Advances in Mathematics 269 (2015) 265–334

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Polynomial-time perfect matchings in dense hypergraphs ☆

Peter Keevash [a,*], Fiachra Knox [b], Richard Mycroft [c]

[a] *Mathematical Institute, University of Oxford, Oxford, UK*
[b] *Department of Mathematics, Simon Fraser University, Burnaby, BC, Canada*
[c] *School of Mathematics, University of Birmingham, Birmingham, UK*

ARTICLE INFO

Article history:
Received 9 July 2013
Accepted 17 October 2014
Available online 4 November 2014
Communicated by Benjamin Sudakov

Keywords:
Hypergraphs
Algorithms
Matchings

ABSTRACT

Let H be a k-graph on n vertices, with minimum codegree at least $n/k + cn$ for some fixed $c > 0$. In this paper we construct a polynomial-time algorithm which finds either a perfect matching in H or a certificate that none exists. This essentially solves a problem of Karpiński, Ruciński and Szymańska; Szymańska previously showed that this problem is NP-hard for a minimum codegree of $n/k - cn$. Our algorithm relies on a theoretical result of independent interest, in which we characterise any such hypergraph with no perfect matching using a family of lattice-based constructions.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

The question of whether a given k-uniform hypergraph (or k-graph) H contains a perfect matching (i.e. a partition of the vertex set into edges), while simple to state, is one of the key questions of combinatorics. In the graph case $k = 2$, Tutte's Theorem [32] gives necessary and sufficient conditions for H to contain a perfect matching, and Edmonds'

☆ Research supported in part by ERC grant 239696 and EPSRC grant EP/G056730/1.
* Corresponding author.
E-mail addresses: keevash@maths.ox.ac.uk (P. Keevash), fjknox@sfu.ca (F. Knox), r.mycroft@bham.ac.uk (R. Mycroft).

http://dx.doi.org/10.1016/j.aim.2014.10.009
0001-8708/© 2014 Elsevier Inc. All rights reserved.

Algorithm [9] finds such a matching in polynomial time. However, for $k \geq 3$ this problem was one of Karp's celebrated 21 NP-complete problems [14]. Results for perfect matchings in hypergraphs have many potential practical applications; one example which has garnered interest in recent years is the 'Santa Claus' allocation problem (see [3]). Since the general problem is intractable provided P $\neq$ NP, it is natural to seek conditions on H which render the problem tractable or even guarantee that a perfect matching exists. In recent years a substantial amount of progress has been made in this direction. One well-studied class of such conditions are minimum degree conditions. This paper provides an algorithm that essentially eliminates the hardness gap between the sparse and dense cases for the most-studied of these conditions.

1.1. Minimum degree conditions

Suppose that H has n vertices and that k divides n (we assume this throughout, since it is a necessary condition for H to contain a perfect matching). In the graph case, a simple argument shows that a minimum degree of $n/2$ guarantees a perfect matching. Indeed, Dirac's Theorem [8] states that this condition even guarantees that H contains a Hamilton cycle. For $k \geq 3$, there are several natural definitions of the minimum degree of H. Indeed, for any set $A \subseteq V(H)$, the *degree* $d(A) = d_H(A)$ of A is the number of edges of H containing A. Then for any $1 \leq \ell \leq k-1$, the *minimum ℓ-degree* $\delta_\ell(H)$ of H is the minimum of $d(A)$ over all subsets $A \subseteq V(H)$ of size ℓ. Two cases have received particular attention: the minimum 1-degree $\delta_1(H)$ is also known as the *minimum vertex degree* of H, and the minimum $(k-1)$-degree $\delta_{k-1}(H)$ as the *minimum codegree* of H.

For sufficiently large n, Rödl, Ruciński and Szemerédi [27] determined the minimum codegree which guarantees a perfect matching in H to be exactly $n/2 - k + c$, where $c \in \{1.5, 2, 2.5, 3\}$ is an explicitly given function of n and k. They also showed that the condition $\delta_{k-1}(H) \geq n/k + O(\log n)$ is sufficient to guarantee a matching covering all but k vertices of H, i.e. one edge away from a perfect matching; their conjecture that $\delta_{k-1}(H) \geq n/k$ suffices for this was recently proved by Han [10]. This provides a sharp contrast to the graph case, where a minimum degree of $\delta(G) \geq n/2 - \varepsilon n$ only guarantees the existence of a matching covering at least $n - 2\varepsilon n$ vertices. There is a large literature on minimum degree conditions for perfect matchings in hypergraphs, see e.g. [1,2,6,7,12, 15,17–25,27,29–31] and the survey by Rödl and Ruciński [26] for details.

Let $\mathbf{PM}(k, \delta)$ be the decision problem of determining whether a k-graph H with $\delta_{k-1}(H) \geq \delta n$ contains a perfect matching. Given the result of [27], a natural question to ask is the following: For which values of δ can $\mathbf{PM}(k, \delta)$ be decided in polynomial time? This holds for $\mathbf{PM}(k, 1/2)$ by the main result of [27]. On the other hand, $\mathbf{PM}(k, 0)$ includes no degree restriction on H at all, so is NP-complete by the result of Karp [14]. Szymańska [28,29] proved that for $\delta < 1/k$ the problem $\mathbf{PM}(k, 0)$ admits a polynomial-time reduction to $\mathbf{PM}(k, \delta)$ and hence $\mathbf{PM}(k, \delta)$ is also NP-complete, while Karpiński, Ruciński and Szymańska [15] showed that there exists $\varepsilon > 0$ such that $\mathbf{PM}(k, 1/2 - \varepsilon)$ is in P. This left a hardness gap for $\mathbf{PM}(k, \delta)$ when $\delta \in [1/k, 1/2 - \varepsilon)$.

In this paper we provide an algorithm which eliminates this hardness gap almost entirely. Moreover, it not only solves the decision problem, but also provides a perfect matching or a certificate that none exists.

Theorem 1.1. *Fix $k \geq 3$ and $\gamma > 0$. Then there is an algorithm with running time $O(n^{3k^2-7k+1})$, which given any k-graph H on n vertices with $\delta_{k-1}(H) \geq (1/k+\gamma)n$, finds either a perfect matching or a certificate that no perfect matching exists.*

A preliminary version of this algorithm (with a slower running time) appeared as an extended abstract [16].

1.2. Lattices and divisibility barriers

Theorem 1.1 relies on a result of Keevash and Mycroft [17] giving fairly general sufficient conditions which ensure a perfect matching in a k-graph. In this context, their result essentially states that if H is a k-graph on n vertices, and $\delta_{k-1}(H) \geq n/k + o(n)$, then H either contains a perfect matching or is close to one of a family of lattice-based constructions termed 'divisibility barriers'. These constructions play a key role in this paper, so we now describe them in some detail.

The simplest example of a divisibility barrier is the following construction, given as one of the two extremal examples in [27].

Construction 1.2. Let A and B be disjoint sets such that $|A|$ is odd and $|A \cup B| = n$, and let H be the k-graph on $A \cup B$ whose edges are all k-sets which intersect A in an even number of vertices.

We consider a partition to include an implicit order on its parts. To describe divisibility barriers in general, we make the following definition.

Definition 1.3. Let H be a k-graph and let $\mathcal{P}$ be a partition of $V(H)$ into d parts. Then the *index vector* $\mathbf{i}_{\mathcal{P}}(S) \in \mathbb{Z}^d$ of a subset $S \subseteq V(H)$ with respect to $\mathcal{P}$ is the vector whose coordinates are the sizes of the intersections of S with each part of $\mathcal{P}$, i.e. $\mathbf{i}_{\mathcal{P}}(S)_X = |S \cap X|$ for $X \in \mathcal{P}$. Further,

(i) $I_{\mathcal{P}}(H)$ denotes the set of index vectors $\mathbf{i}_{\mathcal{P}}(e)$ of edges $e \in H$, and
(ii) $L_{\mathcal{P}}(H)$ denotes the lattice (i.e. additive subgroup) in $\mathbb{Z}^d$ generated by $I_{\mathcal{P}}(H)$.

A *divisibility barrier* is a k-graph H which admits a partition $\mathcal{P}$ of its vertex set V such that $\mathbf{i}_{\mathcal{P}}(V) \notin L_{\mathcal{P}}(H)$; the next proposition shows that such an H contains no perfect matching. To see that this generalises Construction 1.2, let $\mathcal{P}$ be the partition into parts A and B; then $L_{\mathcal{P}}(H)$ is the lattice of vectors (x, y) in $\mathbb{Z}^2$ for which x is even and k divides $x + y$, and $|A|$ being odd implies that $\mathbf{i}_{\mathcal{P}}(V) \notin L_{\mathcal{P}}(H)$.

Proposition 1.4. *Let H be a k-graph with vertex set V. If there is a partition $\mathcal{P}$ of V with $\mathbf{i}_{\mathcal{P}}(V) \notin L_{\mathcal{P}}(H)$ then H does not contain a perfect matching.*

Proof. Suppose M is a matching in H. Then $\mathbf{i}_{\mathcal{P}}(V(M)) = \sum_{e\in M} \mathbf{i}_{\mathcal{P}}(e) \in L_{\mathcal{P}}(H)$. But $\mathbf{i}_{\mathcal{P}}(V) \notin L_{\mathcal{P}}(H)$, so $V(M) \neq V$, i.e. M is not perfect. □

A special case of the main theoretical result of this paper is the following theorem, which states that the converse of Proposition 1.4 holds for sufficiently large 3-graphs as in Theorem 1.1. Thus we obtain an essentially best-possible strong stability 'Andrasfai–Erdős–Sos analogue' for the result of Rödl, Ruciński and Szemerédi [27] in the case $k = 3$.

Theorem 1.5. *For any $\gamma > 0$ there exists $n_0 = n_0(\gamma)$ such that the following statement holds. Let H be a 3-graph on $n \geq n_0$ vertices, such that 3 divides n and $\delta_2(H) \geq (1/3 + \gamma)n$, and suppose that H does not contain a perfect matching. Then there is a subset $A \subseteq V(H)$ such that $|A|$ is odd but every edge of H intersects A in an even number of vertices.*

Theorem 1.5 can be used to decide $\mathbf{PM}(3, 1/3 + \gamma)$, as the existence of a subset A as in the theorem can be checked using (simpler versions of) the algorithms in Section 2. However, the case $k = 3$ is particularly simple because there is only one maximal divisibility barrier; for $k \geq 4$, the next construction shows that the converse of Proposition 1.4 does not hold for general k-graphs as in Theorem 1.1.

Construction 1.6. Let A, B and C be disjoint sets of vertices with $|A \cup B \cup C| = n$, $|A|, |B|, |C| = n/3 \pm 2$ and $|A| = |B| + 2$. Fix some vertex $x \in A$, and let H be the k-graph with vertex set $A \cup B \cup C$ whose edges are

(1) any k-set e with $|e \cap A| = |e \cap B|$ modulo 3, and
(2) any k-set $(x, z_1, \ldots, z_{k-1})$ with $z_1, \ldots, z_{k-1}$ in C.

Construction 1.6 satisfies $\delta_{k-1}(H) \geq n/3 - k - 1$, so if $k \geq 4$ then H meets the degree condition of Theorem 1.1. Moreover, for any partition $\mathcal{P}$ of $V(H)$ we have $\mathbf{i}_{\mathcal{P}}(V(H)) \in L_{\mathcal{P}}(H)$. To see this, suppose on the contrary that it is false, and fix a counterexample $\mathcal{P}$ with as few parts as possible. Then $L_{\mathcal{P}}(H)$ cannot contain any transferral (see Definition 1.8), as otherwise we could merge the corresponding parts of $\mathcal{P}$ to obtain a counterexample with fewer parts. It follows that A, B and C must each be contained within some part of $\mathcal{P}$. We consider the case $\mathcal{P} = (A, B, C)$ and omit the easy cases where $\mathcal{P}$ has two parts (clearly the partition with one part is not a counterexample, as $k \mid n$). Write $\mathbf{i}_{\mathcal{P}}(V(H)) = (x, y, z)$, where $x + y + z = n$. Note that $(0, 0, k)$, $(1, 1, k - 2)$ and $(1, 0, k - 1)$ are all in $I_{\mathcal{P}}(H)$. Then $(x, y, z) = y(1, 1, k - 2) + (x - y)(1, 0, k - 1) + (n/k - x)(0, 0, k) \in L_{\mathcal{P}}(H)$, so $\mathcal{P}$ is not a counterexample.

However, H does not contain a perfect matching. To see this, let M be a matching in H, and note that any edge $e \in M$ has $|e \cap A| = |e \cap B|$ modulo 3, except for at most one edge of M which has $|e \cap A| = |e \cap B| + 1$ modulo 3. Then $\mathbf{i}_{\mathcal{P}}(V(M))_1 - \mathbf{i}_{\mathcal{P}}(V(M))_2 \in \{0, 1\}$ modulo 3, whereas $\mathbf{i}_{\mathcal{P}}(V(H))_1 - \mathbf{i}_{\mathcal{P}}(V(H))_2 = |A| - |B| = 2$, so $V(M) \neq V(H)$, that is, M is not perfect.

1.3. Approximate divisibility barriers

Our starting point will be (a special case of) a result of Keevash and Mycroft [17] on approximate divisibility barriers. First we introduce a less restrictive degree assumption, which follows from the assumption in Theorem 1.1 when $\gamma > 0$ is small. (The reason for doing so will become clear at the end of Section 2.) We will use the following setup throughout the paper.

Setup 1.7. Suppose that $k \geq 3$, that $1/n_0 \ll \varepsilon \ll \gamma \ll 1/k$ and that $n \geq n_0$ satisfies $k \mid n$.

Let H be a k-graph on n vertices such that

(deg) $\delta_1(H) \geq \gamma n^{k-1}$, and
(codeg) at most εn^{k-1} $(k-1)$-sets $A \subseteq V(H)$ have $d_H(A) < (1/k + \gamma)n$.

A result from [17] (stated here as Theorem 4.6), combined with Lemma 7.3, implies that under Setup 1.7, if H does not contain a perfect matching then we can delete $o(n^k)$ edges from H to obtain a subgraph H' for which there exists a partition $\mathcal{P}$ of $V(H')$ such that $\mathbf{i}_{\mathcal{P}}(V(H')) \notin L_{\mathcal{P}}(H')$. Thus if H is far from a divisibility barrier then it has a perfect matching. On the other hand, if H is itself a divisibility barrier then Proposition 1.4 implies that H does not have a perfect matching. However, our algorithm cannot search directly for such partitions $\mathcal{P}$, since the number of edges to be deleted, while small compared to n^k, is still large from a computational perspective.

The main theoretical contribution of this paper is to fill the gap between these cases, by giving a condition for the existence of a perfect matching under Setup 1.7 that is necessary and sufficient, and also efficiently checkable.

1.4. Definitions

Before giving the statement we require several definitions. Firstly, it will be sufficient to consider the following special classes of lattices. To motivate this definition, we remark that all of our lattices will be edge-lattices, any incomplete lattice can be simplified to one that is transferral-free, and our assumptions on H will imply that if its edge-lattice with respect to some partition is transferral-free then it is full. Note also that our definitions depend on k, but we consider this to be fixed throughout the paper. We write $|\mathcal{P}|$ for the number of parts of a partition $\mathcal{P}$.

Definition 1.8. Suppose L is a lattice in $\mathbb{Z}^d$.

(i) We say that $\mathbf{i} \in \mathbb{Z}^d$ is an *r-vector* if it has non-negative co-ordinates that sum to r. We write $\mathbf{u}_j$ for the 'unit' 1-vector that is 1 in co-ordinate j and 0 in all other co-ordinates.
(ii) We say that L is an *edge-lattice* if it is generated by a set of k-vectors.
(iii) We write $L^d_{\max}$ for the lattice generated by all k-vectors.
(iv) We say that L is *complete* if $L = L^d_{\max}$, otherwise it is *incomplete*.
(v) A *transferral* is a non-zero difference $\mathbf{u}_i - \mathbf{u}_j$ of 1-vectors.
(vi) We say that L is *transferral-free* if it does not contain any transferral.
(vii) We say that a set I of k-vectors is *full* if for every $(k-1)$-vector $\mathbf{v}$ there is some $i \in [d]$ such that $\mathbf{v} + \mathbf{u}_i \in I$.
(viii) We say that L is *full* if it contains a full set of k-vectors and is transferral-free.

