in V_j for any $j \neq i$. Without loss of generality we consider the case $i = 1$, i.e. $v \in V_1 \setminus X$. Suppose for a contradiction that $|N(v) \cap V_j| > (p-1)n + \gamma n$ for some j, say $j = p+1$. Then we may greedily form a copy of K_{p+1} in G containing v by choosing $x_2, \ldots, x_{p+1}$ with $x_i \in V_i$ for each i so that each x_i is a neighbour of $v, x_2, \ldots, x_{i-1}$. We have at least $pn - (i-1)(pn - \delta^*(G)) \geq (p-i+1)n - (i-1)\alpha n \geq n/2$ choices for each x_i with $i \in \{2, \ldots, p\}$, and at least $|N(v) \cap V_{p+1}| - (p-1)(pn - \delta^*(G)) \geq (p-1)n + \gamma n - (p-1)(n+\alpha n) \geq \gamma n/2$ choices for x_{p+1}. Thus there are at least $\gamma n^p / 2^p \geq \sqrt{\beta} n^p$ copies of K_{p+1} in G containing v, a contradiction to $v \notin X$.

Now we fix some $v \in V_1 \setminus X$ and use the neighbourhood of v to impose structure on the rest of the graph. We choose a set $S_j \subseteq V_j$ of size $(p-1)n$ which contains or is contained in $N(v) \cap V_j$ for each $j \geq 2$. If $d(S_i, V_j \setminus S_j) < 1 - d_1$ for some $i, j \geq 2$ with $i \neq j$, then as in part (i) we can find at least $2\beta' n^p$ copies of K_p in $\bigcup_{i \geq 2} S_i$. At least $\beta' n^p$ of these are contained in $N(v)$, and so form copies of K_{p+1} with v, another contradiction. So we may suppose that $d(S_i, V_j \setminus S_j) \geq 1 - d_1$ for any $i, j \geq 2$ with $i \neq j$. We now partition V_1 into sets A, B, C as follows. Let A consist of all vertices $u \in V_1$ with $|N(u) \cap (V_j \setminus S_j)| \leq d_2 n$ for every $2 \leq j \leq r$. Let B consist of all vertices $u \in V_1$ with $|N(u) \cap (V_j \setminus S_j)| \geq (1 - d_2)n$ for every $2 \leq j \leq r$. Let $C = V_1 \setminus (A \cup B)$ consist of all remaining vertices of V_1. Next we bound the sizes of each of these sets. By definition of A we have $e(A, V_2 \setminus S_2) \leq d_2 n |A|$, so some vertex in $V_2 \setminus S_2$ has at most $d_2 |A|$ neighbours in A. So $pn - |A| + d_2|A| \geq \delta^*(G) \geq (p-1)n - \alpha n$, from which we obtain $|A| \leq (1 + 2d_2)n$. Next note that by definition of B we have $e(B, V_2 \setminus S_2) \geq (1 - d_2)n|B|$. So at least $n/2$ vertices of $V_2 \setminus S_2$ have at least $(1 - 2d_2)|B|$ neighbours in B. At least one of these vertices is not bad, so by our earlier observation has at most $(p-1)n + \gamma n$ neighbours in V_1. Then $(1 - 2d_2)|B| \leq (p-1)n + \gamma n$, so $|B| \leq (1 + 3d_2)(p-1)n$.

To bound $|C|$ we show that $C \subseteq X$. Consider any vertex $w \in C$. Without loss of generality $|N(w) \cap (V_{p+1} \setminus S_{p+1})| > d_2 n$ and $|N(w) \cap (V_p \setminus S_p)| < (1 - d_2)n$. Choose greedily a vertex $x_j \in S_j$ for each $2 \leq j \leq p$ so that x_j is a neighbour of $w, x_2, \ldots, x_{j-1}$ and satisfies $|N(x_j) \cap (V_{p+1} \setminus S_{p+1})| \geq (1 - \sqrt{d_1})n$. To see that this is possible for each $2 \leq j \leq p$, note that since $d(S_j, V_{p+1} \setminus S_{p+1}) \geq 1 - d_1$, at most $(p-1)\sqrt{d_1} n$ vertices $x_j \in S_j$ fail the latter condition. Note also that at least $|S_j| - |S_j \setminus N(w)| - \sum_{2 \leq i < j} |S_j \setminus N(x_i)| \geq (p-1)n - |S_j \setminus N(w)| - (j-2)(n+\alpha n)$ vertices $x_j \in S_j$ satisfy the neighbourhood condition. For $j < p$ this gives at least $n/2$ suitable choices for x_j. On the other hand, for $j = p$ we have $|N(w) \cap (V_p \setminus S_p)| < (1 - d_2)n$, which implies that $|S_j \setminus N(w)| \leq n - d_2 n + \alpha n$, so we have at least $d_2 n / 2$ suitable choices for x_j. So we may form at least $d_2(n/2)^{p-1}$ copies of K_p containing w in this manner. By construction, each x_j in any such copy has at most $\sqrt{d_1} n$ non-neighbours in $V_{p+1} \setminus S_{p+1}$. Since w has at least $d_2 n$ neighbours in $V_{p+1} \setminus S_{p+1}$, we find

a total of at least $(d_2 n - p\sqrt{d_1}n)d_2(n/2)^{p-1} \geq \sqrt{\beta}n^p$ copies of K_{p+1} in G containing w, so $w \in X$. We deduce that $C \subseteq X$, so $|C| \leq |X| \leq r\sqrt{\beta}n$.

We therefore have $|B| \geq pn - |A| - |C| \geq (p-1)n - 3d_2 n$, so $|B| = (1 \pm 3d_2)(p-1)n$. Let S_1 be a set of size $(p-1)n$ which either contains or is contained in B. Then for any $2 \leq j \leq r$ we have $e(S_1, V_j \setminus S_j) \geq \min\{|B|, |S_1|\}(1-d_2)n \geq (1-4d_2)(p-1)n^2$, so $d(S_1, V_j \setminus S_j) \geq 1 - 4d_2$. Also, at most $4d_2(p-1)n$ vertices of $V_1 \setminus S_1$ lie in $B \cup C$, so for any $2 \leq j \leq r$ we have $e(V_1 \setminus S_1, V_j \setminus S_j) \leq d_2 n^2 + 4d_2(p-1)n^2 \leq 4pd_2 n^2$. But $e(V_1 \setminus S_1, V_j) \geq \delta^*(G)n \geq (p-1-\alpha)n^2$, so we obtain $e(V_1 \setminus S_1, S_j) \geq (p-1-\alpha)n^2 - 4pd_2 n^2$, and so $d(V_1 \setminus S_1, S_j) \geq 1 - 9d_2$. Recall also that $d(S_i, V_j \setminus S_j) \geq 1 - d_1$ for any $i, j \geq 2$ with $i \neq j$. Since $d_1, d_2 \ll d$ we conclude that G is d-splittable with respect to the sets S_j for $j \in [r]$. This is a contradiction, so (ii) holds. $\qquad\square$

5. Avoiding divisibility barriers

Let G be an r-partite graph with vertex classes of size pn such that $\delta^*(G) \geq (p-1)n - \alpha n$, and let $J = J(G)$ be the clique p-complex of G. In the previous section we saw that if G is not d-splittable (for small d), then there is no space barrier to a perfect matching in J_p. In this section we instead consider divisibility barriers. Indeed, we shall see in the second subsection that if $p \geq 3$ and G is not d-splittable, then J_p cannot be close to a divisibility barrier. However, for $p = 2$ there is another possibility, namely that G has the structure of $V^1 \cup V^2$ in Construction 1.2. There we described both V^1 and V^2 as rows, but with the terminology of the previous section they should be considered as a single row. We consider this case in the first subsection. Note that here we have $J_p = J_2 = G$.

5.1. **Pair-complete rows.** Let G be an r-partite graph with vertex classes $V_1, \ldots, V_r$ each of size $2n$. We say that G is d-pair-complete (with respect to $S = \bigcup_{j \in [r]} S_j$) if there exist sets $S_j \subseteq V_j$, $j \in [r]$ each of size n such that $d(S_i, S_j) \geq 1 - d$, $d(V_i \setminus S_i, V_j \setminus S_j) \geq 1 - d$ and $d(S_i, V_j \setminus S_j) \leq d$ for any $i, j \in [r]$ with $i \neq j$. That is, G consists of two halves S and $V \setminus S$, where each half is an almost-complete r-partite graph, and there are few edges between halves. We will show that if G is close to a divisibility barrier, then G is either d-splittable or d-pair-complete. For this we need the following proposition.

Proposition 5.1. Let $r \geq 2$ and H be an r-partite graph whose parts $V_1, \ldots, V_r$ each have size 2. Suppose that $\delta^*(H) \geq 1$, and for any $A \subseteq V$ such that $|A \cap V_j| = 1$ for every $j \in [r]$ there is

(i) an edge ab with $a, b \in A$ or $a, b \notin A$, and
(ii) an edge ab with $a \in A$ and $b \notin A$.

Then for some $V_j = \{x, y\}$ there is a path of even length between x and y.

Proof. Suppose for a contradiction that there is no such path. Then for any $i, j \in [r]$ with $i \neq j$, any vertex $v \in V_i$ must have precisely one neighbour in V_j. Indeed, v must have at least one neighbour in V_j since $\delta^*(H) \geq 1$, but cannot have two since then we obtain a path of length two between these two neighbours in V_j. So for any $i \neq j$ the graph $H[V_i, V_j]$ consists of two disjoint edges. Write $V_1 = \{x_1, y_1\}$, and for each $2 \leq i \leq r$ let $x_i \in V_i$ be adjacent to x_{i-1} and $y_i \in V_i$ be adjacent to y_{i-1}. There are then two possibilities for $H[V_r, V_1]$.

The first case is that x_r is adjacent to x_1 and y_r to y_1. By property (ii) of $A = \{x_1, \ldots, x_r\}$, there must be some edge $x_i y_j$. Fix such an i and j, and consider the paths $x_1 \ldots x_i y_j \ldots y_1$

and $x_1 \ldots x_i y_j \ldots y_r y_1$ between x_1 and y_1. They have lengths $i + j - 1$ and $i + r - j + 1$, which must both be odd, so $i + j$ and r are both even. This argument shows that $i' + j'$ must be even for any edge $x_{i'} y_{j'}$. So by property (i) of $A = \{x_{i'} : i' \text{ even}\} \cup \{y_{i'} : i' \text{ odd}\}$ there must be an edge $x_{i'} x_{j'}$ or $y_{i'} y_{j'}$ such that $i' + j'$ is even. Without loss of generality we may assume the former, and that $j' \neq i$. If $i' = i$ then $x_{j'} x_i y_j y_{j+1} \ldots y_{j'}$ is a path whose length is congruent to $j' - j + 2 \equiv j' + i' - j - i + 2 \equiv 0$ modulo 2, giving a contradiction. On the other hand, if $i' \neq i$ then let P be a path from $x_{j'}$ to x_i which does not contain $x_{i'}$ and whose length has the same parity as $i - j'$ (this must exist since r is even and $x_i, x_{i'}$ and $x_{j'}$ all lie on the cycle $x_1 x_2 \ldots x_r x_1$). Then $x_{i'} x_{j'} P x_i y_j y_{j+1} \ldots y_{i'}$ is a path whose length is congruent to $i - j' + i' - j + 2 \equiv 0$ modulo 2, again giving a contradiction.

The second case is that x_r is adjacent to y_1 and y_r to x_1. Then $x_1 \ldots x_r y_1$ must have odd length, so r is odd. By property (i) of $A = \{x_i : i \text{ even}\} \cup \{y_i : i \text{ odd}\}$ we must have either an edge $x_i y_j$ with $i + j$ odd, or an edge $x_i x_j$ with $i + j$ even, or an edge $y_i y_j$ with $i + j$ even. In the first case $x_1 \ldots x_i y_j \ldots y_1$ is a path of even length $i + j - 1$. On the other hand, for the second case we may assume that $i < j$, whereupon $x_1 \ldots x_i x_j \ldots x_r y_1$ is a path of even length $r - j + i + 1$, and the third case is similar. Thus we have a contradiction in all cases, so the required path exists. $\qquad\square$

Now we can deduce the required structure for divisibility barriers in G when $p = 2$.

Lemma 5.2. *Suppose that $1/n \ll \mu, \alpha \ll d \ll 1/r$ and $r \geq 2$. Let $\mathcal{P}'$ partition a set V into parts $V_1, \ldots, V_r$ each of size $2n$. Suppose G is a $\mathcal{P}'$-partite graph with $\delta^*(G) \geq n - \alpha n$, and that there exists a partition $\mathcal{P}$ refining $\mathcal{P}'$ into parts each of size at least $n - \mu n$ such that $L_{\mathcal{P}}^{\mu}(G)$ is incomplete with respect to $\mathcal{P}'$. Then G is d-splittable or d-pair-complete.*

Proof. We can assume that $L_{\mathcal{P}}^{\mu}(G)$ is transferral-free (recall from Section 2 that this means that $L_{\mathcal{P}}^{\mu}(G)$ does not contain any difference of basis vectors $\mathbf{u}_i - \mathbf{u}_j$, for some $i \neq j$ which index subparts of the same part of $\mathcal{P}'$). Thus for any $i, j \in [r]$ with $i \neq j$, distinct parts A, B of $\mathcal{P}$ contained in V_i, and part C of $\mathcal{P}$ contained in V_j, we cannot have both $e(A, C) \geq \mu(2rn)^2$ and $e(B, C) \geq \mu(2rn)^2$. Since all parts of $\mathcal{P}$ have size at least $n - \mu n$, each part of $\mathcal{P}'$ is partitioned into at most two parts of $\mathcal{P}$. We can assume that $\mathcal{P}$ is a strict refinement of $\mathcal{P}'$, so without loss of generality $\mathcal{P}$ partitions V_1 into two parts V_1^1 and V_1^2. Next we note that there cannot be any part V_j that is not partitioned into two parts by $\mathcal{P}$. For otherwise, letting $A = V_1^1$, $B = V_1^2$ and $C = V_j$, we have $e(A, C) \geq |A|\delta^*(G) \geq \mu(2rn)^2$ and $e(B, C) \geq |B|\delta^*(G) \geq \mu(2rn)^2$, which contradicts $L_{\mathcal{P}}^{\mu}(G)$ being transferral-free, as described above. Thus each V_j is partitioned into two parts V_j^1 and V_j^2 by $\mathcal{P}$. Form an auxiliary graph H on $2r$ vertices, where for each $i \in \{1, 2\}$ and $j \in [r]$ we have a vertex x_j^i of H corresponding to the part V_j^i of $\mathcal{P}$, and we have an edge $x_j^i x_{j'}^{i'}$ if and only if $e(V_j^i, V_{j'}^{i'}) \geq \mu(2rn)^2$. Since $\delta^*(G) \geq n - \alpha n$ and each part of $\mathcal{P}$ has size at least $n - \mu n$, we have $\delta^*(H) \geq 1$. Furthermore, for any edge ab of H, the at least $\mu(2rn)^2$ edges of G corresponding to this edge ensure that $\mathbf{u}_a + \mathbf{u}_b \in L_{\mathcal{P}}^{\mu}(G)$, where we consider a and b to index the parts of $\mathcal{P}$ to which they correspond. It follows that H cannot contain a path of even length between x_j^1 and x_j^2 for any $j \in [r]$. Indeed, if $x_j^1 = y_0, y_1, \ldots, y_{2m} = x_j^2$ are the vertices of such a path then, since $\mathbf{u}_{y_{\ell-1}} + \mathbf{u}_{y_\ell} \in L_{\mathcal{P}}^{\mu}(G)$ for each $\ell \in [2m]$, we have

$$\mathbf{u}_{x_j^1} - \mathbf{u}_{x_j^2} = \mathbf{u}_{y_0} - \mathbf{u}_{y_{2m}} = \sum_{\ell \in [m]} \left((\mathbf{u}_{y_{2\ell-2}} + \mathbf{u}_{y_{2\ell-1}}) - (\mathbf{u}_{y_{2\ell-1}} + \mathbf{u}_{y_{2\ell}}) \right) \in L_{\mathcal{P}}^{\mu}(G),$$

a contradiction to $L_{\mathcal{P}}^{\mu}(G)$ being transferral-free.

We may therefore apply Proposition 5.1 to deduce that there exists some $A \subseteq V(H)$ with $|A \cap \{x_j^1, x_j^2\}| = 1$ for each $j \in [r]$ such that either

  (i) H contains no edges ab with $a, b \in A$ or $a, b \notin A$, or
  (ii) H contains no edges ab with $a \in A$ and $b \notin A$.

For each $j \in [r]$ let $S_j' = V_j^1$ if $x_j^1 \in A$, and V_j^2 otherwise. For $j \in [r]$, let $S_j \subseteq V_j$ be a set of size n that contains or is contained by S_j'. Note that for any $j \neq j'$, we have $e(S_j, V_{j'}) \geq n\delta^*(G) \geq (1-\alpha)n^2$, and similarly $e(V_j \setminus S_j, V_{j'}) \geq (1-\alpha)n^2$. In case (i), for any $j \neq j'$ we have $e(S_j', S_{j'}') \leq \mu(2rn)^2$. This implies $e(S_j, V_{j'} \setminus S_{j'}) \geq (1 - \alpha - 8r^2\mu)n^2$, and so $G[S_j, V_{j'} \setminus S_{j'}]$ has density at least $1 - d$. Since j and j' were arbitrary, we may conclude that G is d-splittable. On the other hand, in case (ii), for any $j \neq j'$ we have $e(S_j', V_{j'} \setminus S_{j'}') \leq \mu(2rn)^2$. This implies $e(S_j, S_{j'}) \geq (1 - \alpha - 8r^2\mu)n^2$, and so $G[S_j, S_{j'}]$ has density at least $1 - d$. Similarly, $G[V_j \setminus S_j, V_{j'} \setminus S_{j'}]$ has density at least $1 - d$, and $G[S_j, V_{j'} \setminus S_{j'}]$ has density at most d for any $j \neq j'$, so G is d-pair-complete. $\square$

5.2. Avoiding divisibility barriers for $p > 2$.

We next show that for $p > 2$, if G is not d-splittable then J_p is not close to a divisibility barrier.

Lemma 5.3. *Suppose that $1/n \ll \mu, \alpha \ll d \ll 1/r$ and $3 \leq p \leq r$. Let $\mathcal{P}'$ partition a set V into vertex classes $V_1, \ldots, V_r$ each of size pn. Suppose G is a $\mathcal{P}'$-partite graph with $\delta^*(G) \geq (p-1)n - \alpha n$ and let $J = J(G)$ be the clique p-complex of G. Suppose $\mathcal{P}$ is a partition refining $\mathcal{P}'$ into parts each of size at least $n - \mu n$ such that $L_{\mathcal{P}}^{\mu}(J_p)$ is incomplete with respect to $\mathcal{P}'$. Then G is d-splittable.*

Proof. We introduce new constants with $\mu, \alpha \ll \mu' \ll c \ll \gamma \ll \gamma' \ll \gamma'' \ll d$. We can assume that $\mathcal{P}$ is a strict refinement of $\mathcal{P}'$, so without loss of generality $\mathcal{P}$ partitions V_1 into parts V_1^i, $i \in [m]$ with $2 \leq m \leq p$. As in the proof of Lemma 5.2, we can also assume that $L_{\mathcal{P}}^{\mu}(J_p)$ is transferral-free, meaning that it does not contain any difference of basis vectors $\mathbf{u}_i - \mathbf{u}_j$, for some $i \neq j$ which index subparts of the same part of $\mathcal{P}'$. Thus for any p vertex classes $V_{i_1}, \ldots, V_{i_p}$, distinct parts U_{i_1}, U_{i_1}' of $\mathcal{P}$ contained in V_{i_1}, and parts $U_{i_j} \subseteq V_{i_j}$ for $2 \leq j \leq p$, it cannot be that $\bigcup_{j=1}^p U_{i_j}$ and $U_1' \cup \bigcup_{j=2}^p U_{i_j}$ both have at least $\mu|V(G)|^p = \mu(rpn)^p$ edges of J_p (i.e. copies of K_p). We use this to deduce the following properties, which control the typical behaviour of neighbourhoods and certain pairwise intersections of neighbourhoods. Note that the bound in (c) is close to the lower bound on $\delta^*(G)$, so it says that G is mostly approximately regular from the point of view of V_1.