Note that $L^d_{\max}$ can be equivalently defined as the lattice of all vectors in $\mathbb{Z}^d$ whose coordinates sum to a multiple of k. Often the dimensions of $\mathbb{Z}^d$ will correspond to parts of a partition $\mathcal{P}$ of a set; in this case we refer to $\mathbb{Z}^{\mathcal{P}}$ and $L^{\mathcal{P}}_{\max}$ instead of $\mathbb{Z}^d$ and $L^d_{\max}$.

Now we come to the structures that appear in our characterisation. To motivate this definition, we remark that the first step in our strategy for finding a perfect matching will be to identify a 'canonical' partition and lattice, then delete the vertices covered by a small matching so that the index of the remaining set is on the lattice.

Definition 1.9. Let H be a k-graph.

(i) A *full pair* $(\mathcal{P}, L)$ for H consists of a partition $\mathcal{P}$ of $V(H)$ into $d \le k-1$ parts and a full edge-lattice $L \subseteq \mathbb{Z}^d$ (possibly distinct from $L_{\mathcal{P}}(H)$).
(ii) A (possibly empty) matching M of size at most $|\mathcal{P}| - 1$ is a *solution* for $(\mathcal{P}, L)$ (in H) if $\mathbf{i}_{\mathcal{P}}(V(H) \backslash V(M)) \in L$; we say that $(\mathcal{P}, L)$ is *soluble* if it has a solution, otherwise *insoluble*.
(iii) A full pair $(\mathcal{P}, L)$ is *C-far* for H if some set of C vertices intersects every edge $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) \notin L$.
(iv) A *C-certificate* for H is an insoluble C-far full pair for H.

We will see later that there is no loss of generality in considering partitions into at most $k-1$ parts because of our codegree assumption, or in requiring solutions to have size at most $|\mathcal{P}| - 1$, as if there is any matching M such that $\mathbf{i}_{\mathcal{P}}(V(H) \backslash V(M)) \in L$ then there is one with $|M| \le |\mathcal{P}| - 1$. In particular, if H has a perfect matching then any full pair is soluble, so there is no C-certificate for H for any $C \ge 0$.

1.5. Main structural result

Now we can state our structural characterisation for the perfect matching problem under Setup 1.7.

Theorem 1.10. *Under Setup 1.7, H has a perfect matching if and only if there is no $2k(k-3)$-certificate for H.*

We remark that most of the work in proving Theorem 1.10 lies in establishing the "if" part of the statement; the "only if" part follows easily from the results in Section 6. Our algorithm for the decision problem is essentially an exhaustive search for a $2k(k-3)$-certificate, although we also need to provide an algorithm to efficiently list the partitions $\mathcal{P}$ that may arise in Definition 1.9. Thus Theorem 1.10 is required to prove correctness of the algorithm. We remark that the constant $2k(k-3)$ is within a factor 2 of being best-possible (see Conjecture 10.3). Furthermore, the constant $1/k$ in Setup 1.7 is best-possible, as shown by the 'space barrier' construction (see the concluding remarks for the definition and further discussion).

1.6. Contents

In the next section we present the algorithmic details of the results in this introduction. That is, we assume Theorem 1.10 and deduce Theorem 1.1. The rest of the paper is devoted to the proof of Theorem 1.10, beginning with Section 3 in which we briefly sketch the ideas of the proof and its key lemmas (Lemmas 7.1 and 7.6). In Section 4 we introduce a number of necessary preliminaries, including results from [17], some convex geometry and well-known probabilistic tools. Section 5 focuses on an important definition, that of 'robust maximality', and some of its properties; this turns out to be the correct notion for the key lemmas. In Section 6 we establish various properties of full lattices, including a characterisation in terms of finite abelian groups. After these preparations, we prove the key lemmas in Section 7. In Section 8 we deduce Theorem 1.10 from Lemma 7.1 and some additional lemmas on subsequence sums in finite abelian groups. Section 9 contains some technical proofs which were deferred from earlier sections. Finally, in Section 10 we state the multipartite versions of our results (omitting the similar proofs) and make some concluding remarks.

1.7. Notation and terminology

A *hypergraph* H consists of a vertex set $V(H)$ and a set $E(H)$ of edges $e \subseteq V(H)$. We frequently identify a hypergraph H with its edge set, for example, writing $e \in H$ for $e \in E(H)$ and $|H|$ for $|E(H)|$. A *k-uniform* hypergraph, or *k-graph*, is a hypergraph in which every edge is a *k-set*, that is, a set of size k. Given a hypergraph H and $A \subseteq V(H)$, the *neighbourhood* of A in H is $H(A) = \{B \subseteq V(H) \setminus A : A \cup B \in H\}$. Note that $|H(A)| = d_H(A)$ is the degree of A in H.

The hypergraph $H[A]$ (as distinct from $H(A)$) is the hypergraph with vertex set A whose edges are the edges of H which are contained in A. We use $H - A$ and $H \setminus A$ interchangeably to denote the hypergraph obtained from H by deleting A and all edges which intersect A (this is identical to $H[V(H) \setminus A]$). If H is a hypergraph and $H' \subseteq H$

we say that H' is a *subgraph* of H (we prefer to avoid the terms 'subhypergraph' and 'sub-k-graph').

We use bold font for vectors and normal font for their co-ordinates, e.g. $\mathbf{v} = (v_1, \dots, v_d)$. We write $\mathbf{0}$ and $\mathbf{1}$ for the vectors whose co-ordinates are all zero and one respectively (the dimension of the vector will be clear from the context). $B^d(\mathbf{v}, r)$ denotes the ball of radius r around $\mathbf{v} \in \mathbb{Z}^d$; we sometimes omit the dimension d. We will often work with vectors $\mathbf{v}$ in $\mathbb{Z}^d$, for some d, in which the coordinates are indexed by some ordered partition $\mathcal{P}$ of a set V with d parts. If X is the jth part of $\mathcal{P}$ for some $j \in [d]$ we write $v_X = v_j$.

We say that an event E holds *with high probability* if $\mathbb{P}(E) = 1 - e^{-\Omega(n^c)}$ for some $c > 0$ as $n \to \infty$; note that when n is sufficiently large, by union bounds we can assume that any specified polynomial number of such events all occur. We write $[r]$ to denote the set of integers from 1 to r, and $x \ll y$ to mean for any $y \geq 0$ there exists $x_0 \geq 0$ such that for any $x \leq x_0$ the following statement holds. Similar statements with more constants are defined similarly. Also, we write $a = b \pm c$ to mean $b - c \leq a \leq b + c$. Throughout the paper we omit floor and ceiling symbols where they do not affect the argument.

2. Algorithms and analysis

We start with the following theorem, which can be used to solve the decision problem of determining whether or not H has a perfect matching. Note that the running time for $k = 3$ is $O(n^3)$, which we cannot reasonably expect to improve, as a faster algorithm would not be able to query all edges of H.

Theorem 2.1. *Under Setup 1.7, Procedure DeterminePM determines whether H contains a perfect matching in time $O(n^{3k^2-8k})$.*

Procedure DeterminePM.

Data: A k-graph H as in Setup 1.7.
Result: Determines whether H has a perfect matching.
if $n < n_0$ **then**
  Examine every set of n/k edges in H, and halt with appropriate output.
foreach *set $S \subseteq V(H)$ of size at most $2k(k-3)$, integer $d \in [k-1]$, full edge-lattice $L \subseteq \mathbb{Z}^d$ and partition $\mathcal{P}$ of $V(H)$ into d parts so that any edge $e \in H$ which does not intersect S has $\mathbf{i}_{\mathcal{P}}(e) \in L$* **do**
  if *there is no matching $M \subseteq H$ of size at most $d-1$ such that $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M)) \in L$* **then**
    Output "no perfect matching" and halt.
Output "perfect matching" and halt.

Procedure DeterminePM is essentially an exhaustive search for a $2k(k-3)$-certificate. It is clear that the ranges of S, d, L and M in the procedure can be listed by brute force in polynomial time. However, brute force cannot be used for $\mathcal{P}$, as there are potentially exponentially many possibilities to consider, so first we provide an algorithm to construct all possibilities for $\mathcal{P}$.

We imagine each vertex class V_j to be a 'bin' to which vertices may be assigned, and keep track of a set U of vertices yet to be assigned to a vertex class. So initially we take each V_j to be empty and $U = V(H)$. The procedure operates as a *search tree*; at certain points the instruction is to branch over a range of possibilities. This means to select one of these possibilities and continue with this choice, then, when the algorithm halts, to return to the branch point, select the next possibility, and so forth. Each branch may produce an output partition; the output of the procedure consists of all output partitions. An informal statement of our procedure is that we generate partitions by repeatedly branching over all possible assignments of a vertex to a partition class, exploring all consequences of each assignment before branching again. Furthermore, we only branch over assignments of vertices which satisfy the following condition. Given a set of assigned vertices, we call an unassigned vertex x *reliable* if there exists a set B of $k-2$ assigned vertices such that $d(x \cup B) \geq (1/k+\gamma)n$.

Lemma 2.2. *Under Setup 1.7, for any $d \in [k-1]$ and full edge-lattice $L \subseteq \mathbb{Z}^d$, there are at most d^{2k-2} partitions $\mathcal{P}$ of $V(H)$ such that $\mathbf{i}_{\mathcal{P}}(e) \in L$ for every $e \in H$, and Procedure ListPartitions lists them in time $O(n^{k+1})$.*

Procedure ListPartitions.

Data: A k-graph H and a full edge-lattice $L \subseteq \mathbb{Z}^d$.
Result: Outputs all partitions $\mathcal{P}$ of $V(H)$ with $\mathbf{i}_{\mathcal{P}}(e) \in L$ for every $e \in H$.

Set $U = V(H)$ and let $\mathcal{P} = (V_1, \dots, V_d)$ be a partition of $V \setminus U$ (so initially $V_i = \emptyset$).
Choose arbitrarily $A \subseteq V(H)$ of size $k-1$ such that $d(A) \geq (1/k+\gamma)n$.
Branch over all possible assignments of vertices in A to $V_1, \dots, V_d$.

while $U \neq \emptyset$ **do**
  if $xy_1 \dots y_{k-1} \in H$ *for some vertices* $x \in U$ *and* $y_1, \dots, y_{k-1} \notin U$ **then**
    Fix $j \in [k]$ such that $\mathbf{i}_{\mathcal{P}}(y_1 \dots y_{k-1}) + \mathbf{u}_j \in L$.
    Assign x to V_j and remove x from U.
  else
    Choose $x \in U$ which is reliable.
    Branch over all possible assignments of x.
if $\mathbf{i}_{\mathcal{P}}(e) \in L$ *for every* $e \in H$ **then** halt with output $\mathcal{P}$.

Proof. First we note that since L is full, the instruction "Fix $j \in [k]$ such that $\mathbf{i}_{\mathcal{P}}(y_1 \dots y_{k-1}) + \mathbf{u}_j \in L$" in Procedure ListPartitions is well-defined, since for any $(k-1)$-vector $\mathbf{v}$ there is precisely one $j \in [k]$ such that $\mathbf{v} + \mathbf{u}_j \in L$.

Next we show that if the number of assigned vertices is at least $(1/k+\gamma)n$ and at most $(1-\gamma)n$ then there is always a reliable unassigned vertex. To see this, note that the number of sets $x \cup B$, where x is unassigned and B is a set of $k-2$ assigned vertices, is at least $\gamma n \binom{n/k+\gamma n}{k-2} > \varepsilon n^{k-1}$ (recall from Setup 1.7 that $\varepsilon \ll \gamma$). Hence some such $x \cup B$ has degree at least $(1/k+\gamma)n$, and so x is reliable.

Observe that after branching initially over all possible assignments of A, at least $(1/k+\gamma)n$ vertices will be assigned (all the neighbours of A) before the procedure branches again. Further, the procedure will no longer branch once there are fewer than γn unassigned vertices remaining. Indeed, in this case for any unassigned vertex x there

are fewer than γn^{k-1} edges containing x and another unassigned vertex. Thus condition (deg) implies that every unassigned vertex x is contained in some edge $xy_1 \dots y_{k-1}$ of H where $y_1, \dots, y_{k-1}$ are assigned. Hence when branching, we may always choose a reliable vertex as stated in the procedure.

The final line of the procedure ensures that any partition $\mathcal{P}$ of $V(H)$ which is output has that property that $\mathbf{i}_{\mathcal{P}}(e) \in L$ for every $e \in H$. The converse is also true: any partition $\mathcal{P}$ of $V(H)$ such that $\mathbf{i}_{\mathcal{P}}(e) \in L$ for every $e \in H$ will be output by some branch of the procedure. To see this, consider the branch of the procedure in which, at each branch point, the vertex x under consideration is assigned to the vertex class in which it lies in $\mathcal{P}$. By our initial remark, every other vertex of H must also be assigned to the vertex class in which it lies in $\mathcal{P}$. We conclude that Procedure ListPartitions indeed runs correctly, returning all partitions $\mathcal{P}$ of $V(H)$ such that $\mathbf{i}_{\mathcal{P}}(e) \in L$ for every $e \in H$.

It remains to bound the number of such partitions. Consider some x over which the procedure branches. Then there can be no edge $xy_1 \dots y_{k-1}$ of H where $y_1, \dots, y_{k-1}$ are assigned. Suppose that this is the case, and let $B = y_1 y_2 \dots y_{k-2}$ be a set of assigned vertices such that $d(\{x\} \cup B) \geq (1/k + \gamma)n$. (Such a B must exist since we chose x to be reliable.) None of the $(1/k + \gamma)n$ vertices v such that $xvy_1 \dots y_{k-2}$ is an edge of H can have been assigned, and each will be assigned before the next branch of the procedure. We conclude that after any branch in the procedure and before the next branch, at least $(1/k + \gamma)n$ vertices are assigned.

Hence the search tree has depth at most k. Since the degree of the root is at most d^{k-1} and every other vertex has d children, the search tree has at most d^{2k-2} leaves. At most one partition is output at each leaf, so at most d^{2k-2} partitions $\mathcal{P}$ will be output by the algorithm, as required. Furthermore, over all branches there will be at most $d^{2k-2}n$ iterations of the *while* loop, and the condition of the first *if* statement takes $O(n^k)$ operations to check, and so the overall running time is $O(n^{k+1})$. □

Remark 2.3. In fact, the step in which the algorithm finds $xy_1 \dots y_{k-1} \in H$ for some vertices $x \in U$ and $y_1, \dots, y_{k-1} \notin U$ can be made more efficient by maintaining a queue of 'new' vertices and only considering k-tuples for which y_1 is the vertex at the front of the queue. Initially the queue consists of the vertices of A. Whenever a vertex is assigned it is added to the back of the queue, and the front vertex y_1 is removed from the queue (i.e. ceases to be new) once we have considered all k-tuples $xy_1 \dots y_{k-1}$ with $x \in U$ and $y_1, \dots, y_{k-1} \notin U$ in which y_1 is the only new vertex. With this modification, the running time of the algorithm can be reduced to $O(n^k)$. Also, for the purpose of Lemma 2.2 we could replace the constant $1/k + \gamma$ in Setup 1.7 by any positive constant independent of n and still deduce that the number of possible partitions is independent of n.

Proof of Theorem 2.1. Let H be a k-graph as in Setup 1.7. We begin by noting that Procedure DeterminePM determines whether H has a perfect matching. Indeed, this is trivial if $n < n_0$, and if $n \geq n_0$ it follows from Theorem 1.10, as Procedure DeterminePM

determines whether there is a $2k(k-3)$-certificate $(\mathcal{P}, L)$ for H. We estimate the running time as follows. There are at most $n^{2k(k-3)}$ choices of sets S, and these can be generated in time $O(n^{2k(k-3)})$. Also, there are only a constant number of choices for d and L, and these can be generated in constant time. Indeed, since L is an edge-lattice, it is generated by a set of k-vectors, and the number of such generating sets is bounded by a function of k.

For each choice of S, d and L, generating the list of choices for $\mathcal{P}$ takes time $O(n^{k+1})$ by Lemma 2.2 applied to $H \setminus S$. Further the number of choices for $\mathcal{P}$ is constant, and for each one it takes time $O(n^{k(k-2)})$ to check for the existence of the matching M. When $k > 3$, $k(k-2) > k+1$ and so we conclude that the running time is $O(n^{3k^2-8k})$, as required. In the case $k = 3$, we use an improved algorithm with running time $O(n^3)$ (which exists as noted in Remark 2.3) to generate the list of choices for $\mathcal{P}$, thus obtaining the result. □

Now we will motivate the proof of our main result, Theorem 1.1. We start by using Procedure DeterminePM to check whether H contains a perfect matching. If H has no perfect matching then this is certified by the $2k(k-3)$-certificate $(\mathcal{P}, L)$ found by Procedure DeterminePM. So suppose that H does contain a perfect matching. How can we find it? A naive attempt at a proof is the following well-known idea. We examine each edge e of H in turn and use the same procedure to test whether deleting the vertices of e would still leave a perfect matching in the remainder of H, in which case we say that e is *safe*. There must be some safe edge e, which we add to our matching, then repeat this process, until the number of vertices falls below n_0, at which point we can find a perfect matching by a constant-time brute force search.

The problem with this naive attempt is that as we remove edges, the minimum codegree may become too low to apply Procedure DeterminePM, and then the process cannot continue. To motivate the solution to this problem, suppose that we have oracle access to a uniformly random edge from some perfect matching. Such an edge is safe, and if we repeatedly remove such random edges, standard large deviation estimates show that with high probability the minimum codegree condition is preserved (replacing γ by $\gamma/2$, say). As an aside, we note that since Linear Programming has a polynomial time algorithm, we can construct a distribution (p_e) on the safe edges such that $\sum_{e:x\in e} p_e = k/n$ for every vertex x; using this distribution instead of the oracle provides a randomised algorithm for finding a perfect matching.

Our actual algorithm is obtained by derandomising the oracle algorithm. Instead of a minimum codegree condition, we bound the sum of squares of codegree 'deficiencies', which is essentially the condition (codeg) of Setup 1.7. We also need to introduce the vertex degree condition (deg), otherwise we do not have an effective bound on the number of partitions $\mathcal{P}$ in Lemma 2.2. Conditions (i) and (ii) in the following lemma effectively serve as proxies for (codeg) and (deg) respectively. Note that by $H - e$ we mean the k-graph obtained from H by deleting all vertices of e and all edges incident to them.

Lemma 2.4. *Suppose that $1/n \ll \varepsilon \ll \gamma \ll 1/k$, that $k \geq 3$ and that H is a k-graph on n vertices. For each set A of $k-1$ vertices of H let $t_A = \max(0, (1/k+\gamma)n - d_H(A))$. Suppose that*

(i) $\sum_{A \in \binom{V(H)}{k-1}} t_A^2 < \varepsilon\gamma^2 n^{k+1}/4 + 3kn^k$,
(ii) $n^{k-1}/3k! - \delta_1(H) + \sum_{A \in \binom{V(H)}{k-1}} \frac{t_A^2}{\sqrt{\varepsilon}\gamma^2 n^2} < \sqrt{\varepsilon} n^{k-1}$,
(iii) *H contains a perfect matching.*

Then we can find, in time at most $O(n^{3k^2-7k})$, an edge $e \in H$ such that (i), (ii) and (iii) also hold for $H - e$ with $n-k$ in place of n.

The proof of Lemma 2.4 involves messy calculations, so we defer the full proof until Section 9 and for now just describe the idea. While $\delta_1(H) > n^{k-1}/3k!$ we only need to maintain condition (i); then an averaging argument shows that the required edge exists. On the other hand, if there is a vertex of small degree we can remove any edge containing it: this so greatly decreases $\sum_{A \in \binom{V(H)}{k-1}} t_A^2$ as to compensate for any further decrease in $\delta_1(H)$.

Proof of Theorem 1.1. We begin by running Procedure DeterminePM to confirm that H contains a perfect matching; if it does not, then we obtain a $2k(k-3)$-certificate for H, which by Theorem 1.10 is a certificate that no perfect matching exists. If H does contain a perfect matching, then we repeatedly apply Lemma 2.4 to delete edges of H (along with their vertices). Initially, condition (iii) of Lemma 2.4 holds by assumption, and conditions (i) and (ii) hold as the codegree assumption implies $t_A = 0$ for any $(k-1)$-set A and $\delta_1(H) \geq k^{-1}\binom{n}{k-1}$. Since Lemma 2.4 ensures that its conditions are preserved after an edge is deleted, we may repeat until n is too small for further application of Lemma 2.4. At this point the number of vertices remaining in H is bounded by a constant depending only on ε, γ and k (this follows from the definition of $\ll$), so we can use the brute-force algorithm to find a perfect matching in the remainder of H in constant time. Together with the deleted edges this forms a perfect matching in H. Since k vertices are deleted from H with each application of Lemma 2.4, at most n/k applications are required in total. Thus the total running time of the algorithm is $O(n^{3k^2-7k+1})$. □

3. Outlines of the proofs

In this section we briefly sketch the proof of Theorem 1.10. The easier direction is that if there is a perfect matching then there is no $2k(k-3)$-certificate. Let H be a k-graph as in Setup 1.7 that contains a perfect matching M^*. To show that there is no $2k(k-3)$-certificate for H, it suffices to show that every $2k(k-3)$-far full pair $(\mathcal{P}, L)$ for H is soluble. To accomplish this, we first show that L is a subgroup of $L^{\mathcal{P}}_{\max}$ with index at most $|\mathcal{P}|$. Then by a simple application of the pigeonhole principle, we can find

a matching M in H of size at most $|\mathcal{P}|-1$ for which $\mathbf{i}_{\mathcal{P}}(M)$ lies in the same coset of L in $L^{\mathcal{P}}_{\max}$ as $\mathbf{i}_{\mathcal{P}}(M^*)$. Since $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M^*)) = \mathbf{0} \in L$ we deduce that $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M)) \in L$, and so M solves $(\mathcal{P}, L)$.

Let us now turn our attention to the remaining implication of the theorem, i.e. that if every $2k(k-3)$-far full pair is soluble then there is a perfect matching. We reduce this to an easier version of the theorem in which we assume that every full pair is soluble. The proof of this reduction relies on properties of subsequence sums in abelian groups, that (roughly speaking) allow us to convert any insoluble full pair into a $2k(k-3)$-far full pair by merging parts. To prove the easier version, we begin by identifying a canonical partition $\mathcal{P}$ of $V(H)$: we choose $\mathcal{P}$ to be 'robustly maximal' with respect to H, which roughly speaking means that it is an approximate divisibility barrier with no other approximate divisibility barrier 'hidden' inside. By reassigning a small number of vertices to different parts we can further ensure that every vertex lies in many edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) \in L$. We will see that $(\mathcal{P}, L)$ is a full pair, so by our assumption it has a solution M. Our key lemma will show that under these conditions $H \setminus V(M)$ has a perfect matching, which together with M forms a perfect matching in H.

We prove the key lemma by reducing it to a partite version, using a random k-partition of H. Thus for much of the paper we work with a k-partite k-graph H whose vertex classes each have size n. Our codegree conditions will apply only to $(k-1)$-sets which contain at most one vertex from any vertex class. We will work with the 'μ-robust' edge-lattice $L^{\mu}_{\mathcal{P}}(H)$, which is defined similarly to $L_{\mathcal{P}}(H)$, except that we require many $(\mu|V(H)|^k)$ edges of a given index vector for it to be included in the generating set (see Definition 4.1).

A key idea in the proof of is that of splitting H into a number of k-partite k-graphs, each of which satisfies a stronger version of the properties of H. More precisely, if almost all partite $(k-1)$-sets of vertices of H have degree at least $n/\ell + \gamma n$ and $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free, then we can find a set of k-partite subgraphs whose vertex sets partition $V(H)$, such that each satisfies a stronger codegree condition: almost all partite $(k-1)$-sets have degree at least $1/(\ell-1)$ proportion of the size of each vertex class. By induction on ℓ we will obtain a perfect matching in each subgraph, which together give a perfect matching in H.

3.1. An analogue for tripartite 3-graphs

To illustrate this idea better, we now outline the proof of an analogous result to Theorem 1.5 for 3-partite 3-graphs. Indeed, let H be a 3-partite 3-graph with vertex classes V_1, V_2 and V_3 each of size n. Suppose that any partite pair in H has codegree at least $(1/3+\gamma)n$. We shall prove that either H contains a perfect matching, or $V(H)$ can be partitioned into two parts A and B such that every edge of H has an even number of vertices in A, but $|A|$ is odd (i.e. there is a divisibility barrier).

We begin by applying Theorem 4.6 to H. This implies that either H contains a perfect matching (in which case we are done) or that there is a partition $\mathcal{P}$ of $V(H)$ into parts of

size at least $(1/3+\gamma/2)n$ which refines the 3-partition of $V(H)$ and is such that $L^{\mu}_{\mathcal{P}}(H)$ is incomplete and transferral-free. Note that this means that $\mathcal{P}$ refines each vertex class V_i into at most two parts; in fact, it is not hard to see that $\mathcal{P}$ must partition each vertex class V_i into precisely two parts, W_i^1 and W_i^2.

We now decompose H into eight 3-partite subgraphs $(H^{ijk})_{i,j,k\in[2]}$, where H^{ijk} consists of all edges of H which contain one vertex from each of W_1^i, W_2^j and W_3^k. So every edge of H lies in precisely one subgraph H^{ijk}. Since $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free, we may assume without loss of generality that almost all of the edges of H lie in one of $H^{111}, H^{122}, H^{212}$ or H^{221} (the remaining subgraphs have very low density). As described earlier, we now observe that almost all partite pairs in each subgraph have codegree at least $1/2+\gamma/2$ proportion of the size of the remaining vertex class in that subgraph. Indeed, our assumption on H tells us that any pair of vertices xy with $x \in W_1^1$ and $y \in W_2^1$ has at least $(1/3+\gamma)n$ neighbours in V_3 (i.e. vertices $z \in V_3$ such that $xyz \in H$). But since H_{112} has very low density, very few such pairs xy can have many neighbours in W_3^2. So almost all of these pairs must have $(1/3+\gamma/2)n \geq (1/2+\gamma/2)|W_{31}|$ neighbours in W_{31}. So H^{111} satisfies a significantly stronger codegree condition than that satisfied by H, ignoring the fact that a small number of pairs, which we will call *bad* pairs, fail this codegree condition. A similar argument applies to H^{122}, H^{212} and H^{221}.

At this point we delete a (small) matching M in H to achieve two aims. Firstly, M will cover all vertices which lie in many bad pairs. Since there are few bad pairs there will only be a small number of such vertices. Secondly, after deleting the vertices covered by M there will be an *even* number of vertices remaining in $W^2 := W_1^2 \cup W_2^2 \cup W_3^2$. This can be done unless $|W^2|$ is odd and every edge of H intersects W^2 in an even number of vertices; in this case we are done, and H contains no perfect matching.

Now we delete the vertices covered by M from H; this only slightly weakens the codegree condition on H^{111}, H^{122}, H^{212} and H^{221} since M does not contain many edges. Then we choose a random partition of the remaining vertices of H into subsets $S_{111}, S_{122}, S_{212}$ and S_{221} under the constraints that (i) each subset contains equally many vertices from each vertex class V_j, and (ii) $S_{ijk} \subseteq W_1^i \cup W_2^j \cup W_3^k$. Such partitions exist since each part W_j^i has size greater than $n/3$ and $|W^2|$ is even (this is a special case of Proposition 6.13). Let H'_{111} be the 3-partite subgraph of H induced by S_{111}, and define H'_{122}, H'_{212}, and H'_{221} similarly. The fact that the partition was chosen randomly implies that, in each subgraph, almost every partite pair has codegree at least $(1/2+\gamma/2)n'$, where n' is the number of vertices in each part, and no vertex lies in many bad pairs. It follows by [1, Theorem 2] (or by Theorem 4.6) that each subgraph contains a perfect matching; together with the deleted matching this yields a perfect matching in H.

3.2. The general case

The case $k = 3$ is particularly simple because there is only one maximal divisibility barrier; recall that for $k \geq 4$, Construction 1.6 shows that a naive generalisation of Theorem 1.5 does not hold. For the general case of Theorem 1.10, we use the ideas

outlined in Section 3.1, but there are additional complications. Firstly, there are now many possibilities for the partition $\mathcal{P}$ returned by applications of Theorem 4.6. Secondly, multiple applications of the above technique are required to successively improve the codegree condition on the k-graphs into which H is decomposed.

These complications will be handled by our partite key lemma (Lemma 7.6), which can be seen as a variation on Theorem 4.6 specialised to k-partite k-graphs with vertex classes of size n in which almost all partite $(k-1)$-sets have codegree at least $(1/\ell+\gamma)n$, with two key differences. One is that, instead of the condition of Theorem 4.6 that $L = L^{\mu}_{\mathcal{P}}(H)$ must be complete, we now only require that $\mathbf{i}_{\mathcal{P}}(V(H)) \in L$. The other is that, whereas the condition of Theorem 4.6 must hold for any partition $\mathcal{P}$ of $V(H)$ into sufficiently large parts, we now only require that $\mathbf{i}_{\mathcal{P}}(V(H)) \in L$ for a single 'canonical' partition $\mathcal{P}$ which meets two additional requirements.

Firstly, we require that every vertex must lie in many edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) \in L$. This condition can be seen as ensuring that each vertex of H lies in the 'correct part' of $\mathcal{P}$. For example, consider the extremal example described in Construction 1.2. If we fix some small $\mu > 0$ and move a single vertex v from part A to part B (but do not change the edge set of H), then the only edges whose index changes are the fewer than n^{k-1} edges which contain v. So $L^{\mu}_{\mathcal{P}}(H)$ is unchanged, and H doesn't have a perfect matching (since H itself is unchanged), but we now have $\mathbf{i}_{\mathcal{P}}(V(H)) \in L^{\mu}_{\mathcal{P}}(H)$, due to v being in the 'wrong part' of $\mathcal{P}$. Secondly, we must assume that L is transferral-free. However this requirement, while necessary, is not sufficient, as the following example will show.

Construction 3.1. Fix $k \geq 5$ and let $\mathcal{P} = \{W_1, W_2\}$ be a partition of V. Let $\mathcal{Q}$ be a refinement of $\mathcal{P}$ which divides W_1 into V_{11} and V_{12} and W_2 into V_{21} and V_{22}. Suppose that $|W_1|$ is even, but that $|V_{11} \cup V_{21}|$ is odd. Let H be the k-graph on V whose edges are precisely the k-subsets of V which contain an even number of vertices in W_1 and an even number of vertices in $V_{11} \cup V_{21}$.

In Construction 3.1, $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free, $\delta_{k-1}(H) \geq (1/k+\gamma)n$ and $\mathbf{i}_{\mathcal{P}}(V(H)) \in L^{\mu}_{\mathcal{P}}(H)$. Thus the conditions relating to $\mathcal{P}$ do not preclude the existence of a perfect matching, but the conditions relating to $\mathcal{Q}$ do. Indeed, H cannot contain a perfect matching, since it is a subgraph of the k-graph described in Construction 1.2, where $A = V_{11} \cup V_{21}$ and $B = V_{12} \cup V_{22}$. To avoid this kind of situation, we insist that there is no strict refinement $\mathcal{Q}$ of $\mathcal{P}$ into not-too-small parts such that $L^{\mu}_{\mathcal{Q}}(H)$ is transferral-free. Note that the trivial partition into a single part satisfies this requirement if and only if H is not close to a divisibility barrier. In fact, we require the stronger property that for some $\mu' \gg \mu$ (i.e. even at a much weaker 'detection threshold') the lattice $L^{\mu'}_{\mathcal{Q}}(H)$ is not transferral-free for any strict refinement $\mathcal{Q}$ of $\mathcal{P}$ into not-too-small parts. This property is 'robust' in the sense that when we delete a small number of vertices from H (as we will need to do), the property is preserved, albeit with weaker constants. (Section 5 is devoted to the study of robust maximality.) If $\mathcal{P}$ satisfies both of these conditions, then Lemma 7.6 states that H must contain a perfect matching.