Claim 5.4.
  (a) *There are at most $\mu'n^2$ pairs (x, y) for which there exist $i \in [m]$ and $j \in [r]$ such that $x \in V_1^i$, $y \in V_1 \setminus V_1^i$ and $|V_j \cap N(x) \cap N(y)| \geq (p-2)n + \mu'n$.*
  (b) *There are at most $\mu'n^3$ triples (x, y, z) for which there exist $i \in [m]$ and $j \in [r]$ such that $x \in V_1^i$, $y \in V_1 \setminus V_1^i$, $z \in V \setminus V_1$, $xz, yz \in G$ and $|V_j \cap N(x) \cap N(z)| \geq (p-2)n + \mu'n$.*
  (c) *For any $2 \leq j \leq r$ there are at most $2\mu'n$ vertices $x \in V_1$ such that $|N(x) \cap V_j| \geq (p-1)n + 2\mu'n$.*

For (a), suppose for a contradiction that there are more than $\mu'n^2$ such pairs. Without loss of generality there are at least $\mu'n^2/rp^2$ pairs (x, y) with $x \in V_1^1$ and $y \in V_1^2$ such that

$|V_p \cap N(x) \cap N(y)| \geq (p-2)n + \mu'n$. For each such pair, we consider greedily choosing $w_j \in V_j$, $2 \leq j \leq p$ such that $xw_2 \ldots w_p$ and $yw_2 \ldots w_p$ are copies of K_p. The number of choices for w_j is $N_j \geq |V_j \cap N(x) \cap N(y)| - \sum_{i=2}^{j-1} |V_j \setminus N(w_i)|$. For $2 \leq j \leq p-1$, we have $N_j \geq pn - (p-1)(pn - \delta^*(G)) \geq (1 - (p-1)\alpha)n > n/2$. Also, $N_p \geq (p-2)n + \mu'n - (p-2)(pn - \delta^*(G)) \geq \mu'n - (p-2)\alpha n \geq \mu'n/2$. Considering all such pairs (x, y), we obtain at least $(\mu')^2 n^{p+1}/2^{p-1} rp^2$ such $(p+1)$-tuples $(x, y, w_2, \ldots, w_p)$. There are at most p^{r+1} possible indices for such a $(p+1)$-tuple, so we may choose $\mu(rpn)^{p+1}$ such $(p+1)$-tuples which all have the same index; let $(x, y, w_2, \ldots, w_p)$ be a representative of this collection. Then there must be at least $\mu(rpn)^p$ edges of J_p with index $\mathbf{i}(\{x, w_2, \ldots, w_p\})$, and at least $\mu(rpn)^p$ edges of J_p with index $\mathbf{i}(\{y, w_2, \ldots, w_p\})$. But this contradicts $L_{\mathcal{P}}^\mu(J_p)$ being transferral-free. A very similar argument applies for (b). Indeed, suppose for a contradiction that there are more than $\mu'n^3$ such triples. Say there are at least $\mu'n^3/r^2p^2$ triples (x, y, z) with $x \in V_1^1$, $y \in V_1^2$, $z \in V_2$ such that $xz, yz \in G$ and $|V_p \cap N(x) \cap N(z)| \geq (p-2)n + \mu'n$. For each such triple we consider greedily choosing $w_j \in V_j$, $3 \leq j \leq p$ such that $xzw_3 \ldots w_p$ and $yzw_3 \ldots w_p$ are copies of K_p. The number of choices for w_j is $N_j \geq |V_j \cap N(x) \cap N(z)| - \sum_{i=3}^{j-1} |V_j \setminus N(w_i)| - |V_j \setminus N(y)|$. Thus the same calculation as in (a) gives a contradiction. For (c), suppose for a contradiction that there are at least $2\mu'n$ such vertices x. For each such x, and each choice of $y \in V_1$ in a different part of $\mathcal{P}$ to x, we have $|V_j \cap N(x) \cap N(y)| \geq |N(x) \cap V_j| - (pn - \delta^*(G)) \geq (p-2)n + \mu'n$. There are at least $n - \mu n$ choices of y for each x, so this contradicts (a). Thus we have proved Claim 5.4.

Now for each $i \in [m]$ and $2 \leq j \leq r$ let X_j^i consist of all vertices of V_j which have at most $|V_1^i| - \gamma n$ neighbours in V_1^i. Bearing in mind the row structure we are aiming for, the intuition is that X_j^i should approximate the jth part of the ith row. We show the following properties that agree with this intuition: the size of X_j^i is roughly correct, and the intended diagonal densities are close to 1.

Claim 5.5.
  (d) *For any $i \in [m]$ and $2 \leq j \leq r$ we have $|X_j^i| \geq |V_1^i| - \gamma'n > n/2$.*
  (e) *For any $i, i' \in [m]$ with $i \neq i'$ and any $2 \leq j \leq r$ we have $d(V_1^i, X_j^{i'}) \geq 1 - c$.*
  (f) *For any $2 \leq j \leq r$ at most γn vertices $v \in V_j$ lie in more than one of the sets X_j^i. Thus $|X_j^i| \leq |V_1^i| + p\gamma'n$ for any $i \in [m]$, and all but $2p\gamma'n$ vertices of V_j lie in $\bigcup_{i \in [m]} X_j^i$.*
  (g) *For any $i, i' \in [m]$ with $i \neq i'$ and any $2 \leq j < j' \leq r$ we have $d(X_j^i, X_{j'}^{i'}) \geq 1 - d/2$.*

For (d), note that $e(V_1^i, V_j) \leq ((p-1)n + 2\mu'n)|V_1^i| + 2\mu'pn^2 \leq (p-1+5p\mu')n|V_1^i|$ by (c). Also,
$$e(V_1^i, V_j) \geq |X_j^i|(|V_1^i| - (pn - \delta^*(G))) + (pn - |X_j^i|)(|V_1^i| - \gamma n)$$
$$\geq pn(|V_1^i| - \gamma n) - |X_j^i|(1 + \alpha - \gamma)n.$$
Therefore $|X_j^i|(1 + \alpha - \gamma)n \geq (1 - 5p\mu')|V_1^i|n - p\gamma n^2$, which gives (d). Next suppose for a contradiction that (e) is false, say that $d(V_1^1, X_2^2) < 1 - c$. Let A be the set of vertices in X_2^2 with fewer than $(1 - c/2)|V_1^1|$ neighbours in V_1^1. Write $|A| = a|X_2^2|$. Then $(1-a)(1-c/2) \leq d(V_1^1, X_2^2) < 1 - c$, so $a \geq c/2$. Thus $|A| \geq cn/4$ by (d). Each vertex in A has fewer than $|V_1^1| - cn/4$ neighbours in V_1^1, and also fewer than $|V_1^2| - \gamma n$ neighbours in V_1^2 by definition of X_2^2. This gives at least $(cn/4)^2 \gamma n$ triples (x, y, z) with $x \in V_1^1$, $y \in V_1^2$, and $z \in X_2^2$ such that

$xz, yz \notin G$. At least $\mu'n^2$ pairs (x, y) therefore lie in at least $2\mu'n$ such triples. For each of these pairs we have $|V_2 \cap N(x) \cap N(y)| \geq pn - 2(pn - \delta^*(G)) + 2\mu'n \geq (p-2)n + \mu'n$. However, this contradicts (a), so (e) holds. For (f), suppose for a contradiction that $A := X_2^i \cap X_2^{i'}$ has size at least $\gamma n/p^2$, for some $i, i' \in [m]$ with $i \neq i'$. By definition, each vertex of A has at most $|V_1^i| - \gamma n$ neighbours in $|V_1^i|$, so $d(A, V_1^i) \leq 1 - \gamma/p$. But then $d(X_2^{i'}, V_1^i) \leq 1 - \gamma^2/p^4$, contradicting (e). So $|A| < \gamma n/p^2$, and summing over all possible values of i and i' we obtain the first statement of (f). This implies that $\sum_{i \in [m]} |X_j^i| \leq pn + m\gamma n$, so $\sum_{i \in [m]} (|X_j^i| - |V_1^i|) \leq m\gamma n$. In combination with (d) this implies $|X_j^i| \leq |V_1^i| + p\gamma'n$ for any $i \in [m]$. Also, (d) gives $\left| \bigcup_{i \in [m]} X_j^i \right| \geq \sum_{i \in [m]} (|V_1^i| - \gamma'n) - m\gamma n \geq pn - 2p\gamma'n$, so (f) holds.

Finally, suppose for a contradiction that (g) is false, say $d(X_2^1, X_3^2) < 1 - d/2$. Without loss of generality we have $|V_1^2| \leq pn/2$. Let A be the set of vertices in X_3^2 with fewer than $(1 - d/4)|X_2^1|$ neighbours in X_2^1. (We re-use A to avoid excessive notation.) Write $|A| = a|X_3^2|$. Then $(1 - a)(1 - d/4) \leq d(X_2^1, X_3^2) < 1 - d/2$, so $a \geq d/4$. Thus $|A| \geq dn/8$ by (d). Each vertex in A has fewer than $|X_2^1| - dn/8$ neighbours in X_2^1, and also fewer than $|V_1^2| - \gamma n$ neighbours in V_1^2 by definition of X_3^2. This gives a set T of at least $(dn/8)^2\gamma n$ triples (x, z, w) with $x \in V_1^2$, $z \in X_2^1$ and $w \in X_3^2$ such that $xw, zw \notin G$. Furthermore, since by (e) we have $d(V_1^2, X_2^1) \geq 1 - c$, all but at most $c(pn)^3$ triples in T have the additional property that $xz \in G$. Let P be the set of pairs (x, z) with $xz \in G$, $x \in V_1^2$, $z \in X_2^1$ that lie in at least $2\mu'n$ triples of T. Then $|T| - c(pn)^3 \leq |P|pn + (pn)^2 2\mu'n$, so $|P| \geq 3\mu'n^2$, say. Since $|V_1^2| \leq pn/2$ and $p \geq 3$, for each $(x, z) \in P$ there are more than $pn/2 - (pn - \delta^*(G)) > n/3$ vertices y such that $y \in V_1 \setminus V_1^2$ and $yz \in G$. Note that this is a key use of the assumption $p \geq 3$, so we had to deal with the case $p = 2$ separately in the previous subsection. There are therefore more than $\mu'n^3$ triples (x, y, z) with $x \in V_1^2$, $y \in V_1 \setminus V_1^2$ and $z \in X_2^1$ such that $xz, yz \in G$ and $(x, z) \in P$. However, for any $(x, z) \in P$ we have $|V_3 \cap N(x) \cap N(z)| \geq pn - 2(pn - \delta^*(G)) + 2\mu'n \geq (p-2)n + \mu'n$, which contradicts (b). Thus (g) holds, proving Claim 5.5.

To complete the proof, we also need to show that the size of each part of V_1 is close to an integer multiple of n. Since each part of V_1 has size at least $n - \mu n$, this will be true if V_1 has a part of size close to $(p-1)n$. So for the final claim we assume that V_1 does not have such a large part; in this case we extend (c) by showing that the bipartite graph induced by any pair of vertex classes is mostly approximately regular. We then show that most vertices in V_1 have sparse non-neighbourhoods, before finally deducing the required statement on the sizes of the parts of V_1.

Claim 5.6. *Suppose V_1 does not have a part of size at least $(p-1)n - \gamma n$. Then*

(h) *for any $j, j' \in [r]$ there are at most cn vertices $z \in V_j$ such that $|N(z) \cap V_{j'}| \geq (p-1)n + 2\mu'n$,*

(j) *for a set V_1' of all but at most γn vertices $x \in V_1$ we have $|V_j \setminus N(x)| = n \pm 2\mu'n$ and $d(V_j \setminus N(x), V_{j'} \setminus N(x)) \leq \gamma$ for any $2 \leq j, j' \leq r$ with $j \neq j'$, and*

(k) *for each i there is an integer p_i such that $|V_1^i| = p_i n \pm \gamma''n$.*

For (h), note that by our assumption on the part sizes of V_1, any such z lies in at least $\gamma^2 n^2/2$ triples (x, y, z) such that $xz, yz \in G$ and x and y lie in different parts of V_1. Any such triple is counted by (b), as $|V_{j'} \cap N(x) \cap N(z)| \geq |N(z) \cap V_{j'}| - (pn - \delta^*(G)) \geq (p-2)n + \mu'n$, so there can be at most cn such vertices z, as claimed. For (j) we introduce the following notation: $N_j(x) := N(x) \cap V_j$ is the set of neighbours of x in V_j, and $N_j^c(x) := V_j \setminus$

$N(x)$ is the set of non-neighbours of x in V_j. Fix some j and j', and suppose $x \in V_1$ is such that $d(N_j^c(x), N_{j'}^c(x)) \geq \gamma$ and $|N_j(x)|, |N_{j'}(x)| \leq (p-1)n + 2\mu'n$. We can estimate $d(N_j(x), N_{j'}^c(x))$ as follows. Write $e(N_j(x), N_{j'}^c(x)) = e(V_j, N_{j'}^c(x)) - e(N_j^c(x), N_{j'}^c(x))$. Then by (h) we have $e(V_j, N_{j'}^c(x)) = \sum_{v \in N_{j'}^c(x)} |N(v) \cap V_j| \leq ((p-1)n + 2\mu'n)|N_{j'}^c(x)| + cn \cdot pn$. Also, $e(N_j^c(x), N_{j'}^c(x)) \geq \gamma|N_j^c(x)||N_{j'}^c(x)| \geq \gamma(n - 2\mu'n)|N_{j'}^c(x)|$ by choice of x. This gives

$$e(N_j(x), N_{j'}^c(x)) \leq (p - 1 + 2\mu' - (1 - 2\mu')\gamma)n|N_{j'}^c(x)| + cpn^2.$$

Since $|N_j(x)| \geq \delta^*(G) \geq (p-1)n - \alpha n$, and $|N_{j'}^c(x)| \geq n - 2\mu'n$ by choice of x, we deduce that $d(N_j(x), N_{j'}^c(x)) \leq 1 - \gamma/2$. Let A be the set of vertices in $N_j(x)$ with fewer than $(1 - \gamma/4)|N_{j'}^c(x)|$ neighbours in $N_{j'}^c(x)$. Write $|A| = a|N_j(x)|$. Then $(1 - a)(1 - \gamma/4) \leq d(N_j(x), N_{j'}^c(x)) \leq 1 - \gamma/2$, so $a \geq \gamma/4$, and $|A| \geq \gamma n/4$. Moreover, by definition of A any $z \in A$ has at least $|N_{j'}(z)| - (1 - \gamma/4)|N_{j'}^c(x)|$ neighbours in $N_{j'}(x)$. Since $|N_{j'}(z)| \geq \delta^*(G) \geq (p-1)n - \alpha n$ and $|N_{j'}^c(x)| \leq pn - \delta^*(G) \leq n + \alpha n$, it follows that for any $z \in A$ we have

$$|V_{j'} \cap N(x) \cap N(z)| \geq (p-1)n - \alpha n - (1 - \tfrac{\gamma}{4})(n + \alpha n)$$
$$\geq (p-2)n + \gamma n/4 - 2\alpha n \geq (p-2)n + \mu'n.$$

Furthermore, since V_1 does not have a part of size at least $(p-1)n - \gamma n$, there must be at least $\gamma n/2$ neighbours y of z which lie in a different part of V_1 to x. There are at least $\gamma n/4$ choices for $z \in A$, so x lies in at least $\gamma^2 n^2/8$ triples (x, y, z) counted in (b). Thus there are at most $\frac{\mu'n^3}{\gamma^2 n^2/8} < \gamma n/2r^2$ such vertices $x \in V_1$ with $d(N_j^c(x), N_{j'}^c(x)) \geq \gamma$ and $|N_j(x)|, |N_{j'}(x)| \leq (p-1)n + 2\mu'n$. Since by (c) at most $4\mu'n$ vertices do not satisfy the latter condition, summing over all $j, j' \in [r]$ gives (j). (Every vertex $x \in V_1$ satisfies $|N_j(x)| \geq \delta^*(x) \geq (p-1)n - \alpha n$.)

For (k), consider any $x, y \in V_1'$ as defined in (j), and let $I_j^{xy} = N_j^c(x) \cap N_j^c(y)$ for $j = 2, 3$. We will show that either $|I_2^{xy}| \leq 3\sqrt{\gamma}n$ or $|I_2^{xy}| \geq (1 - 3\sqrt{\gamma})n$. For suppose that $I_2^{xy} > 3\sqrt{\gamma}n$. Let $B = N_3^c(x) \cup N_3^c(y)$. By definition of V_1' we have $e(I_2^{xy}, B) \leq e(N_2^c(x), N_3^c(x)) + e(N_2^c(y), N_3^c(y)) \leq 3\gamma n|B| \leq \sqrt{\gamma}|I_2^{xy}||B|$, so there is a vertex $z \in I_2^{xy}$ with $|N(z) \cap B| \leq \sqrt{\gamma}|B|$. Then $(1 - \sqrt{\gamma})|B| \leq |B \setminus N(z)| \leq |V_3 \setminus N(z)| \leq n + \alpha n$. This gives $|B| \leq (1 + 2\sqrt{\gamma})n$, so $|I_3^{xy}| = |V_3 \setminus N(x)| + |V_3 \setminus N(y)| - |B| \geq (1 - 3\sqrt{\gamma})n$. Now the same argument interchanging I_2^{xy} and I_3^{xy} shows that $|I_2^{xy}| \geq (1 - 3\sqrt{\gamma})n$, as required.

Next we define a relation $\sim$ on V_1' by $x \sim y$ if $|I_2^{xy}| \geq (1 - 3\sqrt{\gamma})n$. This is an equivalence relation, since if $|I_2^{xy}| \geq (1 - 3\sqrt{\gamma})n$ and $|I_2^{yz}| \geq (1 - 3\sqrt{\gamma})n$, then $|I_2^{xz}| \geq |I_2^{xy}| - |N_2^c(y) \setminus N_2^c(z)| \geq (1 - 3\sqrt{\gamma})n - ((n + \alpha n) - (1 - 3\sqrt{\gamma})n) > 3\sqrt{\gamma}n$, so $|I_2^{xy}| \geq (1 - 3\sqrt{\gamma})n$ as just shown. Let $C_1^1, \ldots, C_1^t$ be the equivalence classes of $\sim$, and arbitrarily choose a representative x_i of each equivalence class C_1^i.

Since each x_i lies in V_1', the sets $N_2^c(x_i)$ each have size $n \pm 2\mu'n$ by (j). Furthermore, any two such sets intersect in at most $3\sqrt{\gamma}n$ vertices, since the representatives x_i each lie in different equivalence classes. We cannot have $t > p$, as then $(p+1)(n - 2\mu'n) - \binom{p+1}{2}3\sqrt{\gamma}n \leq \left|\bigcup_{i=1}^{p+1} N_2^c(x_i)\right| \leq |V_2| \leq pn$ is a contradiction, so we must have $t \leq p$. Next, observe that any vertex in C_1^i has at most $(n + 2\mu'n) - (1 - 3\gamma)n \leq 4\gamma n$ neighbours in $N_2^c(x_i)$. So $e(C_1^i, N_2^c(x_i)) \leq 4\gamma n|C_1^i|$. By averaging, some vertex $v \in N_2^c(x_i)$ therefore has at most $4\gamma n|C_1^i|/(n - 2\mu'n) \leq 5\gamma|C_1^i|$ neighbours in C_1^i. So $(pn - |C_1^i|) + 5\gamma|C_1^i| \geq |N(v) \cap V_1| \geq$

$\delta^*(G) \geq (p-1)n - \alpha n$, which implies $|C_1^i| \leq n + 6\gamma n$. Since the $t \leq p$ equivalence classes C_1^i partition V_1', we deduce that $t = p$ and $|C_1^i| = n \pm \gamma' n$ for any $i \in [t]$.

Now we show that any equivalence class $C_1^{i'}$ must be essentially contained in some part V_1^i of V_1; by symmetry it suffices to show that this is true of C_1^1. So observe that since $|N_2^c(x_1)| \geq n - 2\mu' n$, by (f) we must have $|N_2^c(x_1) \cap X_2^i| \geq n/2p$ for some $i \in [m]$. Fix such an i, and suppose for a contradiction that $|C_1^1 \cap V_1^{i'}| \geq \gamma n$ for some $i' \neq i$. We observed above that any $v \in C_1^1$ has at most $4\gamma n$ neighbours in $N_2^c(x_1)$, so there are at least $n/2p - 4\gamma n \geq n/3p$ vertices of X_2^i which are not neighbours of v. Then $e(V_1^{i'}, X_2^i) \leq |V_1^{i'}||X_2^i| - |C_1^1 \cap V_1^{i'}|n/3p < (1-c)|V_1^{i'}||X_2^i|$, contradicting (e). We conclude that all but at most $p\gamma n$ vertices of C_1^1 lie in V_1^i, and thus that all but at most $p\gamma n$ vertices of any equivalence class lie in the same part of V_1. So for any i, i' we have either $|V_1^i \cap C_1^{i'}| \leq p\gamma n$ or $|V_1^i \cap C_1^{i'}| = |C_1^{i'}| \pm p\gamma n = n \pm 2\gamma' n$; (k) follows immediately since the classes $C_1^1, \ldots, C_1^t$ partition V_1' and $|V_1 \setminus V_1'| \leq \gamma n$ by (j). This completes the proof of Claim 5.6.

To complete the proof of Lemma 5.3, note that there exist integers $p_1, \ldots, p_m$ such that $|V_1^i| = p_i n \pm \gamma'' n$ for each $i \in [m]$. Indeed, if V_1 has a part of size at least $(p-1)n - \gamma n$, then since each part of $\mathcal{P}$ has size at least $n - \mu n$, we may assume that V_1 has two parts V_1^1 and V_1^2 with respective sizes $(p-1)n \pm \gamma n$ and $n \pm \gamma n$. On the other hand, if V_1 has no such part then the required integers p_i exist by (k). We partition V_1 into sets U_1^i with $|U_1^i| = p_i n$ for $i \in [m]$ such that each U_1^i either contains or is contained in some V_1^i. Then U_1^i contains at least $p_i n - \gamma'' n \geq |V_1^i| - 2\gamma'' n$ vertices of V_1^i for any $i \in [m]$. Furthermore, by (d) and (f), for each $2 \leq j \leq r$ we may partition V_j into sets U_j^i with $|U_j^i| = p_i n$ for $i \in [m]$ such that each U_j^i contains at least $|X_j^i| - 2p\gamma' n \geq p_i n - 2\gamma'' n$ vertices from X_j^i. By (e) and (g) we deduce that $d(U_j^i, U_{j'}^{i'}) \geq 1 - d$ whenever $i \neq i'$ and $j \neq j'$. In particular, $d(U_j^1, V_{j'} \setminus U_{j'}^1) \geq 1 - d$ for any $j \neq j'$, so G is d-splittable. $\qquad\square$

Combining Lemmas 5.2 and 5.3, if G is neither d-splittable nor d-pair-complete (if $p = 2$) then there is no divisibility barrier to a perfect matching in J_p. We saw in Lemma 4.2 that there is also no space barrier to a perfect matching in J_p. So Theorem 2.3 implies that G contains a near-balanced perfect matching. The following corollary formalises this argument.

Corollary 5.7. *Suppose that $1/n \ll \gamma \ll \alpha \ll d \ll 1/r$ and $2 \leq p \leq r$. Let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$ each of size pn with $\delta^*(G) \geq (p-1)n - \alpha n$. Suppose also that G is neither d-splittable nor d-pair-complete. Let $J = J(G)$ be the clique p-complex of G. Then J_p contains a γ-balanced perfect matching.*

Proof. Introduce new constants with $1/n \ll \gamma \ll \alpha \ll \mu, \beta \ll d \ll 1/r$. As described in Section 2, the condition $\delta^*(G) \geq (p-1)n - \alpha n$ implies

$$\delta^*(J) \geq (pn, (p-1)n - \alpha n, (p-2)n - 2\alpha n, \ldots, n - (p-1)\alpha n).$$

Suppose that J_p has no γ-balanced perfect matching. Then by Theorem 2.3 (with pn in place of n and p in place of k) we deduce that there is either a space barrier or divisibility barrier. Consider first a space barrier. This means that there exist $p' \in [p-1]$ and $S \subseteq V$ with $|S \cap V_i| = p'n$ for each $i \in [r]$ so that J_p is β-contained in $J_r(S, p')_p$, that is, all but at most $\beta(rpn)^p$ edges of J_p have at most p' vertices in S. However, since G is not d-splittable, Lemma 4.2(i) (with $2\beta(rp)^p$ in place of β) implies that more than $\beta(rp)^p n^p$ copies of K_p in G have more than p' vertices in $G[S]$. Since each copy of K_p in G is an edge of J_p, there cannot be a space barrier.

Now suppose that there is a divisibility barrier. This means that there is some partition $\mathcal{P}$ of $V(J)$ into parts of size at least $\delta^*_{p-1}(J) - \mu pn \geq n - 2p\mu n$ such that $\mathcal{P}$ refines the partition $\mathcal{P}'$ of $V(G)$ into $V_1, \ldots, V_r$ and $L^\mu_{\mathcal{P}}(J_p)$ is incomplete with respect to $\mathcal{P}'$. But if $p \geq 3$ then Lemma 5.3 (with $2p\mu$ in place of μ) implies that G is d-splittable, contradicting our assumption. Similarly, if $p = 2$ then Lemma 5.2 (with 4μ in place of μ) implies that G is d-splittable or d-pair-complete, again contradicting our assumption. We conclude that J_p must contain a γ-balanced perfect matching. $\qquad\square$

6. Finding packings within rows

Recall from the proof outline given in Section 3 that step (ii) in proving Theorem 1.1 is to find a balanced perfect K_{p_i}-packing in each row $G[X^i]$. In this section we demonstrate how this may be achieved. We need to consider two cases. The first case is where $G[X^i]$ is neither d-splittable nor d-pair-complete. Then Corollary 5.7 gives a γ-balanced perfect K_{p_i}-packing in $G[X^i]$. In Lemma 6.2 we show how such a matching can be 'corrected' to a balanced perfect K_{p_i}-packing in this case if $p_i \geq 3$, and also if $p_i = 2$ provided that $G[X^i]$ contains many 4-cycles of a given type. If $p_i = 2$ and $G[X^i]$ does not contain such 4-cycles then it may not be possible to find a balanced perfect matching in $G[X^i]$. Proposition 6.5 will allow us to handle this case by deleting further copies of K_k from G so that the remainder of row i does contain a balanced perfect matching. The second case is where $G[X^i]$ is d-pair-complete. Then we prove Lemma 6.4, which shows that $G[X^i]$ contains a perfect matching provided a parity condition is satisfied. Both here and later we use the fact that, if we add or remove a small number of vertices to or from each block X^i_j of a row X^i of G which is neither d-splittable nor d-pair-complete, then the new row obtained is neither d'-splittable nor d'-pair-complete for $d' \ll d$. This is established by the following proposition.

Proposition 6.1. *Suppose that $1/n, 1/n' \ll \zeta \ll d' \ll d \ll 1/r$ and $r \geq p \geq 1$. Let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$, and for each $j \in [r]$ let $X_j, X'_j \subseteq V_j$ be such that $|X_j| = pn$, $|X'_j| = pn'$ and $|X_j \triangle X'_j| \leq \zeta pn$. Let $X = \bigcup_{j \in [r]} X_j$ and $X' = \bigcup_{j \in [r]} X'_j$; then the following statements hold.*

(i) If $G[X]$ is not d-splittable then $G[X']$ is not d'-splittable.
(ii) If $p = 2$ and $G[X]$ is not d-pair-complete then $G[X']$ is not d'-pair-complete.

Proof. Note that $n' = (1 \pm \zeta)n$. For (i), suppose for a contradiction that $G[X']$ is d'-splittable. Then by definition we may choose $p' \in [p-1]$ and subsets $S'_j \subseteq X'_j$ with $|S'_j| = p'n'$ for $j \in [r]$ such that $e(S'_j, X'_{j'} \setminus S'_{j'}) \geq (1-d')p'(p-p')n'^2$ for any $j' \neq j$. For each $j \in [r]$ we choose $S_j \subseteq X_j$ such that $|S_j| = p'n$ and S_j either contains or is contained in $S'_j \cap X_j$. Note that $|S'_j \triangle S_j| \leq 2\zeta pn$ and $|(X'_j \setminus S'_j) \triangle (X_j \setminus S_j)| \leq 2\zeta pn$. We deduce that for any $j' \neq j$ we have

$$e(S_j, X_{j'} \setminus S_{j'}) \geq e(S'_j, X'_{j'} \setminus S'_{j'}) - |S'_j \setminus S_j|pn - |(X'_{j'} \setminus S'_{j'}) \setminus (X_{j'} \setminus S_{j'})|pn$$
$$\geq (1-d')p'(p-p')n'^2 - 4\zeta p^2 n^2 \geq (1-d)p'(p-p')n^2.$$

Then $G[X]$ is d-splittable with respect to the sets S_j, a contradiction, so this proves (i). For (ii), suppose for a contradiction that $G[X']$ is d'-pair-complete. Then by definition we may choose subsets $S'_j \subseteq X'_j$ with $|S'_j| = n'$ for $j \in [r]$ such that $e(S'_j, S'_{j'}) \geq (1-d')n'^2$, $e(X'_j \setminus S'_j, X'_{j'} \setminus S'_{j'}) \geq (1-d')n'^2$ and $e(S'_j, X'_{j'} \setminus S'_{j'}) \leq d'n'^2$ for any $j' \neq j$. We take $p' = 1$ and

choose S_j for $j \in [r]$ as in (i). Similar calculations as in (i) show that $e(S_j, S_{j'}) \geq (1-d)n^2$ and $e(X_j \setminus S_j, X_{j'} \setminus S_{j'}) \geq (1-d)n^2$ for any $j' \neq j$. We deduce that

$$e(S_j, X_{j'} \setminus S_{j'}) \leq e(S'_j, X'_{j'} \setminus S'_{j'}) + |S_j \setminus S'_j| \cdot 2n + |(X_{j'} \setminus S_{j'}) \setminus (X'_{j'} \setminus S'_{j'})| \cdot 2n$$
$$\leq d'n'^2 + 16\zeta n^2 \leq dn^2.$$

Then $G[X]$ is d-pair-complete with respect to the sets S_j, another contradiction, so this proves (ii). $\qquad\square$

We can now prove the main lemma of this section, which allows us to find balanced perfect clique packings in graphs which are not d-splittable or d-pair-complete.

Lemma 6.2. *Suppose that $1/n \ll \alpha, \nu \ll d \ll 1/r$, $2 \leq p \leq r$ and $r! \mid n$. Let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$ each of size pn, and let J be the clique p-complex of G. Suppose that G contains a spanning subgraph G^* such that G^* is not d-splittable and $\delta^*(G^*) \geq (p-1)n - \alpha n$. If $p \geq 3$, then J_p contains a balanced perfect matching. If instead $p = 2$ then J_p contains a balanced perfect matching if*

(i) G^* *is not d-pair-complete, and*

(ii) *either $r < 4$ or for any distinct i_1, i_2, i_3 in $[r] \setminus \{1\}$ there are at least νn^4 4-cycles $x_1 x_{i_1} x_{i_2} x_{i_3}$ in G with $x_1 \in V_1$ and $x_{i_j} \in V_{i_j}$ for $j \in [3]$.*

Proof. Introduce new constants ε, γ and d' with $1/n \ll \varepsilon \ll \gamma \ll \alpha, \nu \ll d' \ll d \ll 1/r$. Let $\mathcal{I} := \binom{[r]}{p}$, so $\mathcal{I}$ is the family of possible indices of edges of J_p. For any perfect matching M in J_p and any index $A \in \mathcal{I}$, let $N_M(A)$ be the number of edges in M with index A. Since any vertex of J lies in precisely one edge of M, for any $i \in [r]$ we must have

$$\text{(2)} \qquad\qquad\qquad \sum_{A \in \mathcal{I} \,:\, i \in A} N_M(A) = pn.$$

Let $N := rn/\binom{r}{p} = pn/\binom{r-1}{p-1}$, and observe that N is an integer. Our goal is to find a perfect matching M in J_p with $N_M(A) = N$ for every $A \in \mathcal{I}$. To do this, we will apply Theorem 2.3 to find a perfect matching which is near-balanced, but first we need to put aside some configurations that can be used to correct the small differences in the number of edges of each index. This will be unnecessary if $p = r$ or $p = r - 1$, as then any perfect matching in J_p must be balanced. Indeed, for $p = r$ this is trivial, whilst for $p = r - 1$ we note that by (2), for any $i \in [r]$, the number of edges of any perfect matching which do not contain a vertex of V_i is $rn - pn = n$. So for the purpose of finding configurations we may suppose that $p \leq r - 2$ (this is why we only require (ii) for $r \geq 4$).

Fix a set $S \subseteq [r]$ with $|S| = p - 2$ and an ordered quadruple $T = (a, a', b, b')$ of distinct members of $[r] \setminus S$. An (S,T)-*configuration* consists of two vertex-disjoint copies K and K' of K_{p-1}, where K has index $S \cup \{b\}$ and K' has index $S \cup \{b'\}$, and vertices $v \in V_a$ and $v' \in V_{a'}$ such that v and v' are both adjacent to every vertex of $K \cup K'$. Given such an (S,T)-configuration, we can select two vertex-disjoint copies of K_p in G (that is, two disjoint edges of J_p) in two different ways. One way is to take $K \cup \{v\}$ of index $S \cup \{a, b\}$ and $K' \cup \{v'\}$ of index $S \cup \{a', b'\}$; we call this the *unflipped* state. The other way is to take $K \cup \{v'\}$ of index $S \cup \{a', b\}$ and $K' \cup \{v\}$ of index $S \cup \{a, b'\}$; we call this the *flipped* state. Let $\mathcal{W}$ be the set of all pairs (S, T) as above. The first step in our proof is to find a collection $\mathcal{C}$ of pairwise vertex-disjoint configurations in G which contains $p\gamma n$ (S,T)-configurations in G for each $(S, T) \in \mathcal{W}$.

Suppose first that $p \geq 3$. To choose an (S, T)-configuration in this case we first fix $c \in S$ and find vertices $v \in V_a$ and $v' \in V_{a'}$ with $|N(v) \cap N(v') \cap V_c| > (p - 2 + 1/2p)n$. To see that this is possible, let $\mathcal{T}$ be the set of ordered triples (v, v', w) with $v \in V_a$, $v' \in V_{a'}$, $w \in V_c$ and $vw, v'w \in G$. For each w there are at least $\delta^*(G) \geq \delta^*(G^*)$ choices for each of v and v', so $|\mathcal{T}| \geq pn((p-1)n - \alpha n)^2$. Let P be the set of ordered pairs (v, v') that belong to at least $(p - 2 + 1/2p)n$ triples of $\mathcal{T}$, i.e. have $|N(v) \cap N(v') \cap V_c| > (p - 2 + 1/2p)n$. Then $|\mathcal{T}| \leq |P|pn + (pn)^2(p - 2 + 1/2p)n$, so $|P| \geq ((p-1)n - \alpha n)^2 - (p - 2 + 1/2p)pn^2 > n^2/3$. Given such a pair (v, v'), we choose the remaining vertices of the configuration greedily, ending with the two vertices in V_c. For each vertex not in V_c, the number of choices is at least $pn - (p-1)(n - \delta^*(G)) > n/2$. For the two vertices in V_c, the number of choices for each is at least $|N(v) \cap N(v') \cap V_c| - (p-2)(n - \delta^*(G)) > n/3p$. Now we choose the collection $\mathcal{C}$ greedily. At each step, the configurations chosen so far cover at most $2p \cdot |\mathcal{W}| \cdot p\gamma n$ vertices. Since there are at least $n^2/3$ choices for the pair (v, v') and at least $n/3p$ choices for any other vertex we are always able to choose an (S, T)-configuration which is vertex-disjoint from any configuration chosen so far, as required.

Now consider instead the case $p = 2$, for which $J_p = G$. Recall that we can assume $r \geq 4$. Now an (S, T)-configuration consists of a 4-cycle $xyzw$ with $x \in V_a, y \in V_b, z \in V_{a'}$ and $w \in V_{b'}$, where $T = (a, a', b, b')$ (we have $S = \emptyset$). Note that the unflipped state of such a configuration has edges xy and zw, and the flipped state has edges xw and yz. If $a = 1$ then by assumption there are at least νn^4 such 4-cycles in G. For $a \neq 1$ we instead choose *fake configurations*; for $S = \emptyset$ and $T = (a, a', b, b')$ a fake (S, T)-configuration consists simply of vertices $xyzw$ with $x \in V_a, y \in V_b, z \in V_{a'}$ and $w \in V_{b'}$ such that xy and zw are edges. A fake configuration should be thought of as a configuration which cannot be flipped. There are at least $(2n\delta^*(G))^2 \geq n^4$ fake configurations for each T, so similarly to before we may choose $\mathcal{C}$ greedily to consist of genuine (S, T)-configurations if $a = 1$, and fake (S, T)-configurations otherwise. Indeed, at any step the configurations chosen so far cover at most $8|\mathcal{W}|\gamma n$ vertices, so for any $(S, T) \in \mathcal{W}$ at most $8|\mathcal{W}|\gamma n \cdot (2n)^3 < \nu n^4$ (S, T)-configurations share a vertex with a previously-chosen configuration.

Let $V' = \bigcup_{i \in [r]} V_i'$ be the set of all vertices not covered by configurations in $\mathcal{C}$. We now find a matching in J_p covering V'. Note that the configurations in $\mathcal{C}$ cover $2p^2|\mathcal{W}|\gamma n$ vertices in total, equally many of which lie in each vertex class, so for any $i \in [r]$ we have $|V_i'| = pn'$, where $n' := n - 2p^2|\mathcal{W}|\gamma n/r$. Let $G' = G^*[V']$ and let J' be the clique p-complex of G'. Then $\delta^*(G') \geq \delta^*(G^*) - \alpha n \geq (p-1)n' - 2\alpha n$. Furthermore, by Proposition 6.1 G' is not d'-splittable, and if $p = 2$ then G' is not d'-pair-complete. So we may apply Corollary 5.7 with $d', n', 3\alpha$ and ε in place of d, n, α and γ respectively to obtain that J'_p must contain an ε-balanced perfect matching. Extend this matching to a perfect matching M^0 in J_p by adding the configurations in $\mathcal{C}$, all in their unflipped state. This adds equally many edges of each index, so M^0 is ε-balanced, and so $N_{M^0}(A) = (1 \pm \varepsilon)N$ for any index $A \in \mathcal{I}$.

It remains only to flip some configurations to correct these small imbalances in the number of edges of each index. To accomplish this, we shall proceed through the index sets $A \in \mathcal{I}$ in order. For each A we will flip some configurations to obtain a perfect matching with precisely N edges of each index set A' considered in this order up to and including A. At the end of this process we will obtain a perfect matching with precisely N edges of every index set. Let $\mathcal{A}_1 \subseteq \mathcal{I}$ consist of all members of $\mathcal{I}$ of the form $[p-1] \cup \{i\}$ for some $p + 2 \leq i \leq r$, and $\mathcal{A}_2 \subseteq \mathcal{I}$ consist of all members of $\mathcal{I}$ of the form $[p+1] \setminus \{i\}$ for some $i \in [p+1]$. Let $\mathcal{A} = \mathcal{A}_1 \cup \mathcal{A}_2$ and $m = |\mathcal{I}| = \binom{r}{p}$, so $|\mathcal{A}| = r$. Note that if $p = 2$ then $\mathcal{A}$ contains all pairs

$\{1, j\}$ with $2 \leq j \leq r$. Choose any linear ordering $A_1 \leq A_2 \leq \cdots \leq A_m$ of the elements of $\mathcal{I}$ such that

(i) for any $A \in \mathcal{I}$, $x \in A$ and $y \notin A$ with $y < x$ we have $\{y\} \cup A \setminus \{x\} > A$, and
(ii) $\mathcal{A}$ is a terminal segment of $\mathcal{I}$.

This is possible since for any $A \in \mathcal{A}$, $x \in A$ and $y \notin A$ with $y < x$ we have $\{y\} \cup A \setminus \{x\} \in \mathcal{A}$. Note that for any $A \in \mathcal{I} \setminus \mathcal{A}$ there exist $x, y \in A$ and $x', y' \notin A$ such that $x' < x$ and $y' < y$. Crucially, the sets $\{x'\} \cup A \setminus \{x\}$, $\{y'\} \cup A \setminus \{y\}$ and $\{x', y'\} \cup A \setminus \{x, y\}$ each appear after A, by choice of the ordering.

We now proceed through the sets A_i, $i \in [m - r]$ in turn (these are precisely the sets A_i with $A_i \notin \mathcal{A}$). At each step i we will flip at most $2^{i-1}\varepsilon N$ configurations to obtain a perfect matching M^i such that $N_{M^i}(A_j) = N$ for any $j \leq i$ and $|N_{M^i}(A_j) - N| \leq 2^i \varepsilon N$ for any $j > i$. The matching M^0 satisfies these requirements for $i = 0$, so suppose that we have achieved this for $A_1, \ldots, A_{i-1}$, and that we wish to obtain M^i from M^{i-1}. If $N_{M^{i-1}}(A_i) = N$ then we may simply take $M^i = M^{i-1}$, so we may assume $N_{M^{i-1}}(A_i) \neq N$. First suppose that $N_{M^{i-1}}(A_i) > N$. Since $A_i \notin \mathcal{A}$ we may choose $x, y \in A_i$ and $x', y' \notin A_i$ such that $x' < x$ and $y < y'$. Furthermore, if $p = 2$ then we may also require that $x' = 1$. We let $T = (x', x, y', y)$, $S = A_i \setminus \{x, y\}$, and choose a set of $N_{M^{i-1}}(A_i) - N$ unflipped (S, T)-configurations in $\mathcal{C}$ (we shall see later that this is possible). We flip all of these configurations (this is possible if $p = 2$ since $x' = 1$, so these configurations are not fake-configurations). In doing so, we replace $N_{M^{i-1}}(A_i) - N$ edges of M^{i-1} of index $A_i = S \cup \{x, y\}$ and $N_{M^{i-1}}(A_i) - N$ edges of M^{i-1} of index $S \cup \{x', y'\}$ with $N_{M^{i-1}}(A_i) - N$ edges of M^{i-1} of index $S \cup \{x', y\}$ and $N_{M^{i-1}}(A_i) - N$ edges of M^{i-1} of index $S \cup \{x, y'\}$. The number of edges of each other index remains the same; note that this includes any index A_j with $j < i$. Let M^i be formed from M^{i-1} by these flips; then $N_{M^i}(A_j) = N$ for any $j \leq i$ by construction. Also, for any $j > i$ we have

$$|N_{M^i}(A_j) - N| \leq |N_{M^{i-1}}(A_j) - N| + |N_{M^{i-1}}(A_i) - N| \leq 2 \cdot 2^{i-1}\varepsilon N = 2^i \varepsilon N,$$

as required. On the other hand, for $N_{M^{i-1}}(A_i) < N$ we obtain M^i similarly by the same argument with $T = (x', x, y, y')$. To see that it is always possible to choose a set of $N_{M^{i-1}}(A_i) - N$ unflipped (S, T)-configurations from $\mathcal{C}$, note that there are $m - r = \binom{p}{r} - r$ steps of the process, and that at step i exactly $|N_{M^{i-1}}(A_i) - N| \leq 2^{i-1}\varepsilon N$ members of $\mathcal{C}$ are flipped. So in total at most $2^{m-r}\varepsilon N \leq p\gamma n$ members of $\mathcal{C}$ are flipped. Since $\mathcal{C}$ was chosen to contain at least this many pairwise vertex-disjoint (S, T)-configurations for any $(S, T) \in \mathcal{W}$, it will always be possible to make these choices.