We prove Lemma 7.6 by induction on ℓ by a similar argument to that in Section 3.1 (roughly speaking, there we reduced the $\ell = 3$ case to the $\ell = 2$ case, for $k = 3$). That is, we apply Theorem 4.6 to yield a perfect matching (in which case we are done), or a partition $\mathcal{P}$ of $V(H)$ into parts of size at least $(1/\ell+\gamma)n$ such that $L^{\mu}_{\mathcal{P}}(H)$ is incomplete and transferral-free. For the base case of the induction we observe that the latter outcome is impossible for $\ell = 2$. Next, for each $\mathbf{i} \in I^{\mu}_{\mathcal{P}}(H)$, we define $H_{\mathbf{i}}$ to be the 'canonical' induced subgraph of H on the union of the parts $W \in \mathcal{P}$ such that $i_W = 1$, and $\mathcal{P}_{\mathbf{i}}$ to be the restriction of $\mathcal{P}$ to $V(H_{\mathbf{i}})$. By a similar argument to the previous section (formalised in Proposition 7.5), we find that each $H_{\mathbf{i}}$ satisfies a codegree condition similar to that on H, but with $\ell - 1$ in place of ℓ. Our aim is then to find vertex-disjoint subgraphs $\hat{H}_{\mathbf{i}} \subseteq H_{\mathbf{i}}$ whose vertex sets partition $V(H)$, each of which satisfies the conditions of Lemma 7.6 with the stronger codegree condition ($\ell-1$ in place of ℓ). We can then apply the inductive hypothesis to find a perfect matching in each $\hat{H}_{\mathbf{i}}$; together these form a perfect matching in H.

To do this, we first take a robustly maximal refinement $\mathcal{Q}_{\mathbf{i}}$ of $\mathcal{P}_{\mathbf{i}}$ for each $\mathbf{i}$, and delete a small matching that covers all 'bad' vertices. We then need to ensure that our partition into subgraphs $H_{\mathbf{i}}$ satisfies $\mathbf{i}_{\mathcal{Q}_{\mathbf{i}}}(V(\hat{H}_{\mathbf{i}})) \in L^{\mu}_{\mathcal{Q}_{\mathbf{i}}}(H_{\mathbf{i}})$ for each $\mathbf{i}$; accomplishing this is the most technical part of the proof. Observe that whether this condition is satisfied depends only on the values of $|V(\hat{H}_{\mathbf{i}}) \cap Y|$ for parts $Y \in \mathcal{Q}_{\mathbf{i}}$. Thus we need to choose the sizes of the intersections $V(\hat{H}_{\mathbf{i}}) \cap Z$ for parts $Z \in \mathcal{Q}^{\cap}$, where $\mathcal{Q}^{\cap}$ is the 'meet' of the partitions $\{\mathcal{Q}_{\mathbf{i}} \mid \mathbf{i} \in I\}$. We achieve this in three stages. Firstly, in Claim 7.8, we choose rough targets for the size of each $V(\hat{H}_{\mathbf{i}})$; here we use a geometric method which relies on Farkas' Lemma (Theorem 4.7). Secondly, in Claim 7.9, we choose how many vertices each $\hat{H}_{\mathbf{i}}$ will take from each part of the 'join' $\mathcal{Q}^{\cup}$ of the partitions $\mathcal{Q}_{\mathbf{i}}$ (this step deals with the problem illustrated by Construction 3.1). Thirdly, in Claim 7.10, we use Baranyai's Matrix Rounding Theorem to refine this choice to obtain the intersection sizes we require. Finally, we choose the vertex set of each $\hat{H}_{\mathbf{i}}$ at random with the given number of vertices from each part of $\mathcal{Q}^{\cap}$. In Claim 7.11 we demonstrate that with high probability this random selection does indeed give subgraphs $\hat{H}_{\mathbf{i}}$ to which we can apply the inductive hypothesis, completing the proof.

4. Preliminaries

This section contains theoretical background needed for the rest of the paper, organised into the following subsections: 4.1 Partitions, index vectors and lattices, 4.2 Hypergraph matching theory, 4.3 Convex geometry, 4.4 Baranyai's Matrix Rounding Theorem, 4.5 Concentration of probability.

4.1. Partitions, index vectors and lattices

We will frequently speak of a partition $\mathcal{Q}$ of a vertex set V. We use this term in a slightly non-standard way to mean a family of pairwise-disjoint subsets of V whose union is V (so it is possible for a part to be empty, though this will rarely be the

case). Furthermore, we implicitly fix an order on the parts of the partition, so we may consistently speak of, for example, the ith part of $\mathcal{Q}$. Given $V' \subseteq V$, the *restriction* of $\mathcal{Q}$ to V', denoted $\mathcal{Q}[V']$, is the partition of V' with parts $X \cap V'$ for $X \in \mathcal{Q}$.

Many of our results will apply specifically to partite k-graphs. Let $\mathcal{P}$ partition a set V. Then we say that a set $S \subseteq V$ is *$\mathcal{P}$-partite* if S has at most one vertex in any part of $\mathcal{P}$. We say that a hypergraph H on V is *$\mathcal{P}$-partite* if every edge of H is $\mathcal{P}$-partite. In this case we refer to the parts of $\mathcal{P}$ as the *vertex classes* of H. Usually, we will consider k-graphs H which are *k-partite*, i.e. $\mathcal{P}$-partite for some partition $\mathcal{P}$ of $V(H)$ into k parts.

Given a k-graph H and a partition $\mathcal{P}$ of $V(H)$, the lattice $L_{\mathcal{P}}(H)$ in Definition 1.3 can be seen as 'detecting' where edges of H lie with respect to $\mathcal{P}$. However, the information conveyed by $L_{\mathcal{P}}(H)$ is by itself insufficient, as shown by the k-graph formed in Construction 1.6. Indeed, in that instance $L_{\mathcal{P}}(H)$ was complete, but some index vectors did not represent enough edges: specifically, H did not contain two disjoint edges with different numbers of vertices in A and B modulo 3. Thus in the proof of Theorem 1.10 we will frequently want to know which index vectors are represented by many edges; this is achieved by the following definition.

Definition 4.1 *(Robust edge-lattices).* Let H be a k-graph and $\mathcal{P}$ be a partition of $V(H)$ into d parts. Then for any $\mu > 0$,

(i) $I^{\mu}_{\mathcal{P}}(H)$ denotes the set of all $\mathbf{i} \in \mathbb{Z}^d$ such that at least $\mu|V(H)|^k$ edges $e \in H$ have $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$.
(ii) $L^{\mu}_{\mathcal{P}}(H)$ denotes the lattice in $\mathbb{Z}^d$ generated by $I^{\mu}_{\mathcal{P}}(H)$.

The constant μ can be viewed as the 'detection threshold': it specifies the number of edges of a given index which are required for this index to contribute to $L^{\mu}_{\mathcal{P}}(H)$. For any $\mu < \mu'$ it is clear that $L^{\mu'}_{\mathcal{P}}(H) \subseteq L^{\mu}_{\mathcal{P}}(H) \subseteq L_{\mathcal{P}}(H)$. Next we show that robust edge-lattices are insensitive to small perturbations.

Definition 4.2. Let $\mathcal{P}$ be a partition of a set V of size n and H be a k-graph on V.

(i) We say that a k-graph H' on a set $V' \subseteq V$ is *α-close* to H if we have $|V'| \geq (1-\alpha)n$ and $|H \triangle H'| \leq \alpha n^k$.
(ii) We say that a partition $\mathcal{P}'$ of a set $V' \subseteq V$ is *α-close* to $\mathcal{P}$ if for some d we have $|\mathcal{P}'| = |\mathcal{P}| = d$ and $\sum_{i \in [d]} |V_i \triangle V'_i| \leq \alpha n$, where $\mathcal{P} = (V_1, \ldots, V_d)$ and $\mathcal{P}' = (V'_1, \ldots, V'_d)$.

Note that Definition 4.2(i) is asymmetric; that is, it does not imply that H is α-close to H'.

Proposition 4.3. *Under the setup of Definition 4.2, we have $I^{\mu}_{\mathcal{P}}(H) \subseteq I^{\mu-2\alpha}_{\mathcal{P}'}(H')$, so $L^{\mu}_{\mathcal{P}}(H) \subseteq L^{\mu-2\alpha}_{\mathcal{P}'}(H')$. Also, if $\mu \leq 1/k$ then $I^{\mu}_{\mathcal{P}'}(H') \subseteq I^{\mu-3\alpha}_{\mathcal{P}}(H)$, so $L^{\mu}_{\mathcal{P}'}(H') \subseteq L^{\mu-3\alpha}_{\mathcal{P}}(H)$.*

Proof. Consider $\mathbf{i} \in I^{\mu}_{\mathcal{P}}(H)$. By definition, at least μn^k edges $e \in H$ have $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$. Of these edges, at most αn^k are not in H', and at most αn^k of those in H' have $\mathbf{i}_{\mathcal{P}'}(e) \neq \mathbf{i}$. Therefore $\mathbf{i} \in I^{\mu-2\alpha}_{\mathcal{P}'}(H')$. Similarly, consider $\mathbf{i} \in I^{\mu}_{\mathcal{P}'}(H')$. At least $\mu(n-\alpha n)^k > (\mu-\alpha)n^k$ edges $e \in H'$ have $\mathbf{i}_{\mathcal{P}'}(e) = \mathbf{i}$, at most αn^k are not in H, and at most αn^k have $\mathbf{i}_{\mathcal{P}}(e) \neq \mathbf{i}$. Therefore $\mathbf{i} \in I^{\mu-3\alpha}_{\mathcal{P}}(H)$. □

Let $\mathcal{P}$ and $\mathcal{P}'$ be partitions of a set V. We say that $\mathcal{P}'$ *refines* $\mathcal{P}$ if every part of $\mathcal{P}'$ is a subset of a part of $\mathcal{P}$.

Definition 4.4. Let $\mathcal{P}$ and $\mathcal{P}'$ be partitions of a set V such that $\mathcal{P}$ refines $\mathcal{P}'$. Let $\mathbf{i}$ be an index vector with respect to $\mathcal{P}$. Then the *projection* $(\mathbf{i} \mid \mathcal{P}')$ of $\mathbf{i}$ on $\mathcal{P}'$ is defined by

$$\big(\mathbf{i} \mid \mathcal{P}'\big)_W = \sum_{X \in \mathcal{P}, X \subseteq W} \mathbf{i}_X$$

for each $W \in \mathcal{P}'$. We also write $(I \mid \mathcal{P}') = \{(\mathbf{i} \mid \mathcal{P}') : \mathbf{i} \in I\}$ when I is a set of index vectors with respect to $\mathcal{P}$.

As an example under the setup of Definition 4.4, note that if $S \subseteq V$ then $(\mathbf{i}_{\mathcal{P}}(S) \mid \mathcal{P}') = \mathbf{i}_{\mathcal{P}'}(S)$.

Our next result concerns the behaviour of edge-lattices with respect to projection. An informal statement is that if $\mathcal{Q}$ refines $\mathcal{P}$, then the projection of a robust edge-lattice in $\mathbb{Z}^{\mathcal{Q}}$ is contained in a robust edge-lattice in $\mathbb{Z}^{\mathcal{P}}$, and a 'weak converse' holds, namely that any vector in a robust edge-lattice in $\mathbb{Z}^{\mathcal{P}}$ is the projection of some vector in a robust edge-lattice in $\mathbb{Z}^{\mathcal{Q}}$.

Proposition 4.5. *Let $\mathcal{P}$ and $\mathcal{Q}$ be partitions of a set V of n vertices such that $\mathcal{Q}$ refines $\mathcal{P}$, and let H be a k-graph on V.*

(i) *If $\mathbf{i} \in L^{\mu}_{\mathcal{Q}}(H)$ then $(\mathbf{i} \mid \mathcal{P}) \in L^{\mu}_{\mathcal{P}}(H)$.*
(ii) *For any $\mathbf{i} \in L^{\mu}_{\mathcal{P}}(H)$ there exists $\mathbf{i}' \in L^{\mu/m}_{\mathcal{Q}}(H)$ such that $(\mathbf{i}' \mid \mathcal{P}) = \mathbf{i}$, where $m := (k+1)^{|\mathcal{Q}|-|\mathcal{P}|}$.*

Proof. For (i), consider $\mathbf{i} \in I^{\mu}_{\mathcal{Q}}(H)$. At least μn^k edges $e \in H$ have $\mathbf{i}_{\mathcal{Q}}(e) = \mathbf{i}$. Each such edge has $\mathbf{i}_{\mathcal{P}}(e) = (\mathbf{i} \mid \mathcal{P})$, so $(\mathbf{i} \mid \mathcal{P}) \in I^{\mu}_{\mathcal{P}}(H)$. Since projection is linear the result follows.

For (ii), consider $\mathbf{i} \in I^{\mu}_{\mathcal{P}}(H)$. At least μn^k edges $e \in H$ have $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$. Note that there are at most m index vectors $\mathbf{i}'$ with respect to $\mathcal{Q}$ which satisfy $(\mathbf{i}' \mid \mathcal{P}) = \mathbf{i}$, since for every part $X \in \mathcal{P}$ which contains r parts $Y_1, \dots, Y_r$ of $\mathcal{Q}$ there are at most $(k+1)^{r-1}$ ways of partitioning i_X into $i'_{Y_1}, \dots, i'_{Y_r}$. So by the pigeonhole principle there is some $\mathbf{i}'$ with $(\mathbf{i}' \mid \mathcal{P}) = \mathbf{i}$ for which at least $\mu n^k/m$ edges $e \in H$ have $\mathbf{i}_{\mathcal{Q}}(e) = \mathbf{i}'$, that is, $\mathbf{i}' \in I^{\mu/m}_{\mathcal{Q}}(H)$. Again the result follows by linearity of projection. □

4.2. Hypergraph matching theory

In this subsection we describe a central theorem from [17] that plays a key role in this paper. First we need the following definitions. A *k-complex* J is a hypergraph such that $\emptyset \in J$, every edge of J has size at most k, and J is closed downwards, that is, if $e \in J$ and $e' \subseteq e$ then $e' \in J$. We write J_r to denote the r-graph on $V(J)$ formed by edges of size r in J. Also, we use the following notion of partite degree in a multipartite k-complex from [17]. Let $\mathcal{P}$ partition a set V into k parts $V_1, \dots, V_k$, and let J be a $\mathcal{P}$-partite k-complex on V. For each $0 \leq j \leq k-1$ the *partite minimum* j*-degree* $\delta_j^*(J)$ is defined to be the largest m such that any j-edge e has at least m extensions to a $(j+1)$-edge in any part not used by e, i.e.

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

The *partite degree sequence* is then $\delta^*(J) = (\delta_0^*(J), \dots, \delta_{k-1}^*(J))$. Note that we suppress the dependence on $\mathcal{P}$ in our notation: this will always be clear from the context. Minimum degree sequence conditions on a k-complex take the form $\delta^*(J) \geq (a_0, \dots, a_{k-1})$, where the inequality is to be interpreted pointwise. Finally, observe that if J is a $\mathcal{P}$-partite k-complex on V, and $\mathcal{Q}$ is a partition of V which refines $\mathcal{P}$ into d parts, then every edge $e \in J_k$ has $(\mathbf{i}_{\mathcal{Q}}(e) \mid \mathcal{P}) = \mathbf{i}_{\mathcal{P}}(e) = \mathbf{1}$. So we say that a lattice $L \subseteq \mathbb{Z}^d$ is *complete with respect to* $\mathcal{Q}$ if $\mathbf{i} \in L$ for every $\mathbf{i} \in \mathbb{Z}^d$ with $(\mathbf{i} \mid \mathcal{P}) = \mathbf{1}$, and *incomplete with respect to* $\mathcal{Q}$ otherwise. The following is a simplified form of [17, Theorem 2.13].

Theorem 4.6. *(See [17].) Suppose that* $1/n \ll \mu \ll \gamma, 1/k$. *Let* $\mathcal{P}'$ *partition a set* V *into parts* $V_1, \dots, V_k$ *each of size* n. *Suppose that* J *is a* $\mathcal{P}'$*-partite* k*-complex on* V *with*

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

Then either J_k *contains a perfect matching, or* J_k *is close to a divisibility barrier, in that there is some partition* $\mathcal{P}$ *of* $V(J)$ *into* $d \leq k^2$ *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* $\mathcal{P}'$ *and transferral-free.*

4.3. Convex geometry

Given points $\mathbf{v}_1, \dots, \mathbf{v}_r \in \mathbb{R}^d$, we define their *positive cone* as

$$PC\big(\{\mathbf{v}_1, \dots, \mathbf{v}_r\}\big) := \left\{\sum_{j \in [r]} \lambda_j \mathbf{v}_j : \lambda_1, \dots, \lambda_r \geq 0\right\}.$$

A classical result, commonly known as Farkas' Lemma, shows that any point $\mathbf{v}$ outside of the positive cone of these points can be separated from them by a separating hyperplane.

Theorem 4.7 *(Farkas' Lemma). Suppose* $\mathbf{v} \in \mathbb{R}^d \setminus PC(Y)$ *for some finite set* $Y \subseteq \mathbb{R}^d$. *Then there is some* $\mathbf{a} \in \mathbb{R}^d$ *such that* $\mathbf{a} \cdot \mathbf{y} \geq 0$ *for every* $\mathbf{y} \in Y$ *and* $\mathbf{a} \cdot \mathbf{v} < 0$.

We also need the following (slightly rephrased) simple proposition from [17].

Proposition 4.8. *(See [17], Proposition 4.8.) Suppose that* $1/s \ll 1/r, 1/d$. *Let* $X \subseteq \mathbb{Z}^d \cap B^d(\mathbf{0}, r)$, *and let* L_X *be the sublattice of* $\mathbb{Z}^d$ *generated by* X. *Then for any vector* $\mathbf{x} \in L_X \cap B^d(\mathbf{0}, r)$ *we may choose integers* $a_{\mathbf{i}}$ *with* $|a_{\mathbf{i}}| \leq s$ *for each* $\mathbf{i} \in X$ *such that* $\mathbf{x} = \sum_{\mathbf{i} \in X} a_{\mathbf{i}} \mathbf{i}$.

4.4. Baranyai's Matrix Rounding Theorem

The proof of Lemma 7.6 (on which the key lemma, Lemma 7.1, relies) will use Baranyai's Matrix Rounding Theorem [4] (see also [33, Theorem 7.5]).

Theorem 4.9 *(Baranyai's Matrix Rounding Theorem). Let* A *be a real matrix. Then there exists an integer matrix* B *whose entries, row sums, column sums and the sum of all the entries are the entries, row sums, column sums and the sum of all the entries respectively of* A, *each rounded either up or down.*

4.5. Concentration of probability

We will need the following inequalities, known as Chernoff bounds, as applied to sums of Bernoulli random variables (i.e. random variables which take values in $\{0, 1\}$) and hypergeometric random variables. The hypergeometric random variable X with parameters (N, m, n) is defined as $X = |T \cap S|$, where $S \subseteq [N]$ is a fixed set of size m, and $T \subseteq [N]$ is a uniformly random set of size n. If $m = pN$ then X has mean pn. The following is [13, Theorem 2.8].

Lemma 4.10. *Let* X *be a sum of independent Bernoulli random variables and* $0 < a < 3/2$. *Then* $\mathbb{P}(|X - \mathbb{E}X| \geq a\mathbb{E}X) \leq 2\exp\left(-\frac{a^2}{3}\mathbb{E}X\right)$.

Corollary 4.11. *Let* X *be a sum of independent hypergeometric random variables and* $0 < a < 3/2$. *Then* $\mathbb{P}(|X - \mathbb{E}X| \geq a\mathbb{E}X) \leq 2\exp\left(-\frac{a^2}{3}\mathbb{E}X\right)$.

Proof. We follow Remark 2.11 of [13]. By Lemma 1 of [34], each of the hypergeometric random variables may be expressed as a sum of independent (but not identically distributed) Bernoulli random variables. Hence Lemma 4.10 implies the result. □

We will also need a form of the well-known Azuma–Hoeffding inequality. A sequence $(X_0, \ldots, X_n)$ of random variables is a *martingale* if $\mathbb{E}(|X_j|)$ is finite and $\mathbb{E}(X_j | X_0, \ldots, X_{j-1}) = X_{j-1}$ for any $j \in [n]$. The following is (a weaker form of) [13, Theorem 2.25].

Theorem 4.12. *Let $A_0, A_1, \dots, A_n$ be a martingale such that $|A_i - A_{i-1}| \le C$ for every $i \in [n]$. Then $\mathbb{P}(|A_n - A_0| > tC) \le 2\exp(-t^2/2n)$ for any $t > 0$.*

We will apply Theorem 4.12 via the following corollary.

Corollary 4.13. *Let $c > 0$ and H be a k-graph on n vertices with $|H| \ge cn^k$. Let $\mathcal{Q}$ be a partition of $V(H)$ and let $(n_Z)_{Z\in\mathcal{Q}}$ be integers such that $c|Z| \le n_Z \le |Z|$ for each $Z \in \mathcal{Q}$. Suppose that $S \subseteq V(H)$ is chosen uniformly at random subject to the condition that $|S \cap Z| = n_Z$ for each $Z \in \mathcal{Q}$. Then with high probability $|H[S]| = (1 \pm c)\mathbb{E}|H[S]|$.*

Proof. We consider each $S \cap Z$ to be chosen as the first n_Z elements of a random permutation π_Z of Z. We apply Theorem 4.12 with $A_i = \mathbb{E}(|H[S]| \mid \mathcal{F}_i)$, where $\mathcal{F}_i$ is the algebra of events generated by revealing the values of the permutations on the first i elements of $V(H)$. Note that $A_0 = \mathbb{E}|H[S]|$ and $A_n = |H[S]|$. Also, $|H[S]|$ is '$2n^{k-1}$-Lipschitz', in that its value can change by at most $2n^{k-1}$ when any transposition is applied to any of the permutations. It is not hard to deduce that $|A_i - A_{i-1}| \le 2n^{k-1}$ for all $i \in [n]$. Finally, we claim that $\mathbb{E}|H[S]| \ge (c/2)^{k+1}n^k$. Indeed, at least $cn^k/2$ edges of H have at most one vertex within any part of size less than $cn/2$, and for any such edge e, we have $\mathbb{P}(e \subseteq S) \ge (c/2)^k$. So by Theorem 4.12, applied with $C = 2n^{k-1}$ and $t = (c/2)^{k+2}n$, with high probability $|H[S]|$ and $\mathbb{E}|H[S]|$ do not differ by more than $c\mathbb{E}|H[S]|$. □

5. Robust maximality

Given a k-graph H, there may be many approximate divisibility barriers, i.e. partitions $\mathcal{P}$ of $V(H)$ such that $L^\mu_{\mathcal{P}}(H)$ is incomplete for some constant $\mu > 0$; for example, any refinement of such a partition has this property. In this section we will identify a canonical such partition, which is 'robustly maximal'. The definition at which we will arrive has a number of advantages. Firstly, any approximate divisibility barrier in H will imply the existence of at least one robustly maximal partition $\mathcal{P}$. Secondly, we can give a condition which implies the existence of a perfect matching in H and refers to index vectors with respect to $\mathcal{P}$ alone. Thirdly, as the term 'robustly' suggests, the property is insensitive to small modifications. All of these properties will be stated precisely and proved later in the section.

To build up to the definition, we first discuss one way to avoid taking a partition that is too fine. Recall that a lattice $L \subseteq \mathbb{Z}^d$ is *transferral-free* if L does not contain any difference of unit vectors $\mathbf{u}_i - \mathbf{u}_j$ with $i, j \in [d]$ and $i \ne j$. Given any approximate divisibility barrier, we can repeatedly merge parts to obtain one with a transferral-free robust edge-lattice, which we may consider to be its 'simplest' version. Indeed, let H be a k-graph and consider a partition $\mathcal{P}$ of $V(H)$ such that $L^\mu_{\mathcal{P}}(H)$ is incomplete and contains $\mathbf{u}_X - \mathbf{u}_Y$, for some distinct parts X, Y of $\mathcal{P}$. Let $\mathcal{P}'$ be the partition formed from $\mathcal{P}$ by merging the parts X and Y. Then $L^{(k+1)\mu}_{\mathcal{P}'}(H)$ is also incomplete. Indeed, if this is not

the case, then for any $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ we have $(\mathbf{i} \mid \mathcal{P}') \in L^{(k+1)\mu}_{\mathcal{P}'}(H)$, so by Proposition 4.5(ii) there is $\mathbf{i}' \in L^{\mu}_{\mathcal{P}}(H)$ with $(\mathbf{i}' \mid \mathcal{P}') = (\mathbf{i} \mid \mathcal{P}')$. Since $\mathbf{u}_X - \mathbf{u}_Y \in L^{\mu}_{\mathcal{P}}(H)$ we deduce that $\mathbf{i} \in L^{\mu}_{\mathcal{P}}(H)$. However, this implies that $L^{\mu}_{\mathcal{P}}(H)$ is complete, which is a contradiction.

Next we recall from Construction 3.1 that a transferral-free approximate divisibility barrier may have another one 'hidden' inside it, so we will require a maximality property to rule this out. A first attempt might be to say that $\mathcal{P}$ should be maximal such that $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free, but then we would obtain a rather fragile property that is sensitive to small modifications of $\mathcal{P}$, μ and H. Indeed, in the course of our proof we will need to remove small matchings from H in order to cover vertices which are exceptional in various ways. We also want the property to be preserved with high probability by taking an induced subgraph on a randomly chosen set of vertices. These requirements lead naturally to the following key definition.

Definition 5.1. Let H be a k-graph. We say that a partition $\mathcal{P}$ of $V(H)$ is (c, c', μ, μ')-*robustly maximal* with respect to H if

(i) $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free and all parts of $\mathcal{P}$ have size at least $c|V(H)|$,
(ii) $L^{\mu'}_{\mathcal{P}'}(H)$ is not transferral-free for any partition $\mathcal{P}'$ of $V(H)$ with parts of size at least $c'|V(H)|$ that strictly refines $\mathcal{P}$.

We will see in the key lemma that robustly maximal partitions are 'canonical', in that they capture all necessary information on approximate divisibility barriers. The next proposition allows us to refine any transferral-free approximate divisibility barrier to obtain a robustly maximal partition. (Note that if $\mathcal{P}$ is trivial and $\mathcal{P}' = \mathcal{P}$ then H does not have any approximate divisibility barrier.)

Proposition 5.2. *Let $k \geq 2$ be an integer and $c > 0$ be a constant. Let $s = \lfloor 1/c \rfloor$ and fix constants $0 < \mu_1 < \cdots < \mu_{s+1}$ and $c_1, \ldots, c_{s+1} \geq c$. Suppose that H is a k-graph on n vertices, and $\mathcal{P}$ is a partition of $V(H)$ with parts of size at least $c_1 n$ such that $L^{\mu_1}_{\mathcal{P}}(H)$ is transferral-free. Then there exists $t \in [s]$ and a partition $\mathcal{P}'$ of $V(H)$ that refines $\mathcal{P}$ and is $(c_t, c_{t+1}, \mu_t, \mu_{t+1})$-robustly maximal with respect to H.*

Proof. We start by setting $\mathcal{P}^{(1)} = \mathcal{P}$, then we proceed iteratively. At step t, if there is a partition $\mathcal{P}^+$ which strictly refines $\mathcal{P}^{(t)}$ into parts of size at least $c_{t+1}n$ such that the lattice $L^{\mu_{t+1}}_{\mathcal{P}^+}(H)$ is transferral-free, then we let $\mathcal{P}^{(t+1)} = \mathcal{P}^+$. Otherwise we terminate with output $\mathcal{P}^{(t)}$. By definition each $\mathcal{P}^{(t)}$ refines $\mathcal{P}$ and has parts of size at least $c_t n$. Furthermore, if the algorithm terminates at time t, then $L^{\mu_t}_{\mathcal{P}^{(t)}}(H)$ is transferral-free by choice of $\mathcal{P}^{(t)}$, but for any $\mathcal{P}^+$ which strictly refines $\mathcal{P}^{(t)}$ into parts of size at least $c_{t+1}n$ the lattice $L^{\mu_{t+1}}_{\mathcal{P}^+}(H)$ is not transferral-free. That is, $\mathcal{P}^{(t)}$ is $(c_t, c_{t+1}, \mu_t, \mu_{t+1})$-robustly maximal with respect to H. It remains to show that this algorithm must terminate with $t \leq 1/c$. Indeed, $|\mathcal{P}^{(i)}| > |\mathcal{P}^{(i-1)}|$ for $i \geq 2$, but $\mathcal{P}^{(t)}$ can have at most $1/c$ parts, since each part has size at least $c_t n \geq cn$. □

Next we show that robust maximality is insensitive to small modifications.

Lemma 5.3. *Suppose that* $1/n \ll \mu, \mu', \alpha \ll c', c, 1/k$. *Let* $\mathcal{P}$ *be a partition of a set* V *of size* n *and* H *be a* k*-graph on* V. *Let* $V' \subseteq V$, *let* $\mathcal{P}'$ *be a partition of* V' *that is* α*-close to* $\mathcal{P}$, *and let* H' *be a* k*-graph on* V' *that is* α*-close to* H. *If* $\mathcal{P}$ *is* (c, c', μ, μ')*-robustly maximal with respect to* H *then* $\mathcal{P}'$ *is* $(c-\alpha, c'+3\alpha, \mu+3\alpha, \mu'-2\alpha)$*-robustly maximal with respect to* H'.

Proof. Since $\mathcal{P}$ has parts of size at least cn, and $\mathcal{P}'$ is α-close to $\mathcal{P}$, each part of $\mathcal{P}'$ has size at least $(c-\alpha)n$. Also, Proposition 4.3 implies that $L^{\mu+3\alpha}_{\mathcal{P}'}(H') \subseteq L^{\mu}_{\mathcal{P}}(H)$. Since $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free, so is $L^{\mu+3\alpha}_{\mathcal{P}'}(H')$. All that remains is to show that there is no partition $\mathcal{P}^+$ of V' which strictly refines $\mathcal{P}'$ into parts of size at least $(c'+3\alpha)|V'|$ such that the lattice $L^{\mu'-2\alpha}_{\mathcal{P}^+}(H')$ is transferral-free.

Suppose for a contradiction that some such $\mathcal{P}^+$ exists, with d parts $\mathcal{P}^+_1, \dots, \mathcal{P}^+_d$. We use $\mathcal{P}^+$ to form a partition $\mathcal{P}^*$ of V whose parts $\mathcal{P}^*_1, \dots, \mathcal{P}^*_d$ correspond to those of $\mathcal{P}^+$. Note that all but at most αn vertices $u \in V(H)$ are members of $V(H')$ which lie in the same part of $\mathcal{P}'$ as $\mathcal{P}$. We include each such vertex u in the part $\mathcal{P}^*_j$ corresponding to the part $\mathcal{P}^+_j$ of $\mathcal{P}^+$ which contains u. Next, we assign each of the at most αn remaining vertices $v \in V(H)$ to an arbitrary part of $\mathcal{P}^*$ which is a subset of the part of $\mathcal{P}$ containing v. Observe that $\mathcal{P}^*$ is then a refinement of $\mathcal{P}$. Furthermore, $\mathcal{P}^+$ is α-close to $\mathcal{P}^*$, so $L^{\mu'}_{\mathcal{P}^*}(H) \subseteq L^{\mu'-2\alpha}_{\mathcal{P}^+}(H')$ by Proposition 4.3, and this implies that $L^{\mu'}_{\mathcal{P}^*}(H)$ is transferral-free. However, $\mathcal{P}^*$ has parts of size at least $(c'+3\alpha)|V'| - \alpha n \geq c'n$, so the existence of $\mathcal{P}^*$ contradicts the robust maximality of $\mathcal{P}$ with respect to H, completing the proof. □

Our next result shows that with robust maximality we can improve the 'weak converse' of projection from Proposition 4.5 to a (genuine) converse (with weaker parameters).

Proposition 5.4. *Suppose that* $1/n \ll \mu \ll \mu' \ll c', c, 1/k$. *Let* H *be a* k*-graph on* n *vertices,* $\mathcal{P}$ *be a partition of* $V(H)$ *that is* (c, c', μ, μ')*-robustly maximal with respect to* H, *and* $\mathcal{Q}$ *be a partition of* $V(H)$ *which refines* $\mathcal{P}$ *into parts of size at least* $c'n$. *Suppose that* $\mathbf{i} \in L^{\mu'}_{\mathcal{P}}(H)$. *Then* $\mathbf{i}' \in L^{\mu}_{\mathcal{Q}}(H)$ *for any index vector* $\mathbf{i}'$ *with respect to* $\mathcal{Q}$ *such that* $(\mathbf{i}' \mid \mathcal{P}) = \mathbf{i}$.