At the end of this process, we obtain a perfect matching $M := M^{m-r}$ in J_p such that $N_M(A) = N$ for every $A \notin \mathcal{A}$. It remains only to show that we also have $N_M(A) = N$ for any $A \in \mathcal{A}$. We first consider $A \in \mathcal{A}_1$, so $A = [p-1] \cup \{i\}$ for some $i \geq p + 2$, and A is the only index in $\mathcal{A}$ which contains i. Each of the $\binom{r-1}{p-1} - 1$ other sets A' containing i has $N_M(A') = N = pn/\binom{r-1}{p-1}$, so by (2) we also have $N_M(A) = N$. Thus $N_M(A) = N$ for any $A \in \mathcal{A}_1$. Now consider $A \in \mathcal{A}_2$, so $A = [p+1] \setminus \{i\}$ for some $i \in [p+1]$. Note that A is the only member of $\mathcal{A}_2$ which does not contain i. Since $N_M(A') = N$ for any $i \in A' \notin \mathcal{A}_2$, by (2) we have $\sum_{A' \in \mathcal{A}_2 \setminus \{A\}} N_M(A') = pN$. This holds for all $A \in \mathcal{A}_2$, so $N_M(A) = N$ for all $A \in \mathcal{A}_2$. Thus M is a balanced perfect matching in J_p, as required. $\qquad\square$

We also need to be able to find balanced perfect matchings in pair-complete rows. Recall that an r-partite graph G with vertex classes $V_1, \ldots, V_r$ of size $2n$ is *d-pair-complete* if there

exist sets $S_j \subseteq V_j$, $j \in [r]$ each of size n such that $d(S_i, S_j) \geq 1 - d$, $d(V_i \setminus S_i, V_j \setminus S_j) \geq 1 - d$ and $d(S_i, V_j \setminus S_j) \leq d$ for any $i, j \in [r]$ with $i \neq j$. This implies that almost all vertices in S_i have few non-neighbours in S_j and almost all vertices in $V_i \setminus S_i$ have few non-neighbours in $V_j \setminus S_j$. Lemma 6.4 will show that G contains a balanced perfect matching under a similar condition, namely that there are sets $X_j \subseteq V_j$, $j \in [r]$ each of size approximately n such that all vertices in X_i have few non-neighbours in X_j and all vertices in $V_i \setminus X_i$ have few non-neighbours in $V_j \setminus X_j$ for any $i \neq j$, provided that $X = \bigcup_{i \in [r]} X_i$ has even size. Note that we cannot omit the parity requirement on $|X|$, as there may be no edges between X and $V(G) \setminus X$. The proof uses the following characterisation of multigraphic degree sequences by Hakimi [5]. A sequence $d = (d_1, \ldots, d_n)$ with $d_1 \geq \cdots \geq d_n$ is *multigraphic* if there is a loopless multigraph on n vertices with degree sequence d.

Proposition 6.3. *([5]) A sequence $d = (d_1, \ldots, d_n)$ with $d_1 \geq \cdots \geq d_n$ is multigraphic if and only if $\sum_{i=1}^{n} d_i$ is even and $d_1 \leq \sum_{i=2}^{n} d_i$.*

Lemma 6.4. *Suppose that $1/n \ll \zeta \ll 1/r$ and $r - 1 \mid 2n$. Let G be an r-partite graph on vertex classes $V_1, \ldots, V_r$ each of size $2n$. For each $i \in [r]$ let sets X_i and Y_i partition V_i and be such that*

(i) $|X_j|, |Y_j| = (1 \pm \zeta)n$ for any $j \in [r]$,
(ii) for any $i \neq j$, any $x \in X_i$ has at most ζn non-neighbours in X_j and any $y \in Y_i$ has at most ζn non-neighbours in Y_j, and
(iii) $X := \bigcup_{i \in [r]} X_i$ has even size.
Then G contains a balanced perfect matching.

Proof. Choose an integer n' so that $(1 - 5\zeta)n \leq n' \leq (1 - 4\zeta)n$ and $r - 1 \mid n'$. For each $j \in [r]$ let $a_j := |X_j| - n'$. So $3\zeta n \leq a_j \leq 6\zeta n$ for each j, and $a := \sum_{j \in [r]} a_j = |X| - rn'$ is even by (iii) and since rn' is divisible by $r(r - 1)$, which is even. By Proposition 6.3, we can choose pairs (i_ℓ, j_ℓ) with $i_\ell \neq j_\ell$ for $\ell \in [a/2]$ so that each $j \in [r]$ appears in precisely a_j pairs. For each $\ell \in [a/2]$ choose a matching M_ℓ in G which contains an edge of $G[X]$ of index $\{i_\ell, j_\ell\}$, and an edge of $G[Y]$ of each index $A \in \binom{[r]}{2} \setminus \{\{i_\ell, j_\ell\}\}$. We also require that these matchings are pairwise vertex-disjoint. Such matchings may be chosen greedily using (ii), since together they will cover a total of $2 \cdot a/2 \cdot \binom{r}{2} \leq 3\zeta r^3 n \leq n/2$ vertices. Let $M = \bigcup_{\ell \in [a/2]} M_\ell$. Note that M has $a/2$ edges of each index, and so covers $(r - 1)a/2$ vertices from each V_j. For each $j \in [r]$ let $X_j' = X_j \setminus V(M)$ and $Y_j' = Y_j \setminus V(M)$, and let $X' = X \setminus V(M)$ and $Y' = Y \setminus V(M)$. Then $|X_j'| = |X_j| - a_j = n'$ and $|Y_j'| = 2n - |X_j'| - (r - 1)a/2$ for any $j \in [r]$. Since $r - 1$ divides n' and $2n$, we conclude that $r - 1$ divides $|X_j'|$ and $|Y_j'|$ for any $j \in [r]$. So we may partition X' and Y' into sets X_A' and Y_A' for each $A \in \binom{[r]}{2}$, where for each $A = \{i, j\}$, X_A' consists of $n_X = n'/(r - 1)$ vertices from each of X_i' and X_j', and Y_A' consists of $n_Y = |Y_1'|/(r - 1)$ vertices from each of Y_i' and Y_j'. Now for any $A \in \binom{[r]}{2}$, the induced bipartite graph $G[X_A']$ has minimum degree at least $n_X - \zeta n \geq n_X/2$ by (ii), and so contains a perfect matching M_A of size n_X; by the same argument $G[Y_A']$ contains a perfect matching M_A' of size n_Y. Finally, $M \cup \bigcup_A (M_A \cup M_A')$ is a perfect matching in G with $a/2 + n_X + n_Y$ edges of each index, as required. $\square$

However, there are some r-partite graphs G on vertex classes $V_1, \ldots, V_r$ of size $2n$ which satisfy $\delta^*(G) \geq n - \alpha n$ and are neither d-splittable nor d-pair-complete but do not contain a balanced perfect matching. For example, let H be a graph with vertex set $\{x_i, y_i : i \in [r]\}$,

where $x_i x_j$ and $y_i y_j$ are edges for any $i \neq j$ except when $i = 1, j = 2$, and $x_1 y_2$ and $x_2 y_1$ are edges. Form G by 'blowing up' H, that is, replace each x_i and y_i by a set of n vertices, where $r - 1 \mid 2n$, and each edge by a complete bipartite graph between the corresponding sets. Then G contains $2rn$ vertices, and so any balanced perfect matching M in G contains $2n/(r-1)$ edges of each index. Let X be the vertices of G which correspond to some vertex x_i of H. Then $|X| = rn$ and any edge of M covering a vertex of X either covers two vertices of X or has index $\{1, 2\}$. There are exactly $2n/(r-1)$ edges in M of the latter form, so we must have $rn = |X| \equiv 2n/(r-1)$ modulo 2. We conclude that G contains no balanced perfect matching if this congruence fails (e.g. if $r = 5$ and n is even but not divisible by 4). However, $\delta^*(G) = n$, and it is easily checked that G is neither d-splittable nor d-pair-complete if $d \ll 1$. Other examples can be obtained similarly. Note that G does contain a perfect matching by Corollary 5.7, but this cannot be balanced. In later arguments, such graphs will only cause difficulties when the row-decomposition has one row of this form (with $p_i = 2$), and every other row has $p_i = 1$. In such cases the following proposition will enable us to delete further copies of K_k so that the subgraph remaining has a balanced perfect K_{p_i}-packing in every row i.

Proposition 6.5. *Suppose that $1/n \ll \gamma, \alpha \ll 1/r$, $r!$ divides n and $r \geq k \geq 2$. Let G be an r-partite graph on vertex classes $X_1, \ldots, X_r$ of size kn which admits a $(k-1)$-row-decomposition into pairwise-disjoint blocks X_j^i with $|X_j^i| = p_i n$ for $i \in [s]$ and $j \in [r]$, where $p_1 = 2$ and $p_2, \ldots, p_{k-1} = 1$. Suppose that $G[X^1]$ contains a γ-balanced perfect matching M', and that for any $i \neq i'$ and $j \neq j'$ any vertex $v \in X_j^i$ has at most αn non-neighbours in $X_{j'}^{i'}$. Then there exists an integer $D \leq 2\gamma n$ and a K_k-packing M in G such that $r!$ divides $n - D$, M covers $p_i D$ vertices in X_j^i for any $i \in [s]$ and $j \in [r]$, and $G[X^1 \setminus V(M)]$ contains a balanced perfect matching.*

Proof. Note that $|M'| = rn$, and since M' is γ-balanced, $|N_A(M') - N_B(M')| \leq \gamma rn/\binom{r}{2}$ for any $A, B \in \binom{[r]}{2}$. So we may write $M' = M_0 \cup M_1$, where M_0 is a balanced matching in $G[X^1]$, $r \cdot r!$ divides $|M_0|$, and $|M_1| \leq 2\gamma rn$. Note that M_0 covers $2|M_0|/r$ vertices in each X_j^1, $j \in [r]$, so M_1 covers the remaining $2D$ vertices in each X_j^1, where $D := n - |M_0|/r$. Note also that $|M_1| = Dr$, $D \leq 2\gamma n$, and $r!$ divides $|M_0|/r = n - D$. We will construct a sequence $M_1, \ldots, M_{k-1}$, where M_i is a K_{i+1}-packing in $G[\bigcup_{i' \in [i]} X^{i'}]$ that covers $2D$ vertices in X_j^1 and D vertices in $X_j^{i'}$ for each $2 \leq i' \leq i$ and $j \in [r]$. Each M_i will have size $|M_i| = Dr$, and will be formed by adding a vertex of X^i to each copy of K_i in M_{i-1}.

Suppose that we have formed M_i in this manner for some $i \geq 1$, and now we wish to form M_{i+1}. Let Z be the set of ordered pairs (j, q) with $j \in [r]$ and $q \in [D]$. We form a bipartite graph B whose vertex classes are M_i and Z, where a copy K' of K_{i+1} in M_i and a pair (j, q) are connected if $j \notin i(K')$, that is, if K' contains no vertex from X_j. Note that any K' in M_i has degree $(r - i - 1)D$ in B, as there are $r - i - 1$ choices for $j \in [r] \setminus i(K')$ and D choices for $q \in [D]$. The same is true of any pair (j, q) in Z, as $|M_i| = Dr$, and $(i+1)D$ cliques in M_i intersect X_j. So B is a regular bipartite graph, and therefore contains a perfect matching. Let $f : M_i \to Z$ be such that $\{K'f(K') : K' \in M_i\}$ is a perfect matching in B. For each $K' \in M_i$ we extend K' to a copy of K_{i+2} in G by adding a vertex from X_j^{i+1}, where $f(K') = (j, q)$ for some $q \in [D]$. Such extensions may be chosen distinctly, since any $K' \in M_i$ has at least $n - |K'|\alpha n \geq n/2 \geq D$ possible extensions to X_j^{i+1} for any $j \notin i(K')$.

Since every pair in Z was matched to some $K' \in M_i$, the K_{i+2}-packing M_{i+1} covers precisely D vertices from X_j^{i+1} for each $j \in [r]$.

At the end of this process we obtain a K_k-packing $M := M_{k-1}$ in G which covers $2D$ vertices from X_j^1 and D vertices from X_j^i for $i \geq 2$ and $j \in [r]$. By construction M_0 is a balanced perfect K_{p_i}-packing in $G[X^1 \setminus V(M)]$ and $r!$ divides $|M_0|/r = n - D$, as required. $\qquad\square$

7. Covering bad vertices

The final ingredient of the proof is a method to cover 'bad' vertices. The row-decomposition of G will have high minimum diagonal density, which implies that most vertices have high minimum diagonal degree. However, we need to remove those 'bad' vertices which do not have high minimum diagonal degree, so that we can accomplish step (iii) of the proof outline in Section 3, which is to glue together the perfect clique packings in each row to form a perfect K_k-packing in G. Lemma 7.2 will show that we can cover the bad vertices of G by vertex-disjoint copies of K_k, whilst keeping the block sizes balanced and fixing the parity of any pair-complete row, so that each row contains a clique packing covering all of the undeleted vertices. First we need some standard definitions. An $s \times r$ *rectangle* R is a table of rs cells arranged in s rows and r columns. We always assume that $s \leq r$. A *transversal* T in R is a collection of s cells of the grid so that no two cells of T lie in the same row or column. We need the following simple proposition.

Proposition 7.1. *Suppose that R is an $s \times r$ rectangle, where $r \geq s \geq 0$ and $r \geq 1$. Suppose that at most r cells of R are coloured, such that at most one cell is coloured in each column, and at most $r - 1$ cells are coloured in each row. Then R contains a transversal of non-coloured cells.*

Proof. We proceed by induction on s. Note that the proposition is trivial for the cases $s = 0, 1$ and $r = s = 2$. Now assume that $r \geq 3$ and $s \geq 2$. Choose a row with the most coloured cells, and select a non-coloured cell in this row (this is possible since at most $r - 1$ of the r cells in this row are coloured). Let R' be the $(s-1) \times (r-1)$ subrectangle obtained by removing the row and column containing this cell. Then it suffices to find a transversal of non-coloured cells in R'. Note that R' has at most $r - 1$ coloured cells, as if R had any coloured cells, then we deleted at least one coloured cell. Also, since R had at most $r < 2(r - 1)$ coloured cells, R' must have at most $r - 2$ cells coloured in any row, since we removed the row containing the most coloured cells. Since R' contains at most one coloured cell in any column, the required transversal in R' exists by the induction hypothesis. $\qquad\square$

The remainder of this section is occupied by the proof of Lemma 7.2. This is long and technical, so to assist the reader we first give an overview. We start by imposing structure on the graph G, fixing a row-decomposition with high minimum diagonal density. Let X_j^i, $i \in [s]$, $j \in [r]$ be the blocks of this row-decomposition. Then we identify vertices that are 'bad' for one of two reasons: (i) not having high minimum diagonal degree, or (ii) belonging to a pair-complete row and not having high minimum degree within its own half. We assign each bad vertex v to a row in which it has as many blocks as possible that are 'bad' for v, in that they have many non-neighbours of v, so that there are as few bad blocks for v as possible in the other rows. We also refer to the resulting sets W_j^i, $i \in [s]$, $j \in [r]$ as blocks, as although they may not form a row-decomposition, since few vertices are moved

they retain many characteristics of a row-decomposition. If the row-decomposition has type $p = (p_i : i \in [s])$ then $|W^i|$ is approximately proportional to p_i for $i \in [s]$. We establish some properties of the W_j^i's in Claim 7.3. Next, we show how to find various types of K_k in G that will form the building blocks for the deletions. These will be 'properly distributed', in that they have p_i vertices (the 'correct number') in row i for $i \in [s]$, or 'ij-distributed' for some i, j, in that they have 'one too many' vertices in row i and 'one too few' vertices in row j. They will also have an even number of vertices in each half of pair-complete rows so as to preserve parity conditions, except that sometimes we require a clique that is 'properly-distributed outside of row ℓ' to fix the parity of a pair-complete row ℓ. Claim 7.4 analyses a general greedy algorithm for finding copies of K_k, and then Claim 7.5 deduces five specific corollaries on finding building blocks of the above types. In Claim 7.6 we use ij-distributed cliques to balance the row sizes, so that the remainder of row ℓ has size proportional to p_ℓ for $\ell \in [s]$. There are two cases according to whether G has the same row structure as the extremal example; if it does we also ensure in this step that the remainder of each half in pair-complete rows has even size. Next, in Claim 7.7 we put aside an extra K_k-packing that is needed to provide flexibility later in the case when there are at least two rows with $p_i \geq 2$. Then in Claim 7.8 we cover all remaining bad vertices and ensure that the number of remaining vertices is divisible by $rk \cdot r!$. Next, in Claim 7.9 we choose a K_k-packing so that equally many vertices are covered in each part V_j. Then in Claim 7.10 we choose a final K_k-packing so that the remaining blocks all have size proportional to their row size. After deleting all these K_k-packings we obtain G' with an s-row-decomposition X' that satisfies conclusion (i) of Lemma 7.2. To complete the proof, we need to satisfy conclusion (ii), by finding a balanced perfect K_{p_i}-packing in row i for each $i \in [s]$. We need to consider two cases according to whether or not there are multiple rows with $p_i \geq 2$; if there are, then we may need to make some alterations to the K_k-packing from Claim 7.7. Finally, we apply the results of the previous section to find the required balanced perfect clique packings.

Lemma 7.2. *Suppose that $1/n^+ \ll \alpha \ll 1/r$, $r \geq k \geq 3$ and $r > 3$. Let $\mathcal{P}'$ partition a set V into r parts $V_1, \ldots, V_r$ each of size n^+, where rn^+/k is an integer. Suppose that G is a $\mathcal{P}'$-partite graph on V with $\delta^*(G) \geq (k-1)n^+/k$. Suppose also that if rn^+/k is odd and k divides n^+ then G is not isomorphic to the graph $\Gamma_{n^+, r, k}$ of Construction 1.2. We delete the vertices of a collection of pairwise vertex-disjoint copies of K_k from G to obtain $V_1', \ldots, V_r'$, $V' := \bigcup_{j \in [r]} V_j'$ and $G' = G[V']$, such that $|V_j'| = kn'$ for $j \in [r]$, where $r! \mid n'$ and $n' \geq n^+/k - \alpha n$. We can perform this deletion so that G' has an s-row-decomposition X', with blocks $X_j'^i$ for $i \in [s]$, $j \in [r]$ of size $p_i n'$, where $p_i \in [k]$ with $\sum_{i \in [s]} p_i = k$, with the following properties:*

(i) *For each $i, i' \in [s]$ with $i \neq i'$ and $j, j' \in [r]$ with $j \neq j'$, any vertex $v \in X_j'^i$ has at least $p_{i'} n' - \alpha n'$ neighbours in $X_{j'}'^{i'}$.*

(ii) *For every $i \in [s]$ the row $G[X'^i]$ contains a balanced perfect K_{p_i}-packing.*

Proof. Introduce new constants with $1/n \ll d_0 \ll d_1 \ll \cdots \ll d_k \ll \alpha$. Let n be the integer such that $n^+ - k + 1 \leq kn \leq n^+$. Note that $kn - \delta^*(G^+) \leq \lfloor n^+/k \rfloor = n$, so

$$\text{any vertex has at most } n \text{ non-neighbours in any other part.} \qquad (\dagger)$$

Let X be formed from V by arbitrarily deleting $n^+ - kn$ vertices in each part. We fix an s-row-decomposition of $G_1 := G[X]$. Recall that this consists of pairwise disjoint blocks $X_j^i \subseteq V_j$ with $|X_j^i| = p_i n$ for each $i \in [s]$, $j \in [r]$, for some $s \in [k]$ and positive integers

p_i, $i \in [s]$ with $\sum_{i \in [s]} p_i = k$. By Proposition 4.1 we may fix this row-decomposition to have minimum diagonal density at least $1 - k^2 d_{s-1}$, and such that each row $G_1[X^i]$ is not d_s-splittable. Having fixed s and the row-decomposition of G_1, introduce new constants with $d_{s-1} \ll d'_0 \ll d'_1 \ll \cdots \ll d'_{2s+2} \ll d_s$. For each $i \in [s]$ with $p_i = 2$, let $d(i)$ denote the infimum of all d such that $G_1[X^i]$ is d-pair-complete. This gives us at most s values of $d(i)$, so we may choose $t \in [2s+2]$ such that there is no $i \in [s]$ with $d'_{t-2} < d(i) \le d'_t$. We let $d := d'_{t-1}$, $d' := d'_t$, and introduce further new constants $\nu, \eta, \beta, \beta', \zeta, \gamma, \gamma', d''$ and ω such that

$$1/n^+ \ll \nu \ll \eta \ll d \ll \beta \ll \beta' \ll \zeta \ll \gamma \ll \gamma' \ll d'' \ll d' \ll \omega, \alpha \ll 1/r \le 1/k.$$

These are the only constants which we shall use from this point onwards. The purpose of these manipulations is that our fixed row-decomposition has three important properties. Firstly, it has minimum diagonal density at least $1 - d$, since $d = d'_{t-1} \ge k^2 d_{s-1}$. Secondly, any row $G_1[X^i]$ is not d'-splittable, since $d' = d'_t \le d_s$. Thirdly, any row $G_1[X^i]$ which is d'-pair-complete is d-pair-complete, since $d(i) \le d' = d'_t$ by definition of $d(i)$, and so $d(i) \le d'_{t-2} < d'_{t-1} = d$ by choice of t.

Suppose first that $s = 1$, so G_1 has only one row X^1, and $p_1 = k$. Fix $n - r! \le n' \le n$ such that $r! \mid n'$, and let $C = n^+ - kn'$. Then $0 \le C \le kr! + k$. Also, since rn^+/k and $r(kn')/k$ are integers, rC/k is also an integer. We choose rC/k pairwise-disjoint copies of K_k in G which together cover C vertices in each V_j. To see that this is possible note that, by (†), for any $A \in \binom{[r]}{k}$ we may greedily choose the vertices of a copy of K_k in G of index A; this gives at least n choices for each vertex, of which at most $n/2$ (say) have been previously used, so some choice remains. Note that our use of (†) here is not tight, in the sense that the argument would still be valid if n was replaced by $n + o(n)$ in the statement of (†). This will be true of all our applications of (†) except for that in Claim 7.6. We delete all of these copies of K_k from G, and let G' be the resulting graph. We let X'^1_j consist of the kn' undeleted vertices of V_j for each $j \in [r]$. Then the sets X'^1_j for $j \in [r]$ form a 1-row-decomposition of G', which is not d''-splittable by Proposition 6.1. Since $p_1 = k \ge 3$, $G' = G'[X'^1]$ contains a balanced perfect K_k-packing by Lemma 6.2.

We may therefore assume that $s \ge 2$. From each block X^i_j we shall obtain a set W^i_j by moving a small number of 'bad' vertices between blocks, and reinstating the vertices deleted in forming X. As a consequence the sets W^i_j will not form a proper row-decomposition (for example, blocks in the same row may have different sizes). However, since only a small number of vertices will be moved or reinstated, the partition of $V(G)$ into sets W^i_j will retain many of the characteristics of the s-row-decomposition of G_1 into blocks X^i_j. We therefore keep the terminology, referring to the sets W^i_j as 'blocks', and the $W^i = \bigcup_j W^i_j$ and $W_j = \bigcup_i W^i_j$ as 'rows' and 'columns' respectively. Perhaps it is helpful to think of the sets W^i_j as being containers which correspond to the blocks X^i_j, between which vertices may be transferred. It is important to note, however, that the blocks X^i_j will remain unchanged throughout the proof. Furthermore, we shall sometimes refer to the row $G[W^i]$ simply as row i, but we say that a row i is *pair-complete* if $G_1[X^i]$ is d-pair-complete. This means that the truth of the statement 'row i is pair-complete' depends only on our fixed row-decomposition of G_1, and not on the 'blocks' W^i_j or their subsets defined later. Note that if $p_i = 2$ and row i is not pair-complete then $G[X^i]$ is not d'-pair-complete.

We start by identifying the *bad* vertices, which may be moved to a different block. For each $i \in [s]$ and $j \in [r]$ let B_j^i consist of all vertices $v \in X_j^i$ for which there exist $i' \neq i$ and $j' \neq j$ such that $|N(v) \cap X_{j'}^{i'}| \leq (1 - \sqrt{d})p_{i'}n$. We must have $|B_j^i| \leq rk\sqrt{d}p_i n$, otherwise for some $i' \neq i$ and $j' \neq j$ there are more than $\sqrt{d}p_i n$ vertices in X_j^i with at most $(1 - \sqrt{d})p_{i'}n$ neighbours in $X_{j'}^{i'}$. Then $d(X_j^i, X_{j'}^{i'}) < \sqrt{d} \cdot (1 - \sqrt{d}) + (1 - \sqrt{d}) \cdot 1 = 1 - d$ contradicts the minimum diagonal density of G_1.

Next, for each $i \in [s]$ for which row i is pair-complete, by definition there are sets $T_j^i \subseteq X_j^i$ of size n for each $j \in [r]$ such that $d(T_j^i, T_{j'}^i) \geq 1 - d$ and $d(X_j^i \setminus T_j^i, X_{j'}^i \setminus T_{j'}^i) \geq 1 - d$ for any $j \neq j'$. For each $j \in [r]$, we let $B_j'^i$ consist of all vertices $v \in T_j^i$ for which there exists $j' \neq j$ such that $|N(v) \cap T_{j'}^i| \leq (1 - \sqrt{d})n$, and also all vertices $v \in X_j^i \setminus T_j^i$ for which there exists $j' \neq j$ such that $|N(v) \cap (X_{j'}^i \setminus T_{j'}^i)| \leq (1 - \sqrt{d})n$. We must have $|B_j'^i| \leq 2r\sqrt{d}n$, otherwise (without loss of generality) there exists some $j' \neq j$ for which more than $\sqrt{d}n$ vertices in T_j^i have at most $(1-\sqrt{d})n$ neighbours in $T_{j'}^i$. Then $d(T_j^i, T_{j'}^i) < \sqrt{d} \cdot (1-\sqrt{d}) + (1 - \sqrt{d}) \cdot 1 = 1 - d$ contradicts the choice of the sets T_j^i. Thus we have bad sets B_j^i and $B_j'^i$ for $i \in [s]$, $j \in [r]$. We also consider the $n^+ - kn$ deleted vertices in each part to be bad. Let B be the set of all bad vertices. The remaining vertices Y are *good*; let $Y_j^i = X_j^i \setminus B$, $Y^i = \bigcup_{j \in [r]} Y_j^i$ and $Y_j = \bigcup_{i \in [s]} Y_j^i$ for each i and j, so $Y = \bigcup_{i \in [s]} Y^i$.

Let v be any vertex of G. We say that a block X_j^i is *bad with respect to* v if $|N(v) \cap X_j^i| < p_i n - n/2$, that is, if v has more than $n/2$ non-neighbours in X_j^i. So if v is a good vertex, then the set of blocks which are bad with respect to v is a subset of the set of blocks in the same row and column as v. Also, by (†) for any $v \in V(G)$ at most one block in each other column can be bad with respect to v. Similarly as with the notion of pair-completeness, this definition fixes permanently which blocks are bad with respect to a vertex v. We shall later sometimes refer to a 'block' W_j^i being bad with respect to v; this should always be taken to mean that X_j^i is bad which respect to v. We say that a block is *good with respect to* v if it is not bad with respect to v.

We now define the sets W_j^i for each $i \in [s]$ and $j \in [r]$ as follows. Any vertex in Y_j^i is assigned to W_j^i. It remains only to assign the bad vertices; each bad vertex $v \in V_j$ is assigned to W_j^i, where i is a row containing the most blocks X_j^i which are bad with respect to v (if more than one row has the most bad blocks then choose one of these rows arbitrarily). For each pair-complete row i, we also modify the sets T_j^i to form sets S_j^i. Indeed, S_j^i is defined to consist of all vertices of $T_j^i \cap Y_j^i$, plus any vertex in $W_j^i \setminus Y_j^i$ which has at least $n/2$ neighbours in $T_{j'}^i$ for some $j' \neq j$. We let $S^i := \bigcup_{j \in [r]} S_j^i$ for any such i. This completes the phase of the proof in which we impose structure on G. The next claim establishes some properties of the decomposition into 'blocks' W_j^i.

Claim 7.3. *(Structural properties)*

(A1) At most $\beta n/2$ vertices of G are bad.

(A2) We have $Y_j^i \subseteq W_j^i$ and $(p_i - \beta/2)n \leq |Y_j^i| \leq |W_j^i| \leq (p_i + \beta/2)n$ for any $i \in [s]$ and $j \in [r]$. Furthermore, if row i is pair-complete then $|Y_j^i \cap S_j^i|, |Y_j^i \setminus S_j^i| \geq n - \beta n/2$ for any $j \in [r]$.

(A3) Let $v \in Y_j^i$. Then v has at most βn non-neighbours in $W_{j'}^{i'}$ for any $i' \neq i$ and $j' \neq j$. Furthermore, if row i is pair-complete and $j' \neq j$ then v has at most βn non-neighbours in $S_{j'}^i$ if $v \in S_j^i$, and at most βn non-neighbours in $W_{j'}^i \setminus S_{j'}^i$ if $v \notin S_j^i$.

(A4) Let $v \in W_j^i$. Then v has at most $2n/3$ non-neighbours in any block $W_{j'}^{i'}$ which is good with respect to v. Furthermore, if row i is pair-complete then there is some $j' \neq j$ such that v has at most $2n/3$ non-neighbours in $S_{j'}^i$ if $v \in S_j^i$, and at most $2n/3$ non-neighbours in $W_{j'}^i \setminus S_{j'}^i$ if $v \notin S_j^i$.

(A5) For any i with $p_i \geq 2$, there are at least $\gamma' n^{p_i+1}$ copies of K_{p_i+1} in $G[Y^i]$. Furthermore, if row i is pair-complete then there are at least $\gamma' n^3$ copies of K_3 in $G[Y^i \cap S^i]$ and at least $\gamma' n^3$ copies of K_3 in $G[Y^i \setminus S^i]$.

Proof. For (A1), note that since there were at most $rk\sqrt{d}p_i n + 2r\sqrt{d}n$ bad vertices in each X_j^i, the total number of bad vertices is at most $rk(rk^2\sqrt{d}n + 2r\sqrt{d}n) + rk \leq \beta n/2$. For (A2), note that $Y_j^i \subseteq W_j^i$, and any vertex of $X_j^i \setminus Y_j^i$ or $W_j^i \setminus X_j^i$ is bad. Since $|X_j^i| = p_i n$, and there are at most $\beta n/2$ bad vertices by (A1), we conclude that (A2) holds. For (A3), note that since v is good we have $|X_{j'}^{i'} \setminus N(v)| \leq \sqrt{d}p_{i'} n$ for any $i' \neq i$ and $j' \neq j$. Since $|W_{j'}^{i'} \setminus X_{j'}^{i'}| \leq \beta n/2$ by (A1), we conclude that $|W_{j'}^{i'} \setminus N(v)| \leq \beta n$, as required. Similarly, if row i is pair-complete and $v \in S_j^i$, then $v \in T_j^i$, so v being good implies that $|T_{j'}^i \setminus N(v)| \leq \sqrt{d}n$ for any $j' \neq j$. On the other hand, if $v \in W_j^i \setminus S_j^i$, then $v \in X_j^i \setminus T_j^i$, so v being good implies that $|(X_{j'}^i \setminus T_{j'}^i) \setminus N(v)| \leq \sqrt{d}n$ for any $j' \neq j$. Any non-neighbour of v in $S_{j'}^i \triangle T_{j'}^i$ or $W_{j'}^i \setminus X_{j'}^i$ must be a bad vertex; by (A1) this completes the proof of (A3).

Next, for (A4) suppose that $W_{j'}^{i'}$ is good with respect to v. Recall that this means $|X_{j'}^{i'} \setminus N(v)| \leq n/2$. Since any vertex in $W_{j'}^{i'} \setminus X_{j'}^{i'}$ is bad, we find that $|W_{j'}^{i'} \setminus N(v)| \leq n/2 + \beta n/2 \leq 2n/3$ by (A1). So suppose now that row i is pair-complete. If v is good, then the 'furthermore' statement holds by (A3), so we may suppose that $v \in W_j^i \setminus Y_j^i$. If $v \in S_j^i$ then by definition $|N(v) \cap T_{j'}^i| \geq n/2$ for some $j' \neq j$, so $|T_{j'}^i \setminus N(v)| \leq n/2$; then $|S_{j'}^i \setminus N(v)| \leq 2n/3$ by (A1), since any vertex in $S_{j'}^i \triangle T_{j'}^i$ is bad. On the other hand, if $v \notin S_j^i$ then by definition $|N(v) \cap T_{j'}^i| < n/2$, so $|T_{j'}^i \setminus N(v)| > n/2$, for any $j' \neq j$. By (†) this implies $|(X_{j'}^i \setminus T_{j'}^i) \setminus N(v)| < n/2$, and so $|(W_{j'}^i \setminus S_{j'}^i) \setminus N(v)| \leq 2n/3$ by (A1).

Finally, for (A5) suppose first that row i is pair-complete, so $p_i = 2$. Then by (A2) we have $|Y_1^i \cap S_1^i|, |Y_2^i \cap S_2^i|, |Y_3^i \cap S_3^i| \geq n - \beta n/2$. Furthermore, by (A3) any vertex in one of these three sets has at most βn non-neighbours in each of the other two sets. So we may choose vertices $v_1 \in Y_1^i \cap S_1^i$, $v_2 \in Y_2^i \cap S_2^i \cap N(v_1)$ and $v_3 \in Y_3^i \cap S_3^i \cap N(v_1) \cap N(v_2)$ in turn with at least $n - 3\beta n$ choices for each vertex. We conclude that there are at least $n^3/2$ copies of K_3 in $G[Y^i \cap S^i]$. The same argument applied to the sets $Y_1^i \setminus S_1^i$ shows that there are at least $n^3/2$ copies of K_3 in $G[Y^i \setminus S^i]$. On the other hand, if row i is not pair-complete, then we simply wish to find at least $\gamma' n^{p_i+1}$ copies of K_{p_i+1} in $G[Y^i]$. Since $G_1[X^i]$ is not d'-splittable, by Lemma 4.2(ii) (with $2\gamma'$ in place of β) there are at least $2\gamma' n^{p_i+1}$ copies of K_{p_i+1} in $G_1[X^i]$. By (A1) at most $\beta(p_i n)^{p_i+1} \leq \gamma' n^{p_i+1}$ such copies contain a bad vertex; this leaves at least $\gamma' n^{p_i+1}$ copies of K_{p_i+1} in $G[Y^i]$. $\qquad\square$

In the next claim we analyse a general greedy algorithm that takes some fixed clique K'' in which all but at most one vertex is good, and extends it to a copy of K_k with prescribed intersections with the blocks, described by the sets $A_1, \ldots, A_s$.

Claim 7.4. (Extending cliques) *Let K'' be a clique in G on vertices $v_1, \ldots, v_m$, where $v_2, v_3, \ldots, v_m$ are good. Suppose that $A_1, \ldots, A_s \subseteq [r]$ are pairwise-disjoint sets such that for each $q \in [m]$ there is some $i \in [s]$ and $j \in A_i$ such that $v_q \in W_j^i$. Suppose also that for each $i \in [s]$ one of the following five conditions holds:*

(a) for every $j \in A_i$, W_j^i contains some v_q with $q \in [m]$,

(b) $|A_i| \le p_i$ and $V(K'') = \emptyset$,

(c) $|A_i| \le p_i$ and there is some block W_j^i with $j \in A_i$ which is good with respect to v_1 and does not contain a vertex of K'',

(d) $|A_i| \le p_i$ and $v_1 \in W^i$,

(e) $|A_i| < p_i$.

Let $a := \sum_{i \in [s]} |A_i|$. Then for any $b_1, \ldots, b_s \in \{0,1\}$ there are at least ωn^{a-m} copies K' of K_a in $G[Y \cup \{v_1\}]$ which extend K'' and satisfy the following properties.

(i) K' intersects precisely those W_j^i with $i \in [s]$ and $j \in A_i$.

(ii) For any pair-complete row i such that $|A_i| = 2$ and $V(K'') \cap W^i = \emptyset$ we have that $|V(K') \cap S^i|$ is even.

(iii) Consider any pair-complete row i such that $|A_i| = 1$ and write $\{j\} = A_i$. If W_j^i is good with respect to v_1 and does not contain a vertex of K'' then $|V(K') \cap S^i| = b_i$.

Proof. If $V(K'') \ne \emptyset$, then by relabelling the columns W_j if necessary, we may assume that for any $i \in [s]$ for which (c) holds, the block $W_{\max A_i}^i$ is good with respect to v_1 and does not contain a vertex of K''. We also note for future reference that the only properties of good vertices used in the proof of this claim will be those in (A3).

First we define the greedy algorithm for extending K'' to K', and then we will show that we have many choices for the vertex at each step of the algorithm. We proceed through each column V_j, $j \in [r]$ in turn. If j is not in A_i for any $i \in [s]$, then we take no action, since K' will not have a vertex in this column. Similarly, if $v_q \in W_j$ for some $q \in [m]$, then we again take no action, since we already have a vertex of K' in this column, namely v_q; note that $v_q \in W_j^i$ for the unique i such that $j \in A_i$, since the sets A_ℓ are pairwise-disjoint. Now suppose that $j \in A_i$ for some $i \in [s]$, and $V(K'') \cap W_j = \emptyset$. Let $v_1', \ldots, v_{t-1}'$ be the vertices previously chosen by the algorithm (so not including $v_1, \ldots, v_m$). We choose a vertex $v_t' \in Y_j^i \cap \bigcap_{\ell \in [m]} N(v_\ell) \cap \bigcap_{\ell \in [t-1]} N(v_\ell')$, so $\{v_1, \ldots, v_m, v_1', \ldots, v_t'\}$ induces a clique in G. If row i is pair-complete, $|A_i| = 2$, $V(K'') \cap W^i = \emptyset$ and we have previously selected a vertex v_ℓ' in W^i, then we also add the requirement that $v_t' \in S^i$ if and only if $v_\ell' \in S^i$. If instead row i is pair-complete and meets the conditions of (iii) then we instead add the requirement that $v_t' \in S^i$ if and only if $b = 1$. After proceeding through every $j \in [r]$ we have a vertex of W_j^i for every $i \in [s]$ and $j \in A_i$ (some of which are the vertices of K''). We let K' be the subgraph of G induced by these vertices. Then K' is a clique of size a in G which extends K'' and satisfies (i). The additional requirements on the choice of vertices from any pair-complete row i imply that K' must satisfy (ii) and (iii) also.

Having defined the greedy algorithm, we will now show that there are at least $n/4$ choices at each step. First we consider the number of choices for some v_t' in Y_j^i, where row i is not a pair-complete row satisfying the conditions in (ii) or (iii). Note that since we are making

this choice, W_j does not contain a vertex of K'', so (a) does not apply to row i. Let

$$P := \{v_1, \ldots, v_m, v_1', \ldots, v_{t-1}'\} \text{ and } P^i := (P \cap W^i) \setminus v_1.$$

Then $|P^i| \leq |A_i| - 1 \leq p_i - 1$. We need to estimate $|Y_j^i \cap \bigcap_{v \in P} N(v)|$. Note that each $v \in P^i$ has $|Y_j^i \setminus N(v)| \leq n$ by (†). If $V(K'') \neq \emptyset$, then write $N' := Y_j^i \setminus N(v_1)$, so $|N'| \leq n$ also; if $V(K'') = \emptyset$ we let $N' = \emptyset$. Observe that any vertex of $P \setminus \{P^i \cup v_1\}$ is good, either by assumption (for $v_2, \ldots, v_m$) or by selection (since the greedy algorithm only selects vertices from some $Y_{j'}^{i'}$). Then any vertex of $P \setminus \{P^i \cup v_1\} = P \setminus W^i$ lies in $Y^{i'}$ for some $i' \neq i$, and therefore has at most βn non-neighbours in Y_j^i by (A3). So $|Y_j^i \cap \bigcap_{v \in P} N(v)|$ is at least

$$(3) \qquad |Y_j^i| - |P^i|n - |N'| - r\beta n \overset{(A2)}{\geq} p_i n - (p_i - 1)n - n - (r+1)\beta n = -(r+1)\beta n.$$

Whilst this crude bound does not imply that we have even one possible choice for v_t', we will now show that any of the assumptions (b)–(e) improves some part of the bound by at least $n/3$, which implies that there are at least $n/3 - (r+1)\beta n \geq n/4$ choices for v_t'. If (b) pertains to row i (so $V(K'') = \emptyset$), then we have $|N'| = 0$ instead of $|N'| \leq n$. If (d) or (e) pertains to row i, then we have the bound $|P^i| \leq p_i - 2$ in place of $|P^i| \leq p_i - 1$. The same is true if (c) pertains to row i, unless we are choosing the final vertex in W^i, that is $j = \max A_i$. Then our initial relabelling implies that W_j^i is good with respect to v_1, so (A4) gives the bound $|N'| = |Y_j^i \setminus N(v_1)| \leq 2n/3$ in place of $|N'| \leq n$. In all cases the improvement of at least $n/3$ to (3) yields at least $n/3 - (r+1)\beta n \geq n/4$ choices for v_t'.

It remains to consider the number of choices for some v_t' in Y_j^i, where row i is a pair-complete row satisfying the conditions in (ii) or (iii). Suppose first that row i has the conditions of (ii), namely $|A_i| = 2 = p_i$ and $V(K'') \cap W^i = \emptyset$. Suppose also that we have previously selected a vertex $v_\ell' \in W^i$. Then we must ensure that $v_t' \in S_j^i$ if and only if $v_\ell' \in S^i$. Note that these conditions imply that either (b) or (c) pertains to row i. We will show that $|N'| \leq 2n/3$ in either case. In case of (b) this holds because N' is empty. In case of (c), v_t' is the final vertex to be selected in row i, and so our initial relabelling implies that W_j^i is good with respect to v_1; then $|N'| \leq 2n/3$ by definition. Since v_ℓ' is a good vertex (by choice), if $v_\ell' \in S^i$ then by (A3) we have $|S_j^i \setminus N(v_\ell')| \leq \beta n$, or if $v_\ell' \notin S^i$ then by (A3) we have $|(Y_j^i \setminus S_j^i) \setminus N(v_\ell')| \leq \beta n$. In the former case we have

$$\left| Y_j^i \cap S_j^i \cap \bigcap_{v \in P} N(v) \right| \geq |Y_j^i| - |N'| - r\beta n \geq n/4;$$

similarly, in the latter case we obtain $|(Y_j^i \setminus S_j^i) \cap \bigcap_{v \in P} N(v)| \geq n/4$. So in either case there are at least $n/4$ possible choices for v_t'. Finally, suppose that row i has the conditions of (iii), namely $|A_i| = 1$ and W_j^i is good with respect to v_1 (and W_j^i does not contain a vertex of K'' since we are choosing v_t'). Then $|N'| \leq 2n/3$ by (A4), so $|Y_j^i \cap S_j^i \cap \bigcap_{v \in P} N(v)| \geq n/4$ and $|(Y_j^i \setminus S_j^i) \cap \bigcap_{v \in P} N(v)| \geq n/4$, giving us at least $n/4$ choices for v_t', regardless of the value of b.