Proof. Let p be the number of parts of $\mathcal{P}$. We prove the following statement by induction on q: if $\mathcal{Q}$ is a refinement of $\mathcal{P}$ into q parts of size at least $c'n$ and $\mathbf{i}'$ is an index vector with respect to $\mathcal{Q}$ such that $(\mathbf{i}' \mid \mathcal{P}) = \mathbf{i}$ then $\mathbf{i}' \in L^{\mu'/(k+1)^{q-p}}_{\mathcal{Q}}(H)$. The base case $q = p$ is trivial, since then we must have $\mathcal{Q} = \mathcal{P}$, so $\mathbf{i}' = \mathbf{i} \in L^{\mu'}_{\mathcal{Q}}(H)$ by assumption. Assume therefore that we have proved the statement for any refinement $\mathcal{Q}'$ of $\mathcal{P}$ into $q-1$ parts of size at least $c'n$, and that $\mathcal{Q}$ refines $\mathcal{P}$ into q parts of size at least $c'n$. Since $\mathcal{P}$ is (c, c', μ, μ')-robustly maximal we know that $L^{\mu'}_{\mathcal{Q}}(H)$ is not transferral-free, and so there are distinct parts $X, X' \in \mathcal{Q}$ such that $\mathbf{u}_X - \mathbf{u}_{X'} \in L^{\mu'}_{\mathcal{Q}}(H)$. Note that

$(\mathbf{u}_X - \mathbf{u}_{X'} \mid \mathcal{P}) \in L^{\mu'}_{\mathcal{P}}(H)$ by Proposition 4.5(i); since $L^{\mu'}_{\mathcal{P}}(H)$ is transferral-free, this implies that $(\mathbf{u}_X - \mathbf{u}_{X'} \mid \mathcal{P}) = \mathbf{0}$ and hence that X and X' are contained in the same part of $\mathcal{P}$. Let $\mathcal{Q}'$ be formed from $\mathcal{Q}$ by merging X and X' into a single part. Then $\mathcal{Q}'$ is a refinement of $\mathcal{P}$ into $q-1$ parts of size at least $c'n$, so $(\mathbf{i}' \mid \mathcal{Q}') \in L^{\mu'/(k+1)^{q-1-p}}_{\mathcal{Q}'}(H)$ by our inductive hypothesis. By Proposition 4.5(ii) there exists $\mathbf{i}^* \in L^{\mu'/(k+1)^{q-p}}_{\mathcal{Q}}(H)$ with $(\mathbf{i}^* \mid \mathcal{Q}') = (\mathbf{i}' \mid \mathcal{Q}')$. Now $\mathbf{i}^*$ differs from $\mathbf{i}'$ only in the co-ordinates corresponding to X and X'; as $\mathbf{u}_X - \mathbf{u}_{X'} \in L^{\mu'}_{\mathcal{Q}}(H) \subseteq L^{\mu'/(k+1)^{q-p}}_{\mathcal{Q}}(H)$ it follows that $\mathbf{i}' \in L^{\mu'/(k+1)^{q-p}}_{\mathcal{Q}}(H)$, completing the induction. Since $\mathcal{Q}$ as in the statement can have at most $1/c'$ parts, we conclude that $\mathbf{i}' \in L^{\mu'/(k+1)^{1/c'}}_{\mathcal{Q}}(H) \subseteq L^{\mu}_{\mathcal{Q}}(H)$. □

Another important property of robust maximality is that it is preserved by random selection.

Lemma 5.5. *Suppose that $1/n \ll \mu \ll \mu' \ll c', c \ll \eta, 1/k$. Let H be a k-graph on n vertices and $\mathcal{P}$ be a partition of $V(H)$ that is (c, c', μ, μ')-robustly maximal with respect to H. Let $\mathcal{P}'$ be a partition of $V(H)$ that refines $\mathcal{P}$ and let $(n_Z)_{Z\in\mathcal{P}'}$ be integers such that $\eta|Z| \le n_Z \le |Z|$ for each $Z \in \mathcal{P}'$. Suppose that $S \subseteq V(H)$ is chosen uniformly at random subject to the condition that $|S \cap Z| = n_Z$ for each $Z \in \mathcal{P}'$. Then with high probability $\mathcal{P}[S]$ is $(\eta c, 3c'/\eta, \mu/c, (\mu')^3)$-robustly maximal with respect to $H[S]$.*

The proof of this lemma requires the weak hypergraph regularity lemma, so we defer the details to Section 9.

6. Fullness

This section develops the theory of full lattices. In the first subsection we give a characterisation in terms of finite abelian groups. In the second subsection we give an equivalent definition of solubility, via an application of the pigeonhole principle that will also be useful in later sections. In the last subsection we consider a partite form of fullness that will be needed for the partite form of our key lemma.

6.1. The structure of full lattices

In this subsection we characterise full lattices in terms of finite abelian groups. Recall that a set I of k-vectors of dimension d is full if for every $(k-1)$-vector $\mathbf{v}$ there is some $i \in [d]$ such that $\mathbf{v} + \mathbf{u}_i \in I$. Recall also that a lattice L is full if it contains a full set I of k-vectors and is transferral-free. We start by showing that L is generated by I, so is an edge-lattice.

Lemma 6.1. *Suppose $k \ge 3$ and L is a full lattice in $\mathbb{Z}^{\mathcal{P}}$, where $\mathcal{P}$ is a partition of a set V. Let $I \subseteq L$ be a full set of k-vectors and let L' be the lattice generated by I. Then*

(i) *for any* $X_1, X_1', X_2 \in \mathcal{P}$, *there exists* $X_2' \in \mathcal{P}$ *such that* $\mathbf{u}_{X_1} + \mathbf{u}_{X_2} - \mathbf{u}_{X_1'} - \mathbf{u}_{X_2'} \in L'$,
(ii) *for any* $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ *and* $X \in \mathcal{P}$ *there is* $X' \in \mathcal{P}$ *such that* $\mathbf{i} - \mathbf{u}_X + \mathbf{u}_{X'} \in L'$, *and*
(iii) $L = L'$ *is an edge-lattice.*

Proof. To prove (i), we start by fixing any $(k-3)$-vector $\mathbf{i}'$. Since I is full, we can find $Y \in \mathcal{P}$ such that $\mathbf{i}' + \mathbf{u}_{X_1} + \mathbf{u}_{X_2} + \mathbf{u}_Y \in I$. Similarly, we can find $X_2' \in \mathcal{P}$ such that $\mathbf{i}' + \mathbf{u}_{X_1'} + \mathbf{u}_Y + \mathbf{u}_{X_2'} \in I$. Then $\mathbf{u}_{X_1} + \mathbf{u}_{X_2} - \mathbf{u}_{X_1'} - \mathbf{u}_{X_2'}$ is the difference of these two index vectors and hence lies in L'.

For (ii), consider $\mathbf{i}' \in L'$ that minimises $\sum_{Z\in\mathcal{P}} |i'_Z - i_Z|$ subject to $\sum_{Z\in\mathcal{P}} i'_Z = \sum_{Z\in\mathcal{P}} i_Z$. We claim that $\sum_{Z\in\mathcal{P}} |i'_Z - i_Z| \le 2$. Indeed, if not we may choose X_1, X_1', X_2 such that either

(a) $X_1 \neq X_2$, $i_{X_1} - i'_{X_1} > 0$, $i_{X_2} - i'_{X_2} > 0$ and $i_{X_1'} - i'_{X_1'} < 0$, or
(b) $X_1 = X_2$, $i_{X_1} - i'_{X_1} > 1$ and $i_{X_1'} - i'_{X_1'} < 0$.

In either case, we may apply (i) to choose $X_2' \in \mathcal{P}$ such that $\mathbf{i}^* = \mathbf{u}_{X_1} + \mathbf{u}_{X_2} - \mathbf{u}_{X_1'} - \mathbf{u}_{X_2'} \in L'$. Then $\mathbf{i}'+\mathbf{i}^*$ contradicts our choice of $\mathbf{i}'$. Thus the claim holds, so $\mathbf{i}' = \mathbf{i}-\mathbf{u}_Y+\mathbf{u}'_Y$ for some $Y, Y' \in \mathcal{P}$. By (i) again, we can choose X' such that $\mathbf{i}^{**} = \mathbf{u}_X + \mathbf{u}_Y - \mathbf{u}_{X'} - \mathbf{u}_{Y'} \in L'$. Therefore $\mathbf{i} + \mathbf{u}_X - \mathbf{u}_{X'} = \mathbf{i}' + \mathbf{i}^{**} \in L'$, as claimed.

To see (iii), consider any $\mathbf{i} \in L$. By (ii) we have $\mathbf{i}' = \mathbf{i} - \mathbf{u}_X + \mathbf{u}_{X'} \in L'$ for some parts X and X'. Then $\mathbf{u}_X - \mathbf{u}_{X'} = \mathbf{i} - \mathbf{i}' \in L$. Since L is transferral-free we have $X = X'$. Therefore $\mathbf{i} = \mathbf{i}' \in L'$. □

Full lattices have the following maximality property.

Proposition 6.2. *Suppose* L *and* L' *are edge-lattices in* $\mathbb{Z}^d$ *such that* L *is full,* L' *is transferral-free and* $L \subseteq L'$. *Then* $L = L'$.

Proof. Let I and I' be full generating sets for L and L'. Suppose that $L \neq L'$, so $I \neq I'$. Choose $\mathbf{i} \in I' \setminus I$ and write $\mathbf{i} = \mathbf{i}' + \mathbf{u}_i$, where $\mathbf{i}'$ is a $(k-1)$-vector and $i \in [d]$. Since I is full, we have $\mathbf{i}' + \mathbf{u}_j \in I$ for some $j \in [d]$. Note that $j \neq i$, as $\mathbf{i} \notin I$. But $I \subseteq L'$, so L' contains the transferral $\mathbf{u}_i - \mathbf{u}_j = \mathbf{i} - (\mathbf{i}' + \mathbf{u}_j)$. This contradiction shows that $L = L'$. □

Next we define a group of cosets that is naturally associated with any lattice.

Definition 6.3. Suppose L is an edge-lattice in $\mathbb{Z}^{\mathcal{P}}$, where $\mathcal{P}$ is a partition of a set V.

(i) The *coset group* of $(\mathcal{P}, L)$ is $G = G(\mathcal{P}, L) = L^{\mathcal{P}}_{\max}/L$.
(ii) For any $\mathbf{i} \in L^{\mathcal{P}}_{\max}$, the *residue* of $\mathbf{i}$ in G is $R_G(\mathbf{i}) = \mathbf{i} + L$. For any $A \subseteq V$ of size divisible by k, the *residue* of A in G is $R_G(A) = R_G(\mathbf{i}_{\mathcal{P}}(A))$.

Next we show that $|G(\mathcal{P}, L)| = |\mathcal{P}|$ when L is full and $k \ge 3$. Note that this may not hold if $k = 2$, as shown by the example where $|\mathcal{P}| = 2$ and L is generated by $(1, 1)$.

Lemma 6.4. *Suppose $k \geq 3$ and L is a full lattice in $\mathbb{Z}^{\mathcal{P}}$, where $\mathcal{P}$ is a partition of a set V. Then $|G(\mathcal{P}, L)| = |\mathcal{P}|$.*

Proof. Fix any $(k-1)$-vector $\mathbf{i}'$. We claim that every coset $L + \mathbf{v}$ of L in $L^{\mathcal{P}}_{\max}$ contains an index vector $\mathbf{i}' + \mathbf{u}_Y$ for some $Y \in \mathcal{P}$. To see this, note that since L is full there exists $X \in \mathcal{P}$ such that $\mathbf{i}' + \mathbf{u}_X \in L$. Also, by Lemma 6.1(ii) we can choose $Y \in \mathcal{P}$ such that $-\mathbf{v} - \mathbf{u}_X + \mathbf{u}_Y \in L$. But now $\mathbf{i}' + \mathbf{u}_Y = (\mathbf{i}' + \mathbf{u}_X) + (-\mathbf{v} - \mathbf{u}_X + \mathbf{u}_Y) + \mathbf{v} \in L + \mathbf{v}$, as claimed. Furthermore, $L + \mathbf{v}$ cannot contain $\mathbf{i}' + \mathbf{u}_Y$ and $\mathbf{i}' + \mathbf{u}_{Y'}$ for distinct parts Y, Y' of $\mathcal{P}$, as then $\mathbf{u}_Y - \mathbf{u}_{Y'} \in L$, contradicting the fact that L is transferral-free. Therefore $|G(\mathcal{P}, L)| = |\mathcal{P}|$. □

Now we describe the structure of L in terms of its coset group. If G is an abelian group, $g \in G$ and r is a non-negative integer then we write rg for the sum of r copies of g. We show that any full lattice arises from the following construction.

Construction 6.5. Let G be an abelian group. Let $\mathcal{P}$ be a partition of a set V into $|G|$ parts, identified with G. Fix $g_0 \in G$. Let $I(G, g_0)$ be the set of k-vectors $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ with

$$\sum_{g \in G} i_g g = g_0.$$

Let $L(G, g_0)$ be the lattice generated by $I(G, g_0)$.

Lemma 6.6. *$I(G, g_0)$ and $L(G, g_0)$ are full, and $L(G, g_0)$ is the set of index vectors $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ with*

$$\sum_{g \in G} i_g g = \left(k^{-1} \sum_{g \in G} i_g\right) g_0.$$

Proof. First we show that $I(G, g_0)$ is full, i.e. that for every $(k-1)$-vector $\mathbf{v}$ there is $h \in G$ such that $\mathbf{v} + \mathbf{u}_h \in I(G, g_0)$. Indeed, we can take $h = g_0 - \sum_{g \in G} v_g g$. Next let L' be the lattice consisting of all $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ with $\sum_{g \in G} i_g g = (k^{-1} \sum_{g \in G} i_g) g_0$. Note that $L(G, g_0)$ is contained in L'. Furthermore, L' is transferral-free, as for any distinct $g_1, g_2 \in G$ we have $\sum_{g \in G} (\mathbf{u}_{g_1} - \mathbf{u}_{g_2})_g g = g_1 - g_2 \neq 0_G = 0 g_0$, so $\mathbf{u}_{g_1} - \mathbf{u}_{g_2} \notin L'$. Thus $L(G, g_0)$ is transferral-free, and so is full. Finally, $L(G, g_0) = L'$ by Proposition 6.2. □

Theorem 6.7. *Let $k \geq 3$ and suppose L is a full edge-lattice in $\mathbb{Z}^{\mathcal{P}}$, where $\mathcal{P}$ is a partition of a set V. Then there is an identification of $\mathcal{P}$ with $G = G(\mathcal{P}, L)$ and some $g_0 \in G$ such that $L = L(G, g_0)$ and $R_G(\mathbf{i}) = \sum_{g \in G} i_g g - (k^{-1} \sum_{g \in G} i_g) g_0$ for any $\mathbf{i} \in L^{\mathcal{P}}_{\max}$.*

Proof. We fix an arbitrary part $X^0 \in \mathcal{P}$ and identify X^0 with the identity $0_G \in G$. Next we identify each $X \in \mathcal{P}$ with $R_G(\mathbf{u}_X - \mathbf{u}_{X^0})$. Note that for distinct parts X, X' of $\mathcal{P}$ we

have $R_G(\mathbf{u}_X - \mathbf{u}_{X^0}) \neq R_G(\mathbf{u}_{X'} - \mathbf{u}_{X^0})$, otherwise we would have $\mathbf{u}_X - \mathbf{u}_{X'} \in L$, contradicting the fact that L is transferral-free. Furthermore, the identification is bijective by Lemma 6.4.

Let $g_0 = -k\mathbf{u}_{X^0} + L \in G$. Consider any $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ and write $r = k^{-1}\sum_{g\in G} i_g \in \mathbb{Z}$. Then $\mathbf{i} - rk\mathbf{u}_{X^0} = \sum_{X\in\mathcal{P}} i_X(\mathbf{u}_X - \mathbf{u}_{X^0}) \in \sum_{g\in G} i_g g$. Now if $\sum_{g\in G} i_g g = rg_0$ then $\mathbf{i} \in L$. This shows that $L(G, g_0) \subseteq L$. Furthermore, $L(G, g_0)$ is full by Lemma 6.6 and L is transferral-free, so $L = L(G, g_0)$ by Proposition 6.2. Finally, we saw above that $R_G(\mathbf{i} - rk\mathbf{u}_{X^0}) = \sum_{g\in G} i_g g$. Since $R_G(-k\mathbf{u}_{X^0}) = g_0$, we have $R_G(\mathbf{i}) = \sum_{g\in G} i_g g - rg_0$. □

Remark 6.8. The identification of $\mathcal{P}$ with $G = G(\mathcal{P}, L)$ in the proof of Theorem 6.7 is determined up to translation by G, and the element g_0 is determined up to translation by kG. To see this, consider two identifications as in Theorem 6.7, say $\pi^i : \mathcal{P} \to G$ defined by $\pi^i(X) = R_G(\mathbf{u}_X - \mathbf{u}_{X^i})$ for some $X^i \in \mathcal{P}$ for $i = 0, 1$. Then for any $X \in \mathcal{P}$ we have $\pi^1(X) - \pi^0(X) = g'$, where $g' = R_G(\mathbf{u}_X - \mathbf{u}_{X^1}) - R_G(\mathbf{u}_X - \mathbf{u}_{X^0}) = R_G(\mathbf{u}_{X^0} - \mathbf{u}_{X^1}) \in G$ is independent of X. Now set $g_0 = -k\mathbf{u}_{X^0} + L \in G$ and $g_1 = -k\mathbf{u}_{X^1} + L \in G$; then $g_1 - g_0 = kR_G(\mathbf{u}_{X^0} - \mathbf{u}_{X^1}) = kg'$, as claimed.