In conclusion, there are at least $n/4$ suitable choices for each of the $a - m$ vertices chosen by the greedy algorithm in extending K'' to K', giving at least ωn^{a-m} choices for K', as required. $\square$

Now we describe the various types of K_k that will form the building blocks for the deletions. Recall that the number of vertices in each row W^i is approximately proportional to p_i. We say that a copy K' of K_k in G is *properly-distributed* if

  (i) $|V(K') \cap W^i| = p_i$ for each $i \in [s]$, and
  (ii) $|V(K') \cap S^i|$ is even for any pair-complete row $i \in [s]$.

Also, for any $i, j \in [s]$ with $i \neq j$ we say that a copy K' of K_k in G is *ij-distributed* if

  (i) $|V(K') \cap W^i| = p_i + 1$, $|V(K') \cap W^j| = p_j - 1$, and $|V(K') \cap W^\ell| = p_\ell$ for each $\ell \in [s] \setminus \{i, j\}$, and
  (ii) $|V(K') \cap S^\ell|$ is even for any pair-complete row $\ell \neq i, j$.

Note that an *ij-distributed* clique K' has 'one too many' vertices in row i, and 'one too few' in row j. By deleting such cliques we can arrange that the size of each row W^i is *exactly* proportional to p_i. Thereafter we will only delete properly-distributed copies of K_k, so that this property is preserved. Also, condition (ii) in both definitions ensures that we preserve the correct parity of the halves in pair-complete rows. Finally, we say that a copy K' of K_k in G is *properly-distributed outside row ℓ* if

  (i) $|V(K') \cap W^i| = p_i$ for each $i \in [s]$, and
  (ii) $|V(K') \cap S^i|$ is even for any pair-complete row $i \neq \ell$.

Thus K' almost satisfies the definition of 'properly-distributed', except that if row ℓ is pair-complete it may fail the parity condition for the halves. In the next claim we apply Claim 7.4 to finding the building blocks just described.

Claim 7.5. *(Building blocks)* We can find copies of K_k in G as follows.

  (i) *Let $A_1, \ldots, A_s \subseteq [r]$ be pairwise-disjoint with $|A_i| = p_i$ for each $i \in [s]$. Then there are at least ωn^k properly-distributed copies K' of K_k in G such that for any $i \in [s]$, K' intersects W^i in precisely those W^i_j with $j \in A_i$.*
  (ii) *Any vertex $v \in V(G)$ lies in at least $\omega n^{k-1}/4$ properly-distributed copies of K_k in G.*
  (iii) *Let $i, j \in [s]$ be such that $p_i \geq 2$ and $i \neq j$. Then there are at least γn^k ij-distributed copies K' of K_k in $G[Y]$. Furthermore, if row i is pair-complete then for any $b \in \{0, 3\}$ there are at least γn^k such copies K' of K_k with $|V(K') \cap S^i| = b$.*
  (iv) *Let $i, j \in [s]$ be such that $p_i = 1$ and $i \neq j$. Suppose uv is an edge in $G[W^i]$ such that u is a good vertex. Then there are at least ωn^{k-2} ij-distributed copies K' of K_k in G that contain u and v. Furthermore, if row j is pair-complete then for any $b \in \{0, 1\}$ there are at least ωn^{k-2} such copies K' of K_k with $|V(K') \cap S^j| = b$.*
  (v) *Let $i \in [s]$ be such that row i is pair-complete, and uv be an edge in $G[W^i]$ such that u is good. Then there are at least ωn^{k-2} copies of K_k in G which contain both u and v and are properly-distributed outside row i.*

Proof. For (i) we apply Claim 7.4 to $A_1, \ldots, A_s$ with $V(K'') = \emptyset$, so condition (b) pertains to all rows. This gives at least ωn^k copies K' of K_k in G such that K' intersects precisely those W^i_j with $i \in [s]$ and $j \in A_i$, and $|V(K') \cap S^i|$ is even for any pair-complete row i. Since $|A_i| = p_i$ for each $i \in [s]$, each such K' is properly-distributed. Next we consider (iii), as this is also a simple application of Claim 7.4; we will come back to (ii). We begin by choosing a copy K'' of K_{p_i+1} in $G[Y^i]$. By (A5), there are at least $\gamma' n^{p_i+1}$ such copies, and if row i is pair-complete, there are at least $\gamma' n^{p_i+1}$ such copies with precisely b vertices in S^i. Fix any such K'' and let A_i be the set of $q \in [r]$ such that K'' has a vertex in column V_q, so $|A_i| = p_i + 1$. Now choose pairwise-disjoint subsets $A_\ell \subseteq [r] \setminus A_i$ with $A_j = p_j - 1$

and $|A_\ell| = p_\ell$ for every $\ell \in [s] \setminus \{i, j\}$. We may apply Claim 7.4 with K'' and the sets A_ℓ for $\ell \in [s]$, as condition (a) of the claim applies to row i, condition (e) applies to row j, and condition (c) applies to all other rows (since every vertex of K'' is good). We deduce that there are at least ωn^{k-p_i-1} copies K' of K_k which extend K'' such that $|V(K') \cap W^\ell| = |A_\ell|$ for any $\ell \in [s]$ and $|V(K') \cap S^\ell|$ is even for any pair-complete row $\ell \neq i, j$. Each such K' is ij-distributed, so in total we have at least $\gamma' \omega n^k \geq \gamma n^k$ copies K' of K_k with the required properties.

For (ii), (iv) and (v) we proceed similarly, but in each of these cases we have the possibility that the vertex v might be bad, so satisfying requirement (c) in Claim 7.4 is no longer trivial. We consider the blocks as an $s \times r$ rectangle R, and colour those blocks which are bad with respect to v. Our strategy will be to delete some appropriate rows and columns from R, apply Proposition 7.1 to find a transversal T of non-coloured blocks in the remaining subrectangle R', and then use T to choose sets A_i for $i \in [s]$ which meet the conditions of Claim 7.4. (The notation T is not intended to suggest any relationship with the sets T^i_j.) We note the following properties of R:

(R1) any column not containing v has at most one coloured block,
(R2) at most $r - 2$ coloured blocks lie outside the row containing v,
(R3) any row not containing v has at most $r - 3$ coloured blocks.

To see this, recall that we observed (R1) earlier, and (R2) follows because v lies in a row in which it has the most bad blocks. This also implies (R3), as any row not containing v has at most $(r-1)/2 \leq r - 3$ coloured blocks, using the assumption that $r > 3$.

For (ii), we let ℓ and j_ℓ be such that $v \in W^\ell_{j_\ell}$, and suppose first that row ℓ is not pair-complete (we postpone the case where ℓ is pair-complete to the end of the proof). We remove the row and column containing v to obtain an $(s-1) \times (r-1)$ subrectangle R'. Since R' contains at most $r-2$ coloured blocks, with at most one in each column, it contains a transversal T of non-coloured blocks by Proposition 7.1. For each $i \in [s] \setminus \{\ell\}$ let j_i be such that T includes $W^i_{j_i}$. Choose any pairwise-disjoint subsets $A_i \subseteq [r]$ with $j_i \in A_i$ and $|A_i| = p_i$ for each $i \in [s]$. We may apply Claim 7.4 to A_i, $i \in [s]$ with $V(K'') = \{v_1\} = \{v\}$, since condition (d) pertains to row ℓ, and condition (c) pertains to all other rows. This yields at least ωn^{k-1} copies K' of K_k which extend K'' such that $|V(K') \cap S^i|$ is even for any pair-complete row i and $|V(K') \cap W^i| = |A_i| = p_i$ for any $i \in [s]$. Each such K' is a properly-distributed copy of K_k in G containing v.

For (iv), let j_u and j_v be the columns with $u \in V_{j_u}$ and $v \in V_{j_v}$, and let $A_i = \{j_u, j_v\}$. So $|A_i| = 2 = p_i + 1$. Suppose first that $p_j = 1$. In this case we delete rows i and j and columns j_u and j_v from R, leaving an $(s-2) \times (r-2)$ subrectangle R'. Then R' has at most one coloured block in each column, at most $r - 3$ coloured blocks in each row, and at most $r - 2$ coloured blocks in total, so contains a transversal T of non-coloured blocks by Proposition 7.1. For each $\ell \in [s] \setminus \{i, j\}$ let q_ℓ be the column such that T includes $W^\ell_{q_\ell}$, and choose pairwise-disjoint subsets $A_\ell \subseteq [r] \setminus A_i$ for $\ell \in [s] \setminus \{i\}$ such that $q_\ell \in A_\ell$ and $|A_\ell| = p_\ell$ for each $\ell \in [s] \setminus \{i, j\}$, and $A_j = \emptyset$. Now suppose instead that $p_j \geq 2$. In this case we delete row i and columns j_u and j_v from R, leaving an $(s-1) \times (r-2)$ subrectangle R'. Then R' has at most one coloured block in each column, at most $r - 3$ coloured blocks in each row, and at most $r - 2$ coloured blocks in total. Since $p_j \geq 2$ we have $r - 2 \geq k - 2 \geq s - 1$, so we may again apply Proposition 7.1 to find a transversal T of non-coloured blocks in R'. For each $\ell \in [s] \setminus \{i\}$ let q_ℓ be the column such that T includes $W^\ell_{q_\ell}$, and choose pairwise-disjoint

subsets $A_\ell \subseteq [r] \setminus A_i$ for $\ell \in [s] \setminus \{i\}$ such that $q_\ell \in A_\ell$ for any $\ell \in [s] \setminus \{i\}$, $|A_\ell| = p_\ell$ for each $\ell \in [s] \setminus \{i, j\}$, and $|A_j| = p_j - 1$. In either case we may apply Claim 7.4 with $V(K'') = \{u, v\}$, since conditions (a), (e) and (c) of Claim 7.4 pertain to rows i, j and all other rows respectively. We deduce that there are at least ωn^{k-2} copies K' of K_k in G which extend K'' such that $|V(K') \cap W^\ell| = |A_\ell|$ for any $\ell \in [s]$ and $|V(K') \cap S^\ell|$ is even for any pair-complete row $\ell \neq i, j$. Each such K' is ij-distributed. Furthermore, if row j is pair-complete, then we can apply Claim 7.4(iii), as $W_{q_j}^j \in T$ is good with respect to v by definition of T, to see that there are at least ωn^{k-2} such copies K' with $|V(K') \cap S^j| = b$, as required.

For (v), we again let j_u and j_v be the columns with $u \in V_{j_u}$ and $v \in V_{j_v}$ and $A_i = \{j_u, j_v\}$. So $|A_i| = 2 = p_i$. We delete row i and columns j_u and j_v from R, leaving an $(s-1) \times (r-2)$ subrectangle R'. Then R' has at most one coloured block in each column, at most $r-3$ coloured blocks in each row, and at most $r-2$ coloured blocks in total. Since $p_i = 2$ we have $r - 2 \geq k - 2 \geq s - 1$, so as before Proposition 7.1 yields a transversal T of non-coloured blocks in R'. We let $W_{q_\ell}^\ell$, $\ell \in [s] \setminus \{i\}$ be the blocks of T, and choose pairwise-disjoint subsets $A_\ell \subseteq [r] \setminus A_i$ such that $q_\ell \in A_\ell$ and $|A_\ell| = p_\ell$ for any $\ell \in [s] \setminus \{i\}$. We may apply Claim 7.4 with $V(K'') = \{u, v\}$, since conditions (a) and (c) of Claim 7.4 pertain to row i and all other rows respectively. So there are at least ωn^{k-2} copies K' of K_k in G which extend K'' such that $|V(K') \cap W^\ell| = |A_\ell| = p_\ell$ for any $\ell \in [s]$ and $|V(K') \cap S^\ell|$ is even for any pair-complete row $\ell \neq i$. Each such K_k is properly-distributed outside row i.

Finally we consider the case of (ii) where row ℓ (containing v) is pair-complete. By (A4) and (A2) there are at least $n/4$ vertices $u \in Y^\ell \cap N(v)$ such that $|\{u, v\} \cap S^\ell|$ is even. For any such u, by (v) there are at least ωn^{k-2} copies K' of K_k which extend uv and are properly-distributed outside row ℓ. Each such K' is properly-distributed since $V(K') \cap S^\ell = \{u, v\}$, so v lies in at least $\omega n^{k-1}/4$ properly-distributed copies of K_k. $\qquad\square$

We now use Claim 7.5 to select several K_k-packings in G whose deletion leaves a subgraph G' which satisfies the conclusions of the lemma. The choice of these packings will vary somewhat according to the row structure of G. We say that G has the *extremal row structure* if it has the same row structure as the graphs $\Gamma_{n,r,k}$ of Construction 1.2, that is, if there is some row $i \in [s]$ with $p_i \geq 2$ which is pair-complete, and $p_{i'} = 1$ for any $i' \neq i$; this case requires special attention. The first step is the following claim, which balances the row sizes, so that the remainder of row i has size proportional to p_i for $i \in [s]$. This is the only step of the proof that requires the exact minimum degree condition (whether or not G has the extremal row structure).

Claim 7.6. (Balancing rows) *There is a K_k-packing M_1 in G such that $|M_1| \leq \beta r k^2 n$ and $|W^i \setminus V(M_1)| = p_i(rn^+/k - |M_1|)$ is a constant multiple of p_i for each $i \in [s]$. If G has the extremal row structure we can also require that $|S^i \setminus V(M_1)|$ is even for the unique $i \in [s]$ with $p_i = 2$.*

Proof. We write $a_i := |W^i| - p_i rn^+/k$ for each $i \in [s]$, where we recall that rn^+/k is an integer, so a_i is an integer. Then $\sum_{i \in [s]} a_i = |V(G^+)| - rn^+ = 0$. Let $I^+ = \{i : a_i > 0\}$, $I^- = \{i : a_i < 0\}$ and $a := \sum_{a_i > 0} a_i = -\sum_{a_i < 0} a_i$. Recall that $n^+ - k + 1 \leq kn \leq n^+$ and $(1 + \beta/2)p_i rn \geq |W^i| = \sum_{j \in [r]} |W_j^i| \geq (1 - \beta/2)p_i rn$ for each i by (A2). This gives $|a_i| \leq \beta r k n$ for each i and $a \leq \beta r k^2 n$. We choose sequences $(i_\ell : \ell \in [a])$ and $(i'_\ell : \ell \in [a])$ so that each $i \in I^+$ occurs a_i times as some i_ℓ, and each $i \in I^-$ occurs $-a_i$ times as some i'_ℓ.

We divide the remainder of the proof into two cases according to whether or not G has the extremal row structure.

Case 1: G does not have the extremal row structure. We start by showing that for any $i \in I^+$ with $p_i = 1$ there is a matching E^i in $G[W^i]$ of size a_i, each of whose edges contains a good vertex. We let $a_{ij} := |W^i_j| - n^+/k$ for each $j \in [r]$, so $|a_{ij}| \leq \beta n$ for each j by (A2); note that the a_{ij} are not necessarily integers. Then $|W^i_j| = n^+/k + a_{ij} \geq n + a_{ij}$. By (†), any $v \in W^i_{j'}$ for any $j' \neq j$ has at least a_{ij} neighbours in W^i_j. We fix a matching M in the bipartite graph $G[W^i_1 \cup W^i_2]$ which is maximal with the property that every edge contains a good vertex. We will show that $|M| \geq a_{i1} + a_{i2}$. For suppose otherwise. Then we may fix $u \in Y^i_1$ and $v \in Y^i_2$ which are not covered by M. As noted above, u has at least a_{i2} neighbours in W^i_2, and v has at least a_{i1} neighbours in W^i_1. Since u and v are good vertices, each of these neighbours must be covered by M, by maximality of M. Thus there must be an edge $w_1 w_2$ in M with $w_1 \in N(v)$ and $w_2 \in N(u)$. Now removing $w_1 w_2$ from M and adding the edges $w_1 v$ and $w_2 u$ contradicts the maximality of M. We conclude that $|M| \geq a_{i1} + a_{i2}$. Now we greedily extend M to a maximal matching E^i in $G[W^i]$ for which every edge of $E^i \setminus M$ contains a vertex of Y^i_1. Then E^i covers at least a_{ij} vertices in W^i_j for each $j \geq 3$, and so has total size at least $\sum_{j \in [r]} a_{ij} = a_i$; by removing edges we may take E^i to have size exactly a_i, as claimed.

We now proceed through each $\ell \in [a]$ in turn. For each ℓ, if $p_{i_\ell} = 1$ then we choose an edge $e \in E^{i_\ell}$ which has not been used for any $\ell' < \ell$ (this is possible as i_ℓ occurs $a_{i_\ell} = |E^{i_\ell}|$ times as some $i_{\ell'}$). We extend e to an $i_\ell i'_\ell$-distributed copy K^ℓ of K_k in G, so that K^ℓ is vertex-disjoint from any $K^{\ell'}$ previously selected for $\ell' < \ell$ and from any of the edges of $\bigcup_{i \in I^+} E^i$ other than e itself. Then the extension of e to K^ℓ must avoid a set of at most $ka + 2a \leq \beta' n$ 'forbidden vertices'. By Claim 7.5(iv), there are at least $\omega n^{k-2} > \beta' n (rn^+)^{k-3}$ extensions of e to an $i_\ell i'_\ell$-distributed copy of K_k, so we may choose K' as desired. Similarly, if $p_{i_\ell} \geq 2$ then we choose an $i_\ell i'_\ell$-distributed copy K^ℓ of K_k in G which is vertex-disjoint from any previously chosen copy of K_k and from the matchings E^i. As before this means we must avoid at most $\beta' n$ forbidden vertices. By Claim 7.5(iii), there are at least $\gamma n^k \geq \beta' n (rn^+)^{k-1}$ $i_\ell i'_\ell$-distributed copies of K_k in G, so we can choose K^ℓ as desired. At the end of this process we have a K_k-packing $M_1 := \{K^\ell : \ell \in [a]\}$ with $|M_1| = a \leq \beta r k^2 n$.

Now consider any $i \in [s]$. If $i \in I^+$ (so $a_i > 0$) then each copy of K_k in M_1 had p_i vertices in W^i, except for a_i copies which had $p_i + 1$ vertices in W^i (these are the K^ℓ with $i_\ell = i$). On the other hand, if $i \notin I^+$ (so $a_i \leq 0$) then each copy of K_k in M_1 had p_i vertices in W^i, except for $-a_i$ copies which had $p_i - 1$ vertices in W^i (these are the K^ℓ with $i'_\ell = i$). In any case $|W^i \cap V(M_1)| = p_i |M_1| + a_i = p_i a + |W^i| - p_i r n^+/k$. Then $|W^i \setminus V(M_1)| = p_i(rn^+/k - a)$ is a constant multiple of p_i for each $i \in [s]$, so the proof of Case 1 is complete.

Case 2: G has the extremal row structure. Let i^* be the unique row such that $p_{i^*} = 2$, so row i^* is pair-complete and $p_i = 1$ for any $i \neq i^*$. Recall that in this case we must satisfy the additional condition that $|S^{i^*} \setminus V(M_1)|$ is even. We divide the proof into two further cases according to whether or not $a = 0$.

Case 2.1: $a > 0$. We start by choosing copies K^ℓ of K_k for $\ell < a$ exactly as in Case 1. For $\ell = a$ we consider three further cases, according to the value of i^*.

Case 2.1.1: $i^* = i_a$. As in Case 1, we use Claim 7.5(iii) to choose K^a to be an $i_a i'_a$-distributed copy of K_k in G that does not include a forbidden vertex. However, we also

choose K^a so that $|S^{i^*} \setminus \bigcup_{\ell \in [a]} V(K^\ell)|$ is even. This is possible as by Claim 7.5(iii) there are at least γn^k $i_a i'_a$-distributed copies K' of K_k in G for each choice of parity of $|V(K') \cap S^{i_a}|$.

Case 2.1.2: $i^* = i'_a$. As in Case 1, we use Claim 7.5(iv) to extend an edge $e \in E^{i_a}$ to an $i_a i'_a$-distributed copy K^a of K_k in G avoiding all forbidden vertices. Here we again impose the additional requirement that $|S^{i^*} \setminus \bigcup_{\ell \in [a]} V(K^\ell)|$ is even. This is possible as by Claim 7.5(iv) there are at least ωn^{k-2} extensions of e to an $i_a i'_a$-distributed copy K' of K_k in G for each choice of parity of $|V(K') \cap S^{i'_a}|$.

Case 2.1.3: $i^* \neq i_a, i'_a$. In this case, instead of choosing the extension K^a of $e \in E^{i_a}$ to be $i_a i'_a$-distributed, we extend e to an $i_a i^*$-distributed copy K^a of K_k in G which avoids any forbidden vertices (using Claim 7.5(iv) as before). The K_k-packing M_1 then covers 'one too many' vertices in $W^{i'_a}$ and 'one too few' in W^{i^*}. To correct this imbalance, we apply Claim 7.5(iii) to choose an $i^* i'_a$-distributed copy K' of K_k in G such that K' is vertex-disjoint from M_1 and $|S^{i^*} \setminus (V(K') \cup V(M_1))|$ is even. We add K' to M_1.