6.2. Solubility

Recall that a full pair $(\mathcal{P}, L)$ for a k-graph H is soluble if there exists a matching M in H of size at most $|\mathcal{P}|-1$ such that $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M)) \in L$. Here we show that omitting the size condition on M gives an equivalent condition.

Lemma 6.9. *Let $(\mathcal{P}, L)$ be a full pair for a k-graph H, where $k \geq 3$. Then $(\mathcal{P}, L)$ is soluble if and only if there exists a matching M in H such that $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M)) \in L$.*

The proof uses the following application of the pigeonhole principle.

Proposition 6.10. *Let $G = (X, +)$ be an abelian group of order m, and suppose that elements $x_i \in X$ for $i \in [r]$ are such that $\sum_{i\in[r]} x_i = x'$. Then $\sum_{i\in I} x_i = x'$ for some $I \subseteq [r]$ with $|I| \leq m - 1$.*

Proof. It suffices to show that if $r > m-1$ then $\sum_{i\in[r]} x_i = \sum_{i\in I} x_i$ for some $I \subseteq [r]$ with $|I| < r$. To see this, note that there are $r + 1 > m$ partial sums $\sum_{i\in[j]} x_i$ for $0 \leq j \leq r$, so by the pigeonhole principle some two must be equal, that is, there exists $j_1 < j_2$ so that $\sum_{i\in[j_1]} x_i = \sum_{i\in[j_2]} x_i$. Then

$$\sum_{i\in[r]} x_i = \sum_{i\in[r]\setminus\{j_1+1,\dots,j_2\}} x_i + \sum_{i\in[j_2]} x_i - \sum_{i\in[j_1]} x_i = \sum_{i\in[r]\setminus\{j_1+1,\dots,j_2\}} x_i,$$

as required. □

Proof of Lemma 6.9. If $(\mathcal{P}, L)$ is soluble then such a matching M exists by definition. Conversely, suppose such a matching M exists. Write $G = G(\mathcal{P}, L)$ and note

that $|G| = |\mathcal{P}|$ by Lemma 6.4. Since $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M)) \in L$ we have $\sum_{e \in M} R_G(e) = R_G(V(H))$. By Proposition 6.10 there exists a submatching M' of M of size at most $|G| - 1 = |\mathcal{P}| - 1$ such that $\sum_{e \in M'} R_G(e) = R_G(V(H))$. Hence $R_G(V(H) \setminus V(M')) = 0$, i.e. $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M')) \in L$. □

6.3. Partite fullness

In the partite form of the key lemma we need the following partite form of fullness, in which we consider index vectors of sets that are partite with respect to some fixed partition.

Definition 6.11. Let $\mathcal{P}'$ be a partition of a set V into k parts and $\mathcal{P}$ be a refinement of $\mathcal{P}'$.

(i) We say that a set I of k-vectors with respect to $\mathcal{P}$ is *full* with respect to $\mathcal{P}'$ if $(\mathbf{i} \mid \mathcal{P}') = \mathbf{1}$ for every $\mathbf{i} \in I$, and for every $X \in \mathcal{P}'$ and $(k-1)$-vector $\mathbf{v}$ with respect to $\mathcal{P}$ such that $(\mathbf{v} \mid \mathcal{P}') = \mathbf{1} - \mathbf{u}_X$, there is some $Y \in \mathcal{P}$ with $Y \subseteq X$ such that $\mathbf{v} + \mathbf{u}_Y \in I$.
(ii) We write $L^{\mathcal{P}\mathcal{P}'}_{\max}$ for the lattice of vectors $\mathbf{i} \in \mathbb{Z}^{\mathcal{P}}$ such that $(\mathbf{i} \mid \mathcal{P}')$ is a multiple of $\mathbf{1}$.
(iii) We say that a lattice $L \subseteq L^{\mathcal{P}\mathcal{P}'}_{\max}$ is *full* with respect to $\mathcal{P}'$ if it is transferral-free and contains a set of k-vectors that is full with respect to $\mathcal{P}'$.

The next proposition records some properties of partite fullness. We just give the proofs of (i) and (ii), as the proofs of (iii)–(v) are very similar to those of Lemmas 6.1 and 6.4.

Proposition 6.12. *Let $k \geq 3$, let $\mathcal{P}'$ be a partition of a set V into k parts, $\mathcal{P}$ be a refinement of $\mathcal{P}'$, and $L \subseteq L^{\mathcal{P}\mathcal{P}'}_{\max}$ be a full lattice with respect to $\mathcal{P}'$. Let $I \subseteq L$ be a set of k-vectors that is full with respect to $\mathcal{P}'$ and let L' be the lattice generated by I. Then the following properties hold for some integer r.*

(i) *Each part of $\mathcal{P}'$ is refined into exactly r parts by $\mathcal{P}$.*
(ii) *For every part X of $\mathcal{P}$, there are exactly r^{k-2} vectors $\mathbf{i} \in I$ such that $i_X = 1$. In particular, $|I| = r^{k-1}$.*
(iii) *For any $\mathbf{i} \in L^{\mathcal{P}\mathcal{P}'}_{\max}$ and $X \in \mathcal{P}$, there is $X' \in \mathcal{P}$ such that $\mathbf{i} - \mathbf{u}_X + \mathbf{u}_{X'} \in L'$.*
(iv) *$|L^{\mathcal{P}\mathcal{P}'}_{\max}/L| = r$.*
(v) *$L = L'$ is an edge-lattice.*

Proof of (i) and (ii). Fix $X \in \mathcal{P}'$ and let r be the number of parts into which $\mathcal{P}$ refines X. For any other $X' \in \mathcal{P}'$, we will construct a bijection between the parts of $\mathcal{P}$ contained in X and those contained in X'. Let $\mathbf{v}$ be a non-negative index vector with respect to $\mathcal{P}$ such that $(\mathbf{v} \mid \mathcal{P}') = \mathbf{1} - \mathbf{u}_X - \mathbf{u}_{X'}$. For each $Y \in \mathcal{P}$ which is contained in X, we apply the property that I is full to the vector $\mathbf{v} + \mathbf{u}_Y$ to obtain a part $Y' \in \mathcal{P}$ which is contained

in X', such that $\mathbf{v}+\mathbf{u}_Y+\mathbf{u}_{Y'}\in I$. Now observe that since L is transferral-free, Y' must be unique. Thus $Y\mapsto Y'$ is a bijection, so (i) holds. For (ii), we observe that the number of such $(k-2)$-vectors $\mathbf{v}$ is r^{k-2} and there is a one-to-one correspondence between them and vectors $\mathbf{i}\in I$ such that $i_Y=1$. By considering the r parts of $\mathcal{P}$ contained in X we obtain $|I|=r^{k-1}$. □

Finally, we require the following consequence of Farkas' Lemma.

Proposition 6.13. *Let $\mathcal{P}'$ partition a set V into parts $V_1,\dots,V_k$ each of size n, and let $\mathcal{P}$ be a partition refining $\mathcal{P}'$ into parts of size at least n/k. Suppose I is a set of k-vectors with respect to $\mathcal{P}$ which is full with respect to $\mathcal{P}'$. Then $\mathbf{i}_{\mathcal{P}}(V)\in PC(I)$.*

Proof. Suppose for a contradiction that $\mathbf{i}_{\mathcal{P}}(V)\notin PC(I)$. Then by Theorem 4.7 we may fix $\mathbf{a}\in\mathbb{R}^{|\mathcal{P}|}$ such that $\mathbf{a}\cdot\mathbf{i}\geq 0$ for every $\mathbf{i}\in I$ and $\mathbf{a}\cdot\mathbf{i}_{\mathcal{P}}(V)<0$. For each $i\in[k]$, let $X_i^1,\dots,X_i^{b_i}$ be the parts of $\mathcal{P}$ which are subsets of V_i, and let $a_i^1,\dots,a_i^{b_i}$ be the corresponding coordinates of $\mathbf{a}$, with the labels chosen so that $a_i^1\leq\cdots\leq a_i^{b_i}$. Fix $i\in[k]$ for which $a_i^{b_i}-a_i^1$ is minimised, so in particular $a_i^{b_i}-a_i^1\leq\frac{1}{k}\sum_{j\in[k]}a_j^{b_j}-a_j^1$. By assumption we may choose $\mathbf{i}\in I$ such that $\mathbf{i}=\mathbf{u}_{X_i^s}+\sum_{j\neq i}\mathbf{u}_{X_j^1}$ for some $s\in[b_i]$. Then

$$\begin{aligned}0>\mathbf{a}\cdot\mathbf{i}_{\mathcal{P}}(V)&\geq\sum_{j\in[k]}na_j^1+\frac{n}{k}\big(a_j^{b_j}-a_j^1\big)\\&\geq n\bigg(\sum_{j\in[k]}a_j^1+\big(a_i^{b_i}-a_i^1\big)\bigg)\geq n\mathbf{a}\cdot\mathbf{i}\geq 0,\end{aligned}$$

a contradiction. So $\mathbf{i}_{\mathcal{P}}(V)\in PC(I)$. □

7. The key lemmas

The following key lemma will be used in the proof of Theorem 1.10. It provides a simple condition for finding a perfect matching under Setup 1.7 when we have a robustly maximal partition: it suffices that the index of the vertex set is in the robust edge-lattice and every vertex is in many edges with index in the robust edge-lattice.

Lemma 7.1. *Suppose that $k\geq 3$ and $1/n\ll\varepsilon\ll\mu\ll\mu'\ll c,d\ll\gamma,1/k$. Let H be a k-graph on a set V of size kn and $\mathcal{P}$ be a partition of V. Suppose that*

(i) *at most εn^{k-1} $(k-1)$-sets $S\subseteq V$ have $d_H(S)<(1+\gamma)n$,*
(ii) *$\mathcal{P}$ is (c,c,μ,μ')-robustly maximal with respect to H,*
(iii) *any vertex is in at least dn^{k-1} edges $e\in H$ with $\mathbf{i}_{\mathcal{P}}(e)\in L^{\mu}_{\mathcal{P}}(H)$, and*
(iv) *$\mathbf{i}_{\mathcal{P}}(V)\in L^{\mu}_{\mathcal{P}}(H)$.*

Then H contains a perfect matching.

We will prove Lemma 7.1 by taking a random k-partition of H and deducing it from the partite key lemma that we prove in this section. However, showing that the conditions on H transfer to the k-partite subgraph induced by the random partition is technical and non-trivial, and so we defer the proof of Lemma 7.1 to Section 9. In this section, after some preparatory results in the first subsection, we prove the partite key lemma (Lemma 7.6) in the second subsection.

7.1. Preliminaries

The following result forms the base case of the partite key lemma (Lemma 7.6). This is the case where our k-graph H is far from any divisibility barrier. Note that here any edge $e \in H$ must have $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{1}$. So condition (iv) is trivial, and condition (iii) simply states that every vertex lies in at least dn^{k-1} edges. However, we state the lemma in this form for ease of comparison to the full version of Lemma 7.6.

Lemma 7.2. *Suppose that $k \geq 3$ and $1/n \ll \varepsilon \ll \mu \ll \mu' \ll c, d \ll \gamma, 1/k$. Let $\mathcal{P}$ partition a set V into vertex classes $V_1, \dots, V_k$ of size n and let H be a $\mathcal{P}$-partite k-graph on V. Suppose that*

(i) *at most εn^{k-1} $\mathcal{P}$-partite $(k-1)$-sets $S \subseteq V$ have $d_H(S) < (1/k + \gamma)n$,*
(ii) *$\mathcal{P}$ is (c, c, μ, μ')-robustly maximal with respect to H,*
(iii) *any vertex is in at least dn^{k-1} edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) \in L^{\mu}_{\mathcal{P}}(H)$, and*
(iv) *$\mathbf{i}_{\mathcal{P}}(V) \in L^{\mu}_{\mathcal{P}}(H)$.*

Then H contains a perfect matching.

The proof requires the following lemma that will enable us to apply Theorem 4.6.

Lemma 7.3. *Suppose that $1/n \ll \varepsilon \ll \alpha \ll 1/k$. Let $\mathcal{P}$ partition a vertex set V into parts $V_1, \dots, V_k$ each of size n. Suppose H is a $\mathcal{P}$-partite k-graph on V such that at most εn^{k-1} $\mathcal{P}$-partite $(k-1)$-sets $S \subseteq V$ have $d_H(S) < D$. Then there exists a k-complex J on V such that $J_k \subseteq H$ and $\delta^*(J) \geq ((1 - \sqrt{\varepsilon})n, (1-\alpha)n, \dots, (1-\alpha)n, D - \alpha n)$.*

Proof. Let $\beta = \alpha/k^k$. Call a $\mathcal{P}$-partite $(k-1)$-set $A \subseteq V$ *bad* if $d_H(A) < D$, and *good* otherwise. We define bad $\mathcal{P}$-partite sets A of size i recursively for $i = k-2, k-3, \dots, 0$ by saying that A is *bad* if it there are more than βn vertices x in some part of $\mathcal{P}$ such that $A \cup \{x\}$ is bad. If a $\mathcal{P}$-partite set A is not bad we say it is *good*. We claim that for any $0 \leq i \leq k-1$ the number of bad i-sets is at most $\sqrt{\varepsilon}\binom{n}{i}$. To see this, we show by induction on i that the number of bad $(k-i)$-sets is at most $\varepsilon(k/\beta)^i\binom{n}{k-i}$. The base case $i = 1$ holds by assumption on H, since $\beta \ll 1/k$. The induction step follows since there are at least $\beta n/k$ times as many bad $(k-i)$-sets as bad $(k-i-1)$-sets. Thus the claim holds.

We define the k-complex J as follows. We take $J_0 = \{\emptyset\}$, which is good by the above claim, and define J_i recursively for $1 \leq i < k$ as the family of good i-sets A such that every $(i-1)$-subset of A is an element of J_{i-1}. Since the number of bad singletons (i.e. bad 1-sets) is at most $\sqrt{\varepsilon}n$, we then have $\delta_0^*(J) \geq (1-\sqrt{\varepsilon})n$. For the remaining inequalities on $\delta^*(J)$ we now prove by induction that $\delta_i^*(J) \geq (1-k^i\beta)n \geq (1-\alpha)n$ for each $0 \leq i \leq k-2$. The base case $i=0$ is already done. For the induction step, suppose that $\delta_i^*(J) \geq (1-k^i\beta)n$ for some $0 \leq i \leq k-3$. Suppose for a contradiction that there exists $e \in J_{i+1}$ such that $|J_{i+2}(e) \cap V_j| < (1-k^{i+1}\beta)n$ for some $j \in [k]$ such that e is disjoint from V_j.

Let F be the set of good sets $e^+ = e \cup \{w\}$ such that $w \in V_j$ and $e^+ \notin J_{i+2}$. Then $|F| \geq (k^{i+1}-1)\beta n$, as by choice of e there are at least $k^{i+1}\beta n$ choices of w such that $e^+ \notin J_{i+2}$, and at most βn choices make e^+ bad, since $e \in J_{i+1}$ is good. Next note that any $e^+ \in F$ contains some $(i+1)$-set $e^* \notin J_{i+1}$, otherwise we would have added e^+ to J_{i+2}. Let F' be the set of such $(i+1)$-sets e^* for all $e^+ \in F$ (choosing arbitrarily if there is more than one choice for some e^+). Then $|F'| = |F|$, as each e^+ is determined by its choice of e^* (we have $e^+ = e \cup e^*$).

Note that each $e^* \in F'$ intersects e in i vertices, so there is some i-set $e^- \subseteq e$ which is contained in at least $|F'|/(i+1) > k^i\beta n$ sets of F'. Then $|J_{i+1}(e^-) \cap V_j| < (1-k^i\beta)n$, as F' is disjoint from J_{i+1}. But since $e \in J_{i+1}$ we have $e^- \in J_i$, and so $\delta_i^*(J) < (1-k^i\beta)n$. This contradiction establishes the induction step.