In each of the cases 2.1.1, 2.1.2, 2.1.3 we have $|W^i \setminus V(M_1)| = p_i(rn^+/k - a)$ for each $i \in [s]$ as in Case 1. We also have $|S^{i^*} \setminus V(M_1)|$ even by construction, and $|M_1| \leq \beta r k^2 n$, noting in Case 2.1.3 that we could have improved the bound on a to $a \leq \beta r k^2 n - 1$. This completes the proof in Case 2.1.

Case 2.2: $a = 0$. Note that $a_i = 0$ for each $i \in [s]$, so $|W^i|/p_i = rn^+/k$ is constant initially. If $|S^{i^*}|$ is even then $M_0 = \emptyset$ has the required properties, so we may assume $|S^{i^*}|$ is odd. Now suppose that for some $i \neq i^*$ there is an edge uv in $G[W^i]$ such that u is good. By Claim 7.5(iv) we may extend uv to an ii^*-distributed copy K^1 of K_k in G. Now apply Claim 7.5(iii) to choose an i^*i-distributed copy K^2 of K_k in G which does not intersect K^1 so that $|S^{i^*} \setminus (V(K^1) \cup V(K^2))|$ is even. We then have $M_1 = \{K^1, K^2\}$ with the required properties. So we may assume that no such edge exists. Now suppose instead that $G[W^{i^*}]$ contains an edge uv such that u is good and $|\{u,v\} \cap S^{i^*}| = 1$. By Claim 7.5(v) we may extend uv to a copy K' of K_k which is properly-distributed outside row i^*. We may take $M_1 = \{K'\}$, as then $|V(M_1) \cap S^i| = p_i$ for any $i \in [s]$ and $|S^{i^*} \setminus V(M_1))|$ is even. So we may assume that no such edge exists. It follows that $|W_j^i| \leq n$ for any $i \neq i^*$ and $j \in [r]$. Indeed, if $|W_j^i| > n$ then choose $j' \neq j$ and a good vertex $u \in Y_{j'}^i$. Then u has a neighbour $v \in W_j^i$, since u has at most n non-neighbours in any part, but we assumed no such edge uv exists. Likewise, it follows that $|S_j^{i^*}|, |W_j^{i^*} \setminus S_j^{i^*}| \leq n$ for any $j \in [r]$. For if (say) $|S_j^{i^*}| > n$, then choose $j' \neq j$ and $u \in Y_{j'}^{i^*} \setminus S_{j'}^{i^*}$; then u has a neighbour $v \in S_j^{i^*}$, giving an edge uv that we assumed did not exist.

Now for any $j \in [r]$ we have $|W_j^{i^*}| \leq 2n$, $|S_j^{i^*}| \leq n$, $|W_j^{i^*} \setminus S_j^{i^*}| \leq n$, and $|W_j^i| \leq n$ for any $i \neq i^*$. Since $|W_j| = |V_j| = n^+ \geq kn$, we have $n^+ = kn$, and equality holds in each of these inequalities. Then k divides n^+ and $rn^+/k = rn = |S^{i^*}|$ is odd. Now, any good vertex $u \in Y_j^i$ for $i \in [s]$ and $j \in [r]$ has $\delta^*(G) \geq (k-1)n$ neighbours in $V_{j'}$ for any $j' \neq j$. If $i \neq i^*$ then none of these can lie in $W_{j'}^i$, so we conclude that $N(u) \cap V_{j'} = V_{j'} \setminus W_{j'}^i$. Similarly, we deduce that any $u \in Y_j^{i^*} \cap S_j^{i^*}$ has $N(u) \cap V_{j'} = V_{j'} \setminus (W_{j'}^{i^*} \setminus S_{j'}^{i^*})$ and any $u \in Y_j^{i^*} \setminus S_j^{i^*}$ has $N(u) \cap V_{j'} = V_{j'} \setminus S_{j'}^{i^*}$ for any $j' \neq j$. It follows that for any $i \neq i'$ and $j \neq j'$ any vertex $v \in W_j^i$ has $|N(v) \cap W_{j'}^{i'}| \geq |Y_{j'}^{i'}| \geq (1-\beta)p_{i'}n$. Furthermore, any vertex $v \in S_j^{i^*}$ has $|N(v) \cap S_{j'}^{i^*}| \geq |Y_{j'}^{i^*} \cap S_{j'}^{i^*}| \geq (1-\beta)n$, and any vertex $v \in W_j^i \setminus S_j^i$ has

$|N(v) \cap (W^i_{j'} \setminus S^i_{j'})| \geq |Y^i_{j'} \setminus S^i_{j'}| \geq (1 - \beta)n$. So every vertex of G satisfies property (A3), which was shown to hold for all good vertices of G in Claim 7.3. This was the only property of good vertices used in the proof of Claim 7.4. So we may consider every vertex of G to be good when applying Claim 7.4, and therefore also when applying Claim 7.5. The argument above then shows that we may choose M_1 as desired if there exists any edge uv in W^i for any $i \neq i^*$, or any edge uv in W^{i^*} such that $|\{u, v\} \cap S^{i^*}| = 1$. Since G is not isomorphic to $\Gamma_{n^+,r,k}$, there must be some such edge. This completes the proof of Case 2.2, and so of the claim. $\qquad\square$

In the next claim we put aside an extra K_k-packing M_2 that is needed to provide flexibility later in the case when there are at least two rows with $p_i \geq 2$. If instead G has at most one row i with $p_i \geq 2$ then we take $M_2 = \emptyset$. The proof of the claim is immediate from Claim 7.5(iii) so we omit it.

Claim 7.7. (Preparing multiple rows with $p_i \geq 2$) *Suppose that at least two rows of G have $p_i \geq 2$. Then there is a K_k-packing M_2 vertex-disjoint from M_1 such that*

(i) *every vertex covered by M_2 is good, and*

(ii) *M_2 consists of $\lceil \eta n \rceil$ ij-distributed copies of K_k in $G[Y]$ for each ordered pair (i, j) with $i, j \in [s]$, $i \neq j$ and $p_i, p_j \geq 2$.*

Note that (ii) implies that $|M_2| \leq \beta n$ and $|W^i \setminus V(M_1 \cup M_2)|/p_i$ is constant for $i \in [s]$ by choice of M_1. The latter property will be preserved when we remove all subsequent packings, as they will only consist of properly-distributed cliques. Next we cover all remaining bad vertices and ensure that the number of remaining vertices is divisible by $rk \cdot r!$. For this, let B' denote the set of all bad vertices in G which are not covered by $M_1 \cup M_2$, and note that by (A1) we have $|B'| \leq \beta n/2$.

Claim 7.8. (Covering bad vertices and ensuring divisibility) *There is a K_k-packing M_3 vertex-disjoint from $M_1 \cup M_2$, consisting of at most βn properly-distributed copies of K_k, such that*

(i) *$\bigcup_{i \in [3]} M_i$ covers every vertex of B', and*

(ii) *the number of vertices not covered by $\bigcup_{i \in [3]} M_i$ is divisible by $rk \cdot r!$.*

Proof. Fix $\beta n/2 \leq C \leq \beta n$ so that $C \equiv rn^+/k - |M_1 \cup M_2|$ modulo $r \cdot r!$ (recall that rn^+/k is an integer). We will greedily choose M_3 to consist of C copies of K_k. As long as some vertex $v \in B'$ remains uncovered, we choose some such v and select a properly-distributed copy of K_k containing v. Once every vertex in B' is covered we repeatedly choose a properly-distributed copy of K_k until we have C copies of K_k in total. At any step there are at most $k(|M_1 \cup M_2| + C) \leq \beta' n$ forbidden vertices covered by $M_1 \cup M_2$ or a previously-chosen member of M_3, so there will always be a suitable choice available for the next member of M by Claim 7.5(i) or (ii). Fix such an M_3. Then $|V(G) \setminus \bigcup_{i \in [3]} V(M_i)|/k = rn^+/k - |M_1 \cup M_2| - C \equiv 0$ modulo $r \cdot r!$, so $|V(G) \setminus \bigcup_{i \in [3]} V(M_i)|$ is divisible by $kr \cdot r!$. $\qquad\square$

The penultimate K_k-packing is chosen so that equally many vertices are covered in each part V_j.

Claim 7.9. (Balancing columns) *There is a K_k-packing M_4 vertex-disjoint from $\bigcup_{i \in [3]} M_i$, consisting of properly-distributed cliques, such that*

(i) $|V_j \cap \bigcup_{i \in [4]} V(M_i)| = k|\bigcup_{i \in [4]} M_i|/r$ for any $j \in [r]$,

(ii) $|\bigcup_{i \in [4]} V(M_i)| \le \beta' n / 2$, and

(iii) $rk \cdot r!$ divides $|M_4|$.

Proof. We let $a'_j := |\bigcup_{i \in [3]} V(M_i) \cap V_j| - k|\bigcup_{i \in [3]} M_i|/r$ for each $j \in [r]$. So $\sum_{j \in [r]} |V(M_i)| = 0$, and $|a'_j| \le |\bigcup_{i \in [3]} V(M_i)| \le 2\beta r k^3 n$ for each $j \in [r]$. Note also that each a'_j must be an integer since r divides both $|V(G) \setminus \bigcup_{i \in [3]} V(M_i)|$ (by choice of M_3) and $|V(G)|$ (by assumption). Similarly to the proof of Claim 7.6, we let $J^+ := \{j \in [r] : a'_j > 0\}$, $J^- := \{j \in [r] : a'_j < 0\}$, and $a' := \sum_{j \in J^+} a'_j = -\sum_{j \in J^-} a'_j$. We form sequences $j_1, \ldots, j_{a'}$ and $j'_1, \ldots, j'_{a'}$ such that each $j \in J^+$ occurs a'_j times as some j_q, and each $j \in J^-$ occurs $-a'_j$ times as some j'_q. Now, for each $q \in [a']$ choose sets $A_q, A'_q \in \binom{[r]}{k}$ such that $j_q \in A_q$ and $A'_q = (A_q \setminus \{j_q\}) \cup \{j'_q\}$. For each $A \in \binom{[r]}{k}$ let N_A and N'_A be the number of times that A appears as some A_q or A'_q respectively. Note that $N_A \le a' \le 2\beta r^2 k^3 n$, so $2\beta r^2 k^3 n + N'_A - N_A \ge 0$ for any $A \in \binom{[r]}{k}$. Fix an integer C' such that $kr \cdot r!$ divides C' and $2\beta r^2 k^3 n \le C' \le 3\beta r^2 k^3 n$. We choose M_4 to consist of $C' + N'_A - N_A$ properly-distributed copies of K_k in G with index A for each $A \in \binom{[r]}{k}$. Since $\sum_{A \in \binom{[r]}{k}} (N'_A - N_A) = 0$, this will give us $|M_4| = C'\binom{r}{k}$; note that $rk \cdot r!$ then divides $|M_4|$. We also require that these copies are pairwise vertex-disjoint, and vertex-disjoint from $\bigcup_{i \in [3]} M_i$. We may choose M_4 greedily, since by Claim 7.5(i), for any $A \in \binom{[r]}{k}$ there are ωn^k copies of K_k in G with index A, and when choosing each copy we only need to avoid the at most $2\beta r k^3 n + \binom{r}{k} C'$ vertices covered by $\bigcup_{i \in [3]} M_i$ or previously chosen members of M_4. Now, since for any $q \in [a']$ we have $A_q \setminus A'_q = \{j_q\}$ and $A'_q \setminus A_q = \{j'_q\}$, for any $j \in [r]$ we have

$$|V(M_4) \cap W_j| = \sum_{A \in \binom{[r]}{k} : j \in A} (C' + N'_A - N_A) = C'\binom{r-1}{k-1} - a'_j = \frac{k|M_4|}{r} - a'_j.$$

Then $|V_j \cap \bigcup_{i \in [4]} V(M_i)| = k|M_4|/r - a'_j + |\bigcup_{i \in [3]} V(M_i) \cap V_j| = k|\bigcup_{i \in [4]} M_i|/r$ for any $j \in [r]$. Finally, $|\bigcup_{i \in [4]} V(M_i)| \le C'\binom{[r]}{k} + 2\beta r k^3 n \le \beta' n/2$. $\square$

The final K_k-packing is chosen so that the remaining blocks all have size proportional to their row size.

Claim 7.10. (Balancing blocks) *There is a K_k-packing M_5 vertex-disjoint from $\bigcup_{i \in [4]} M_i$, consisting of properly-distributed cliques, such that defining $M = \bigcup_{i \in [5]} M_i$, $X'^i_j := W^i_j \setminus V(M)$, $X'^i = \bigcup_{j \in [r]} X'^i_j$ and $X' = \bigcup_{i \in [s]} X'^i = V(G) \setminus V(M)$, we have $|X'^i_j| = p_i n'$ for any $i \in [s]$ and $j \in [r]$, where $n' = |X'|/kr$ is an integer divisible by $r!$ with $n' \ge n - \zeta n/2$. Thus X'^i_j forms a row-decomposition of $G' := G[X']$ of type p. Furthermore, any vertex $v \in X'^i_j$ has at most $\beta n \le \alpha n'$ non-neighbours in any $X'^{i'}_{j'}$ with $i' \ne i$.*

Proof. For each $i \in [s]$ and $j \in [r]$ we let $W'^i_j = W^i_j \setminus V(\bigcup_{i \in [4]} M_i)$, $W'^i := \bigcup_{j \in [r]} W'^i_j$, $W'_j := \bigcup_{i \in [s]} W'^i_j$, and $W' = \bigcup_{j \in [r]} W'_j$. We may fix an integer D so that $|W'^i| = p_i D$ for any $i \in [s]$, since $|W^i \setminus V(M_1 \cup M_2)|/p_i$ is constant for $i \in [s]$ and each clique in $M_3 \cup M_4$ is properly-distributed. Recall also that $|W'_j|$ is constant for each $j \in [r]$ by choice of M_4. Let

$Q = (q_{ij})$ be the s by r integer matrix whose (i, j) entry is $q_{ij} = |W'^i_j| - p_iD/r$. Then each row of Q sums to zero. Furthermore, since $|W'_j|$ is constant for each $j \in [r]$ we have

$$\sum_{i \in [s]} q_{ij} = \sum_{i \in [s]} (|W'^i_j| - p_iD/r) = |W'_j| - |W'|/r = 0,$$

i.e. each column of Q also sums to zero. We also have $\sum_{i,j} |q_{ij}| \leq \beta'n$ using (A2) and Claim 7.9(ii).

We write $Q = \sum_{A \in \mathcal{Q}} A$, where $\mathcal{Q}$ is a multiset of matrices. Each matrix in $\mathcal{Q}$ is of the form Q^{abcd}, for some $a, c \in [s]$ and $b, d \in [r]$, defined to have (i, j) entry equal to 1 if $(i, j) = (a, b)$ or $(i, j) = (c, d)$, equal to -1 if $(i, j) = (a, d)$ or $(i, j) = (c, b)$, and equal to zero otherwise. To see that such a representation is possible, we repeat the following step. Suppose $q_{ab} > 0$ for some a, b. Since each row and column of Q sum to zero, we may choose c, d such that $q_{ad} < 0$ and $q_{cb} < 0$. Then $Q' := Q - Q^{abcd}$ is an s by k integer matrix in which the entries of each row and column sum to zero. Also, writing $Q' = (q'_{ij})$, we have $\sum_{i,j} |q'_{ij}| \leq \sum_{i,j} |q_{ij}| - 2$. By iterating this process at most $\sum_{i,j} |q_{ij}|/2$ times we obtain the all-zero matrix, whereupon we have expressed Q in the required form with $|\mathcal{Q}| \leq \beta'n$, counting with multiplicity.

Let $\mathcal{P}$ denote the set of all families $\mathcal{A}$ of pairwise vertex-disjoint subsets $A_i \subseteq [r]$ with $|A_i| = p_i$ for $i \in [s]$. To implement a matrix $Q^{abcd} \in \mathcal{Q}$ we fix any two families $\mathcal{A}, \mathcal{A}'$ such that $b \in A_a$ and $d \in A_c$, and $\mathcal{A}'$ is formed from $\mathcal{A}$ by swapping b and d. That is, $A'_i = A_i$ if $i \in [s] \setminus \{a, c\}$, $A'_a = (A_a \setminus \{b\}) \cup \{d\}$ and $A'_c = (A_c \setminus \{d\}) \cup \{b\}$. For each $\mathcal{A} \in \mathcal{P}$ let $Q_\mathcal{A}$ be the number of times it is chosen as $\mathcal{A}$ for some Q^{abcd}, and $Q'_\mathcal{A}$ the number of times it is chosen as $\mathcal{A}'$ for some Q^{abcd}. Fix an integer C'' such that $kr \cdot r!$ divides C'' and $rk\beta'n \leq C'' \leq 2rk\beta'n$. For each $\mathcal{A} \in \mathcal{P}$ let $N_\mathcal{A} = C'' + Q_\mathcal{A} - Q'_\mathcal{A}$, so $N_\mathcal{A} \geq 0$.

Now we greedily choose M_5 to consist of $N_\mathcal{A}$ copies of K' for each $\mathcal{A} \in \mathcal{P}$, each of which will intersect each W^i in precisely those W^i_j such that $j \in A_i$. Then we will have $|M_5| = \sum_{\mathcal{A} \in \mathcal{P}} N_\mathcal{A} = C''|\mathcal{P}| \leq \zeta n/2$. When choosing any copy of K_k we must avoid the vertices of copies of K_k which were previously chosen for M_5, or which lie in $\bigcup_{i \in [4]} M_i$; there are at most $k\zeta n$ such vertices. By Claim 7.5(i), for any $\mathcal{A} \in \mathcal{P}$ there are at least ωn^k properly-distributed copies of K_k in G which intersect each W^i in precisely those W^i_j such that $j \in A_i$, so we can indeed choose M_5 greedily. This defines M, X'^i_j, X'^i, X'_j, X', n', G' as in the statement of the claim. Note that $|M| \leq \zeta n$, and $kr \cdot r!$ divides $|X'|$ by Claims 7.8(ii) and 7.9(iii) and the choice of C''.

Finally, consider the number of vertices used in W^a_b, where $a \in [s]$, $b \in [r]$. If $\mathcal{A} \in \mathcal{P}$ is chosen uniformly at random, then A_a is a uniformly random subset of size p_a, so contains b with probability p_a/r. So if we chose C'' copies of each $\mathcal{A} \in \mathcal{P}$ we would choose p_aN vertices in W^a_b, where $N := C''|\mathcal{P}|/r$. However, since we choose $N_\mathcal{A}$ copies of each $\mathcal{A} \in \mathcal{P}$, we need to adjust by $Q_\mathcal{A} - Q'_\mathcal{A}$. These are chosen so that for each matrix $Q^{abcd} \in \mathcal{Q}$ we choose one more vertex in each of W^a_b and W^c_d, and one fewer vertex in each of W^a_d and W^c_b. Since $Q = \sum_{A \in \mathcal{Q}} A$ we thus use $p_aN + q_{ab}$ vertices in W^a_b. This gives $|X'^a_b| = |W'^a_b| - p_aN - q_{ab} = p_a(D/r - N)$. Writing $n' = D/r - N$, we have $n' = |X'|/kr$, so n' is an integer divisible by $r!$. Note also that $n - n' \leq |V(M)|/kr \leq \zeta n/2$. Lastly, by choice of M_3 every bad vertex is covered by M, so any vertex $v \in X'^i_j$ has at most $\beta n \leq \alpha n'$ non-neighbours in any $X'^{i'}_{j'}$ with $i' \neq i$ and $j' \neq j$ by (A3).   □

After deleting M as in Claim 7.10, we obtain G' with an s-row-decomposition X' that satisfies conclusion (i) of Lemma 7.2. To complete the proof, we need to satisfy conclusion (ii), by finding a balanced perfect K_{p_i}-packing in row i for each $i \in [s]$.

Observe first that it is straightforward to find such a packing in row i if $p_i \neq 2$. Indeed, for each $i \in [s]$ there is a trivial balanced perfect K_1-packing in $G[X'^i]$, namely $\{\{v\} : v \in X'^i\}$; this gives the desired packing if $p_i = 1$. Now suppose that $p_i \geq 3$ for some $i \in [s]$, and recall that $G_1[X^i]$ was not d'-splittable (this is true of any row of G_1). For any $j \in [r]$ we have $|X_j^i \triangle W_j^i| \leq 2\beta n$ by (A2) and $|W_j^i \triangle X'^i_j| = p_i n - p_i n' \leq (\zeta/2)p_i n$ by choice of n', so $|X_j^i \triangle X'^i_j| \leq \zeta p_i n$. Proposition 6.1 then implies that $G[X'^i]$ is not d''-splittable, so $G[X'^i]$ contains a balanced perfect K_{p_i}-packing by Lemma 6.2.

It follows that if there is no row i with $p_i = 2$, then we have conclusion (ii), so the proof is complete in this case. We may therefore assume that at least one row satisfies $p_i = 2$, and consider two cases according to whether or not any other row has $p_i \geq 2$.

Case 1: There is exactly one row i with $p_i \geq 2$.

Case 1.1: G has the extremal row structure. This means that G has one pair-complete row i^*, and $p_i = 1$ for any $i \neq i^*$. There is a trivial balanced perfect K_1-packing in $G[X'^i]$ for any $i \neq i^*$, so it remains only to find a balanced perfect matching in $G[X'^{i^*}]$. Since row i^* is pair-complete, we chose sets $S_j^{i^*}$ for $j \in [r]$ when forming the sets W_j^i. Let $S'^{i^*}_j := S_j^{i^*} \cap X'^i_j = S_j^{i^*} \setminus V(M)$ for each $j \in [r]$. Then

$$(1+\beta)n \overset{(A2)}{\geq} |W_j^{i^*}| - |Y_j^{i^*} \setminus S_j^{i^*}| \geq |S_j^{i^*}| \geq |S'^{i^*}_j| \geq |S_j^{i^*}| - |M| \overset{(A2)}{\geq} (1 - \beta/2)n - \zeta n,$$

and so $|S'^{i^*}_j| = (1 \pm 2\zeta)n'$ for each $j \in [r]$. Recall also that in this case we required that $|S^{i^*} \setminus V(M_1)|$ was even when choosing M_1, and M_2 was empty. Furthermore, any copy K' of K_k in $M_3 \cup M_4 \cup M_5$ was chosen to be properly-distributed, so in particular $|K' \cap S^{i^*}|$ is even. We conclude that $|S^{i^*} \setminus V(M)|$ is even. Since every bad vertex of G was covered by $M_3 \subseteq M$, by (A3) for any $j' \neq j$ any vertex in $S'^{i^*}_j$ has at most $\beta n \leq 2\zeta n'$ non-neighbours in $S'^{i^*}_{j'}$, and any vertex in $X'^{i^*}_j \setminus S'^{i^*}_j$ has at most $\beta n \leq 2\zeta n'$ non-neighbours in $X'^{i^*}_{j'} \setminus S'^{i^*}_{j'}$ by (A3). By Lemma 6.4 (with 2ζ in place of ζ) we conclude that $G[X'^{i^*}]$ contains a balanced perfect matching, completing the proof in this case.

Case 1.2: $p_\ell = 2$ for some ℓ, $p_{i'} = 1$ for any $i' \neq \ell$, but row ℓ is not pair-complete. Recall that this means that $G_1[X^\ell]$ was not d'-pair-complete. Moreover, just as for rows with $p_i \geq 3$, we have that $G_1[X^\ell]$ was not d'-splittable and that $|X_j^\ell \triangle X'^\ell_j| \leq \zeta(2n)$ for any $j \in [r]$. Proposition 6.1 therefore implies that $G[X'^\ell]$ is neither d''-splittable nor d''-pair-complete, so $G[X'^\ell]$ contains a ν-balanced perfect matching by Corollary 5.7. Then Proposition 6.5 implies that there exists an integer $D \leq 2\nu n'$ and a K_k-packing M^* in G such that $r!$ divides $n'' := n' - D$, M^* covers $p_i D$ vertices in X'^i_j for any $i \in [s]$ and $j \in [r]$, and $G[X'^\ell \setminus V(M^*)]$ contains a balanced perfect matching. Note that since $n' \geq n - \zeta n/2$ we have $n'' \geq n^+/k - \zeta n$. We add the copies of K_k in M^* to M, and let X''^i_j, X''^i, X'', G'' be obtained from X'^i_j, X'^i, X', G' by deleting the vertices covered by M^*. This leaves an s-row-decomposition of G'' into blocks X''^i_j of size $p_i n''$, in which $G[X''^\ell]$ contains a balanced perfect matching. As before $G[X''^i]$ contains a trivial balanced K_1-packing for every $i \neq \ell$. Finally, any vertex $v \in X'^i_j$ has lost at most $k|M^*| \leq \beta n$ neighbours in $X'^{i'}_{j'}$ for any $i \neq i'$ and $j \neq j'$, so still has at least $p_{i'} n' - 2\beta n \geq p_{i'} n'' - \alpha n''$. So the enlarged matching M,

restricted blocks X''^i_j and G'' satisfy (i) and (ii) with n'' in place of n', which completes the proof in this case.

Case 2: There are at least two rows i with $p_i \geq 2$. In this case we modify the cliques in M_2 and the blocks X'^i_j so that after these modifications $G[X'^i]$ contains a balanced perfect K_{p_i}-packing for each $i \in [s]$. We proceed through each $i \in [s]$ in turn. When considering any $i \in [s]$ we leave all blocks $X'^{i'}_j$ with $i' \neq i$ unaltered. We observed previously that if $p_i \neq 2$, then $G[X'^i]$ contains a balanced perfect K_{p_i}-packing, so no modifications are required to achieve the desired packings in such rows.

This leaves only those rows $i \in [s]$ with $p_i = 2$ to consider. Suppose first that row i is pair-complete. As in Case 1.1 we let $S'^i_j = S^i_j \cap X'^i_j$ for each $j \in [r]$, which gives $|S'^i_j| = (1 \pm 2\zeta)n'$ for each $j \in [r]$, and for any $j' \neq j$, any vertex in S'^i_j has at most $2\zeta n'$ non-neighbours in $S'^i_{j'}$, and any vertex in $X'^i_j \setminus S'^i_j$ has at most $2\zeta n'$ non-neighbours in $X'^i_{j'} \setminus S'^i_{j'}$. If $|S'^i|$ is even then $G[X'^i_j]$ contains a balanced perfect matching by Lemma 6.4. So we may suppose that $|S'^i|$ is odd. Fix any i' with $i' \neq i$ and $p_{i'} \geq 2$. We choose any $i'i$-distributed copy K' of K_k in M_2, let x be the vertex of K' in X'^i, and let j be such that $x \in X'^i_j$. By Claim 7.7(i), every vertex in K' is good, so at least $2n' - k\beta n$ vertices $y \in X'^i_j$ are adjacent to every member of $V(K') \setminus \{x\}$. So we may choose a vertex $y \in X'^i_j \cap \bigcap_{v \in V(K') \setminus \{x\}} N(v)$ such that $|\{x, y\} \cap S^i_j|$ is odd. We replace K' in M by the copy of K_k in G induced by $\{y\} \cup V(K') \setminus \{x\}$, and replace y by x in X'^i_j and G'. We also delete y from S'^i_j if $y \in S'^i_j$, and add x to S'^i_j if $x \in S^i_j$. So $|S'^i_j|$ is now even. Since x is good, we may now apply Lemma 6.4 as in Case 1.1 to find a balanced perfect matching in $G[X'^i]$.

Finally suppose that $p_i = 2$ and row i is not pair-complete, that is $G_1[X^i]$ was not d'-pair-complete. Since $G_1[X^i]$ was also not d'-splittable, as before Proposition 6.1 implies that $G[X'^i]$ is neither d''-splittable nor d''-pair-complete. Fix any i' with $i' \neq i$ and $p_{i'} \geq 2$. Recall that M_2 contains $\lceil \eta n \rceil$ copies K' of K_k in $G[Y]$ which are $i'i$-distributed. We can fix $q \in [r]$ so that at least $\eta n/r$ such K' have exactly one vertex in Y^i_q. We assign arbitrarily $\eta n/r$ such K' to each ordered triple (i_1, i_2, i_3) of distinct elements of $[r] \setminus \{q\}$, so that at least $\eta n/r^4$ of the K' are assigned to each triple. Now, fix any triple (j_1, j_2, j_3) and any K' which was assigned to it. Let x be the vertex of K' in Y^i_q, and consider paths $xx_1x_2x_3y$ of length 4 in G with $x_\ell \in X'^i_{j_\ell}$ for $\ell \in [3]$ and $y \in X'^i_q \cap \bigcap_{v \in V(K') \setminus \{x\}} N(v)$. Since every vertex of $K' \setminus \{x\}$ is good and does not lie in W^i, at most $k\beta n$ vertices $y \in X'^i_q$ fail to be adjacent to all of $V(K') \setminus \{x\}$. Choosing x_1, x_2, x_3 and y in turn, recalling (†) and $n' \geq n - \zeta n/2$, there are at least $2n' - n \geq (1 - \zeta)n'$ choices for each x_ℓ, and at least $2n' - n - (k-1)\beta n \geq (1 - \zeta)n'$ choices for y. We obtain at least $n^4/2$ such paths, and so we may fix some $y = y(x)$ which lies in at least $n^3/5$ such paths. For each of these $n^3/5$ paths $xx_1x_2x_3y$ we add a 'fake' edge between y and x_1. Then allowing the use of fake edges, y lies in at least $n^3/5$ 4-cycles $x_1x_2x_3y$ of length 4 in G with $x_\ell \in X'^i_{j_\ell}$ for $\ell \in [3]$. We introduce fake edges in this manner for every K' in M_2 which was assigned to the triple (j_1, j_2, j_3), for every ordered triple (j_1, j_2, j_3) of distinct elements of $[r] \setminus \{q\}$. Let G^* be the graph formed from $G[X'^i]$ by the addition of fake edges. Then by construction, for any triple (j_1, j_2, j_3) there are at least $(\eta n/r^4)(n^3/5) \geq \nu n^4$ 4-cycles $yx_1x_2x_3$ in G^* with $y \in X'^i_q$ and $x_\ell \in X'^i_{j_\ell}$ for $\ell \in [3]$. Furthermore, G^* contains a spanning subgraph $G[X'^i]$ which is neither d''-splittable nor d''-pair-complete, and has $\delta^*(G[X'^i]) \geq 2n' - n \geq n' - \zeta n$. Then G^* contains a balanced perfect matching M^* by

Lemma 6.2 (with q in place of 1). Of course, M^* may contain fake edges. However, any fake edge in M^* is of the form $y(x)x_1$, where x_1 is a neighbour of x, and x lies in some K' in M_2. Since $y(x)$ is uniquely determined by x and M^* is a matching, at most one edge in M^* has the form $y(x)x_1$ for each x. Furthermore, by choice of $y = y(x)$, $\{y\} \cup V(K') \setminus \{x\}$ induces a copy K'_y of K_k in G, and xx_1 is an edge. So we may replace K' in M_2 by K'_y, replace y in X'^i_q by x (note that x is good), and the fake edge yx_1 in M^* by the edge xx_1 of G. We carry out these substitutions for every fake edge in M^*, at the end of which M^* is a perfect matching in $G[X'^i]$, which is balanced since each edge was replaced with another of the same index.

When considering row i we only replace cliques of M_2 that are $i'i$-distributed for some $i' \neq i$. These cliques uniquely determine i, so do not affect the replacements for other rows. We may therefore proceed through every $i \in [s]$ in this manner. After doing so, the modified blocks X'^i_j are such that $G[X'^i]$ contains a balanced perfect K_{p_i}-packing for any $i \in [s]$, i.e. this row-decomposition of the modified G' satisfies condition (ii) of the lemma. Note that we still have $|X'^i_j| = p_i n'$ for any $i \in [s]$. Since we only replaced good vertices with good vertices, every vertex in any modified block X'^i_j is good, and so condition (i) of the lemma holds as in Claim 7.10. This completes the proof of Lemma 7.2. $\qquad \blacksquare$

8. Completing the proof of Theorem 1.1

In this section we combine the results of previous sections to prove Theorem 1.1. For this we also use a theorem of Daykin and Häggkvist [3], which gives a minimum vertex degree condition which is sufficient to ensure the existence of a perfect matching in a k-partite k-graph whose vertex classes each have size n.

Theorem 8.1 ([3]). *Suppose that G is a k-partite k-graph whose vertex classes each have size n, in which every vertex lies in at least $\frac{k-1}{k}n^{k-1}$ edges. Then G contains a perfect matching.*

We can now give the proof of Theorem 1.1, as outlined in Section 3, which we first restate.

Theorem 1.1. *For any $r \geq k$ there exists n_0 such that for any $n \geq n_0$ with $k \mid rn$ the following statement holds. Let G be an r-partite graph whose vertex classes each have size n such that $\delta^*(G) \geq (k-1)n/k$. Then G contains a perfect K_k-packing, unless rn/k is odd, $k \mid n$, and $G \cong \Gamma_{n,r,k}$.*

Proof. First suppose that $k = 2$, so a perfect K_k-packing is a perfect matching. If $r = 2$ then G is a bipartite graph with minimum degree at least $n/2$, so has a perfect matching. For $r \geq 3$ the result follows from Tutte's theorem, which states that a graph G on the vertex set V contains a perfect matching if and only if for any $U \subseteq V$ the number of odd components (i.e. connected components of odd size) in $G[V \setminus U]$ is at most $|U|$. To see that this implies the theorem for $k = 2$, suppose for a contradiction that there is some $U \subseteq V$ for which $G[V \setminus U]$ has more than $|U|$ odd components. Clearly $|U| < |V|/2 = rn/2$. So by averaging U has fewer than $n/2$ vertices in some V_j. Since $\delta^*(G) \geq n/2$, every $v \in V \setminus V_j$ has a neighbour in $V_j \setminus U$, so $G[V \setminus U]$ has at most $|V_j \setminus U| \leq n$ components. So we must have $|U| < n$. But then U must have fewer than $n/r \leq n/3$ vertices in some V_j, so any $v \in V \setminus V_j$ has more than $n/6$ neighbours in V_j. It follows that $G[V \setminus U]$ has at most 5 components, so $|U| < 5$. So any $v \in V \setminus V_j$ actually has more than $n/2 - 5 > n/3$ neighbours in V_j, so $G[V \setminus U]$ has at most 2 components. So $|U| \leq 1$. If $|U| = 1$ then $|V \setminus U|$ is odd, so $G[V \setminus U]$

cannot have 2 odd components. The only remaining possibility is that $|U| = 0$ and G has 2 odd components. Let C^1 and C^2 be these components, and for each $i \in [2]$ and $j \in [r]$ let V_j^i be the vertices of V_j covered by C^i. Then $|V_j^i| \geq \delta^*(G) \geq n/2$, so we deduce that $|V_j^i| = n/2$ and $G[V_j^i, V_{j'}^i]$ is a complete bipartite graph for any $i \in [2]$ and $j \neq j'$. So $|C^1| = rn/2$ is odd, 2 divides n and G is isomorphic to $\Gamma_{n,r,2}$, contradicting our assumption.

We may therefore assume that $k \geq 3$. If $r = k = 3$ then Theorem 1.1 holds by the result of [9]. So we may assume that $r > 3$. We introduce a new constant d with $1/n \ll d \ll 1/r$. Since $r > 3$ and $r \geq k \geq 3$ we may apply Lemma 7.2 (with n and d in place of n^+ and α) to delete the vertices of a collection of pairwise vertex-disjoint copies of K_k from G. Letting V' be the set of undeleted vertices, we obtain $G' = G[V']$ and an s-row-decomposition of G' into blocks $X_j'^i$ of size $p_i n'$ for $i \in [s]$ and $j \in [r]$, for some $s \in [k]$ and $p_i \in [k]$ with $\sum_{i \in [s]} p_i = k$, such that $r! \mid n'$, $n' \geq n/k - dn$ and

(i) for each $i, i' \in [s]$ with $i \neq i'$ and $j, j' \in [r]$ with $j \neq j'$, any vertex $v \in X_j'^i$ has at least $p_{i'} n' - dn'$ neighbours in $X_{j'}'^{i'}$, and

(ii) for every $i \in [s]$ the row $G[X'^i]$ contains a balanced perfect K_{p_i}-packing M^i.

Note that we must have $|M^i| = rn'$ for any $i \in [s]$.

Now we implement step (iii) of the proof outline, by constructing auxiliary hypergraphs, perfect matchings of which describe how to glue together the perfect K_{p_i}-packings in the rows into a perfect K_k-packing of G. We partition $[k]$ arbitrarily into sets A_i with $|A_i| = p_i$ for $i \in [s]$. Let Σ denote the set of all injective functions $\sigma : [k] \to [r]$. For each $i \in [s]$ we partition M^i into sets E_σ^i of size $N := \frac{rn'(r-k)!}{r!}$ for $\sigma \in \Sigma$, so that each member of E_σ^i has index $\sigma(A_i)$. To see that this is possible, fix any $B \in \binom{[r]}{p_i}$. Since M^i is balanced, $rn'/\binom{r}{p_i}$ members of M^i have index B. Note that there are $\frac{p_i!(r-p_i)!}{(r-k)!}$ members of Σ with $\sigma(A_i) = B$. Since $\frac{p_i!(r-p_i)!}{(r-k)!} \cdot N = rn'/\binom{r}{p_i}$ we may choose the sets E_σ^i as required. For every $\sigma \in \Sigma$, we form an auxiliary s-partite s-graph H_σ with vertex classes E_σ^i for $i \in [s]$, where a set $\{e_1, e_2, \ldots, e_s\}$ with $e_i \in E_\sigma^i$ for each $i \in [s]$ is an edge of H_σ if and only if $xy \in G$ for any $i \neq j$, $x \in e_i$ and $y \in e_j$. Thus H_σ has N vertices in each vertex class, and $e_1 e_2 \ldots e_s$ is an edge of H_σ if and only if $\bigcup_{j \in [s]} V(e_j)$ induces a copy of K_k in G.

Next we show that each H_σ has high minimum degree. Fix $\sigma \in \Sigma$ and $i \in [s]$. Then for any $e_i \in E_\sigma^i$, e_i is a copy of K_{p_i} in $G[X'^i]$ with index $\sigma(A_i)$, and so by (i) each vertex $x \in V(e_i)$ has at most dn' non-neighbours in each $X_j'^\ell$ with $\ell \neq i$ and $j \notin \sigma(A_i)$. So at most $p_i dn'$ vertices of $X_j'^\ell$ are not neighbours of some vertex of e_i. Now we can estimate the number of $(s-1)$-tuples $(e_j \in E_\sigma^j : j \in [s] \setminus \{i\})$ so that $\{e_1, \ldots, e_s\}$ is not an edge of H_σ. There are fewer than k choices for $j \in [s] \setminus \{i\}$, at most $p_i dn'$ elements $e_j \in E_\sigma^j$ that contain a non-neighbour of some vertex of e_i, and at most N^{s-2} choices for $e_{j'} \in E_\sigma^{j'}$, $j' \in [s] \setminus \{i, j\}$. So the number of edges of H_σ containing e_i is at least $N^{s-1} - kp_i dn' N^{s-2} \geq \frac{k-1}{k} N^{s-1}$. Since e_i was arbitrary, Lemma 8.1 gives a perfect matching in each H_σ. This corresponds to a perfect K_k-packing in G covering the vertices of $\bigcup_{i \in [s]} V(E_\sigma^i)$, where $V(E_\sigma^i)$ denotes the vertices covered by members of E_σ^i. Combining these perfect K_k-packings, and adding the pairwise vertex-disjoint copies of K_k deleted in forming G', we obtain a perfect K_k-packing in G. $\square$

9. Concluding remarks

By examining the proof, one can obtain a partial stability result, i.e. some approximate structure for any r-partite graph G with vertex classes each of size n, where $k \mid rn$, such that $\delta^*(G) \geq (k-1)n/k - o(n)$, but G does not contain a perfect K_k-packing. To do this, note that under this weaker minimum degree assumption, the n in (†) must replaced by $n + o(n)$. We now say that a block X_j^i is bad with respect to v if v has more than $n/2 + o(n)$ non-neighbours in X_j^i, so it is still true that at most one block in each column is bad with respect to a given vertex. Then each of our applications of (†) proceeds as before, except for in Claim 7.6, where we used the exact statement of (†) (i.e. the exact minimum degree hypothesis). This was needed to choose a matching E^i in $G[W^i]$ of size a_i, each of whose edges contains a good vertex, for each $i \in I^+$ with $p_i = 1$. If we can choose such matchings E^i then the rest of the proof to give a perfect K_k-packing still works under the assumption $\delta^*(G) \geq (k-1)n/k - o(n)$. So we can assume that there is some $i \in I^+$ with $p_i = 1$ for which no such matching exists. Since the number of bad vertices and a_i are $o(n)$, it follows that W^i is a subset of size $rn/k + o(n)$ with $o(n^2)$ edges, i.e. we have a sparse set of about $1/k$-proportion of the vertices. On the other hand, this is essentially all that can be said about the structure of G, as any such G with an independent set of size $rn/k + 1$ cannot have a perfect K_k-packing.

References

[1] P. Catlin, On the Hajnal-Szemerédi theorem on disjoint cliques, *Utilitas Math.* **17** (1980), 163–177.

[2] B. Csaba and M. Mydlarz, Approximate Multipartite Version of the Hajnal–Szemerédi Theorem, *J. Combinatorial Theory, Series B* **102** (2012), 395–410.

[3] D.E. Daykin and R. Häggkvist, Degrees giving independent edges in a hypergraph, *Bull. Austral. Math. Soc.* **23** (1981), 103–109.

[4] E. Fischer, Variants of the Hajnal-Szemerédi Theorem, *J. Graph Theory* **31** (1999), 275–282.

[5] S. Hakimi, On the realizability of a set of integers as degrees of the vertices of a graph, *SIAM J. Appl. Math.* **10** (1962), 496–506.

[6] P. Keevash and R. Mycroft, A geometric theory for hypergraph matching, *Mem. Amer. Math. Soc.* **233** (2014), number 1098.

[7] D. Kühn and D. Osthus, Embedding large subgraphs into dense graphs, *Surveys in Combinatorics 2009*, Cambridge University Press, 2009, 137–167.

[8] A. Lo and K. Markström, A multipartite version of the Hajnal-Szemerédi theorem for graphs and hypergraphs, *Comb. Probab. Comput.* **22** (2013), 97–111.

[9] C. Magyar and R. Martin, Tripartite version of the Corrádi-Hajnal theorem, *Disc. Math.* **254** (2002), 289–308.

[10] R. Martin and E. Szemerédi, Quadripartite version of the Hajnal-Szemerédi theorem, *Disc. Math.* **308** (2008), 4337–4360.

Peter Keevash   Richard Mycroft
Mathematical Institute   School of Mathematics
University of Oxford   University of Birmingham
Oxford, UK   Birmingham, UK

E-mail addresses: `Peter.Keevash@maths.ox.ac.uk, r.mycroft@bham.ac.uk`