Now we define J_k to be the set of all edges $e \in H$ such that $e^- \in J_{k-1}$ for every $e^- \subseteq e$ of size $k-1$. It remains to show that $\delta_{k-1}^*(J) \geq D - \alpha n$, which we do similarly to the induction step. Suppose there exists $e \in J_{k-1}$ such that $|J_k(e)| < D - \alpha n$ and let F be the set of edges $e^+ \in H \setminus J_k$ such that $e \subseteq e^+$. Since e is good, we must have $|F| \geq \alpha n$. Next note that any $e^+ \in F$ contains some $(k-1)$-set $e^* \notin J_{k-1}$, otherwise we would have added e^+ to J_k. Again let F' be the set of such $(k-1)$-sets e^* for all $e^+ \in F$. Then $|F'| = |F|$ (as before, each e^+ is determined by its choice of e^*).

Note that each $e^* \in F'$ intersects e in $k-2$ vertices, so there is some $(k-2)$-set $e^- \subseteq e$ which is contained in at least $|F'|/(k-1) > k^{k-1}\beta n$ sets of F'. Then $|J_{k-1}(e^-)| < (1-k^{k-1}\beta)n$, as F' is disjoint from J_{k-1}. But since $e \in J_{k-1}$ we have $e^- \in J_{k-2}$, and so $\delta_{k-2}^*(J) < (1-k^{k-1}\beta)n$. This contradiction completes the proof. □

Proof of Lemma 7.2. Introduce a new constant α with $\varepsilon \ll \alpha \ll \mu$. We apply Lemma 7.3 with $D = (1/k+\gamma)n$ to obtain a $\mathcal{P}$-partite k-complex J on V with $J_k \subseteq H$ and

$$\delta^*(J) \geq \big((1-\sqrt{\varepsilon})n, (1-\alpha)n, \ldots, (1-\alpha)n, (1/k+\gamma-\alpha)n\big).$$

Next we choose a matching M in H of size at most $k\sqrt{\varepsilon}n$ which includes all of the at most $k\sqrt{\varepsilon}n$ vertices $x \in V$ for which $\{x\}$ is not an edge of J; we can construct M greedily by the vertex degree assumption. We write $V' = V \setminus V(M)$, $n' = n - |M|$, and $J' = J[V']$ and verify that J' satisfies the conditions of Theorem 4.6 with n', 2μ and $\gamma/2$ in place of n, μ and γ respectively.

The degree sequence condition holds as

$$\delta^*(J') \geq (n', (1-2\alpha)n', \dots, (1-2\alpha)n', (1/k+\gamma-2\alpha)n').$$

(Note that only the first and last co-ordinates are close to being tight.) Furthermore, $\mathcal{P}[V']$ is $(c/2, 2c, 2\mu, \mu'/2)$-robustly maximal with respect to J'_k by Lemma 5.3. Now suppose that $\mathcal{P}'$ is a partition of V' into parts of size at least $\delta^*_{k-1}(J') - 2\mu n' \geq n'/k \geq 2c|V'|$ which refines $\mathcal{P}[V']$. Since J'_k is $\mathcal{P}$-partite every edge $e \in J'_k$ has $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{1}$; in particular we have $\mathbf{1} \in L^{\mu'/2}_{\mathcal{P}[V']}(J'_k)$. So by Proposition 5.4 we have $\mathbf{i} \in L^{2\mu}_{\mathcal{P}'}(J'_k)$ for any index vector $\mathbf{i}$ with respect to $\mathcal{P}'$ such that $(\mathbf{i} \mid \mathcal{P}) = \mathbf{1}$, i.e. $L^{2\mu}_{\mathcal{P}'}(J'_k)$ is complete with respect to $\mathcal{P}[V']$.

Now J'_k has a perfect matching by Theorem 4.6. Together with M, we have a perfect matching in H. □

Next we recall from the proof outline the form of our inductive approach, which uses the following 'canonical' induced subgraphs.

Definition 7.4 *(Canonical subgraphs).* Suppose that H is a k-graph on V, $\mathcal{P}$ is a partition of V, and $\mu > 0$. For each $\mathbf{i} \in I^{\mu}_{\mathcal{P}}(H)$, we define $H_{\mathbf{i}}$ to be the induced subgraph of H on the union of the parts $W \in \mathcal{P}$ such that $i_W > 0$, and $\mathcal{P}_{\mathbf{i}}$ to be the restriction of $\mathcal{P}$ to $V(H_{\mathbf{i}})$.

Note that $H_{\mathbf{i}}$ contains all the edges of H of index $\mathbf{i}$ and is $\mathcal{P}_{\mathbf{i}}$-partite whenever H is $\mathcal{P}$-partite. In the following proposition we establish various properties of H and these subgraphs $H_{\mathbf{i}}$. In (i), we show that the robust edge-lattice of H is full, so the properties of Proposition 6.12 also hold. In (ii), we show that $H_{\mathbf{i}}$ inherits a similar codegree condition to that of H (with slightly weaker constants), which is essential for the induction. Note that a minimum codegree condition would not be inherited, so this is one reason why for much of this paper we work with a condition on the codegree of most $(k-1)$-sets (we also need this condition for Lemma 2.4). The next two parts are useful boosting properties for the edge detection parameter and the part sizes. We will apply these with $d = 1/k$ and $c \ll 1/k$, so (iii) shows that the part sizes of $\mathcal{P}$ are in fact much larger than our original assumption, and (iv) shows that the robust edge-lattice is unchanged for a wide range of μ. Finally, recall from our proof outline that we choose matching edges to cover bad vertices; (v) will show that there are few bad vertices.

Proposition 7.5. *Suppose that $1/n \ll \mu, \varepsilon \ll \psi, d, c, 1/k$. Let $\mathcal{P}'$ partition a set V into k parts $V_1, \dots, V_k$, where $cn \leq |V_i| \leq n$ for each $i \in [k]$. Let H be a $\mathcal{P}'$-partite k-graph on V and $\mathcal{P}$ be a partition of V which refines $\mathcal{P}'$. Suppose that*

(α1) *at most εn^{k-1} $\mathcal{P}'$-partite $(k-1)$-sets $S \subseteq V$ have $d_H(S) < dn$, and*
(α2) *$\mathcal{P}$ has parts each of size at least cn and $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free.*

Then we have the following properties.

(i) $I^\mu_{\mathcal{P}}(H)$ *is full with respect to* $\mathcal{P}'$.
(ii) *For any* $\mathbf{i} \in I^\mu_{\mathcal{P}}(H)$ *at most* ψn^{k-1} $\mathcal{P}_{\mathbf{i}}$*-partite* $(k-1)$*-sets* $S \subseteq V(H_{\mathbf{i}})$ *have* $|H_{\mathbf{i}}(S)| < (d-\psi)n$.
(iii) *Each part of* $\mathcal{P}$ *has size at least* $(d-\psi)n$.
(iv) $I^{dc^{k-1}/2k^k}_{\mathcal{P}}(H) = I^\mu_{\mathcal{P}}(H)$.
(v) *For any* $\mathbf{i} \in I^\mu_{\mathcal{P}}(H)$ *at most* ψn *vertices of* $V(H_{\mathbf{i}})$ *lie in fewer than* $c^{k-2}dn^{k-1}/2$ *edges of* $H_{\mathbf{i}}$.

Proof. Let θ satisfy $\mu, \varepsilon \ll \theta \ll \psi, d, c, 1/k$. For (i), let $\mathbf{v}$ be a non-negative index vector with respect to $\mathcal{P}$ such that $(\mathbf{v} \mid \mathcal{P}') = \mathbf{1} - \mathbf{u}_{V_k}$. Then it suffices to show that there is some part X_k of $\mathcal{P}$ with $X_k \subseteq V_k$ such that H contains at least $\mu|V|^k$ edges of index $\mathbf{v} + \mathbf{u}_{X_k}$. To see this, write $\mathbf{v} = \sum_{j\in[k-1]} \mathbf{u}_{X_j}$ for parts $X_j \subseteq V_j$ of $\mathcal{P}$ and note that the number of $(k-1)$-sets $S = \{x_1, \dots, x_{k-1}\}$ with $x_j \in X_j$ for each $j \in [k-1]$ is at least $\prod_{j\in[k-1]} |X_j| \geq (cn)^{k-1}$. Hence at least $(c^{k-1} - \varepsilon)n^{k-1}$ such $(k-1)$-sets have $d_H(S) \geq dn$, and so at least $(c^{k-1} - \varepsilon)dn^k$ edges of H contain one vertex from each of $X_1, \dots, X_{k-1}$. Since $\mathcal{P}$ refines V_k into at most $1/c$ parts, some part $X_k \subseteq V_k$ of $\mathcal{P}$ must be as required.

For (ii), fix $\mathbf{i} \in I^\mu_{\mathcal{P}}(H)$, and for $j \in [k]$ let $W_j \subseteq V_j$ be the part of $\mathcal{P}$ for which $i_{W_j} = 1$ (so $V(H_{\mathbf{i}}) = \bigcup_{j\in[k]} W_j$). Let $\mathcal{S}$ be the family of sets $S = \{x_1, \dots, x_{k-1}\}$ with $x_i \in W_i$ for each $i \in [k-1]$. Also define $\mathcal{S}_1$ to consist of those $S \in \mathcal{S}$ with $d_H(S) \geq dn$ but $d_{H_{\mathbf{i}}}(S) < (d-\theta)n$, and $\mathcal{S}_2$ to consist of those $S \in \mathcal{S}$ with $d_H(S) < dn$. Then as above we have $|\mathcal{S}| \geq (cn)^{k-1}$, whilst $|\mathcal{S}_2| < \varepsilon n^{k-1}$ by $(\alpha 1)$. To bound $|\mathcal{S}_1|$, note that each $S \in \mathcal{S}_1$ is contained in at least θn edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e)$ equal to some $\mathbf{i}' \neq \mathbf{i}$, and since $\mathbf{i}' - \mathbf{i}$ is a transferral for any such $\mathbf{i}'$ we have $\mathbf{i}' \notin I^\mu_{\mathcal{P}}(H)$. Now there are at most $c^{-k}\mu(kn)^k$ such edges of H, as each part V_j of $\mathcal{P}'$ is refined into at most $1/c$ parts of $\mathcal{P}$, so there are at most c^{-k} possible values of $\mathbf{i}_{\mathcal{P}}(e)$ for an edge e of H. We conclude that $|\mathcal{S}_1| \leq c^{-k}\mu(kn)^k/\theta n \leq \theta n^{k-1}/(k+1)$, so $\mathcal{T} = \mathcal{S}_1 \cup \mathcal{S}_2$ has size $|\mathcal{T}| \leq \theta n^{k-1}/k < |\mathcal{S}|$. Since any $S \in \mathcal{S} \setminus \mathcal{T}$ has $d_{H_{\mathbf{i}}}(S) \geq (d-\theta)n$, by symmetry this proves a stronger form of (ii), with θ in place of ψ.

For (iii), consider any $Y \in \mathcal{P}$ and fix $\mathbf{i} \in I^\mu_{\mathcal{P}}(H)$ such that $i_Y = 1$; this exists by (i) and Proposition 6.12(ii). Without loss of generality $Y \subseteq V_k$. Then with notation as in (ii), for any $S \in \mathcal{S} \setminus \mathcal{T}$ we have $|Y| \geq d_{H_{\mathbf{i}}}(S) \geq (d-\theta)n$. The same argument implies that $|H_{\mathbf{i}}| \geq |\mathcal{S} \setminus \mathcal{T}|(d-\theta)n \geq c^{k-1}d|V(H)|^k/2k^k$, and so we have (iv).

Finally, note that the number of vertices in W_1 that belong to at least $(cn)^{k-2}/3$ sets $S \in \mathcal{T}$ is at most $|\mathcal{T}| \cdot 3/(cn)^{k-2} \leq \psi n/k$. Since $\mathcal{P}$ has parts each of size at least cn, any other vertex of W_1 lies in at least $2(cn)^{k-2}/3 \cdot (d-\theta)n \geq dc^{k-2}n^{k-1}/2$ edges of $H_{\mathbf{i}}$. The same argument applies to any part W_j with $j \in [k]$, so this proves (v). □

7.2. The partite key lemma

In the remainder of this section we prove the following partite form of the key lemma.

Lemma 7.6. *Let k and ℓ be integers with $k \geq \ell \geq 2$ and $k \geq 3$, and suppose that $1/n \ll \varepsilon \ll \mu \ll \mu' \ll c, d \ll \gamma, 1/k$. Let $\mathcal{P}'$ partition a set V into vertex classes $V_1, \dots, V_k$ of size n, let H be a $\mathcal{P}'$-partite k-graph on V and let $\mathcal{P}$ be a refinement of $\mathcal{P}'$. Suppose that*

($\beta 1$) *at most εn^{k-1} $\mathcal{P}'$-partite $(k-1)$-sets $S \subseteq V$ have $d(S) < (1/\ell + \gamma)n$,*
($\beta 2$) *$\mathcal{P}$ is (c, c, μ, μ')-robustly maximal with respect to H,*
($\beta 3$) *any vertex is in at least dn^{k-1} edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) \in L^{\mu}_{\mathcal{P}}(H)$, and*
($\beta 4$) *$\mathbf{i}_{\mathcal{P}}(V) \in L^{\mu}_{\mathcal{P}}(H)$.*

Then H contains a perfect matching.

We remark that the case $\ell = k$ implies the cases $\ell = 2, \dots, k-1$, and hence we set $\ell = k$ in all later applications of the lemma. The other cases are only required for the proof of the lemma itself.

Proof. We fix k and proceed by induction on ℓ. Let $s = \lfloor 1/c \rfloor$; then we use the following hierarchy of constants:

$$1/n \ll 1/D \ll \varepsilon \ll \mu \ll \mu'_0 \ll \mu'_1 \ll \cdots \ll \mu'_{s^k} \ll \mu' \ll c, d \ll \eta, 1/K \ll \gamma, 1/k.$$

In several places during the proof we will delete small matchings from H; for convenient notation we let $\mathcal{P}'$ and $\mathcal{P}$ also denote the restrictions of $\mathcal{P}'$ and $\mathcal{P}$ to the undeleted vertices.

Step 1: The case $\mathcal{P} = \mathcal{P}'$. In this case we have a perfect matching by Lemma 7.2. In particular, this gives the base case $\ell = 2$ of the induction, as by Proposition 7.5(iii) each part of $\mathcal{P}$ has size at least $n/\ell + \gamma n/2 > n/2$, so we must have $\mathcal{P} = \mathcal{P}'$.

Step 2: The canonical subgraphs. For the rest of the proof we assume that $\mathcal{P} \neq \mathcal{P}'$, that $3 \leq \ell \leq k$ and that the lemma holds with $\ell - 1$ in place of ℓ. We write

$$I = I^{\mu}_{\mathcal{P}}(H) \quad \text{and} \quad L = L^{\mu}_{\mathcal{P}}(H).$$

Our strategy will be to split H up randomly into a number of vertex-disjoint k-partite subgraphs, each of which satisfies the conditions of the lemma (with weaker constants) when ℓ is replaced by $\ell - 1$. The inductive hypothesis will then imply the existence of a perfect matching in each subhypergraph, and taking the union of these matchings gives

a perfect matching in H. In this step we lay the groundwork by analysing the canonical subgraphs from which the random subgraphs will be chosen; the analogues of properties $(\beta 1)$, $(\beta 2)$ and $(\beta 3)$ for these subgraphs will follow from this analysis.

We first observe that H, $\mathcal{P}'$ and $\mathcal{P}$ meet the conditions of Proposition 7.5 with $1/\ell+\gamma$ in place of d. Therefore I is full with respect to $\mathcal{P}'$. We can therefore apply Proposition 6.12 to obtain an integer r such that $|I| = r^{k-1}$, each part of $\mathcal{P}'$ is partitioned into exactly r parts by $\mathcal{P}$, and for any part X of $\mathcal{P}$ there are exactly r^{k-2} vectors $\mathbf{i} \in I$ such that $i_X = 1$. Also, by Proposition 7.5(iii), every part of $\mathcal{P}$ has size at least $(1/\ell+\gamma/2)n$. Since $\mathcal{P} \neq \mathcal{P}'$ this implies $2 \leq r < \ell$.

Next we recall Definition 7.4: for each $\mathbf{i} \in I$, the canonical subgraph $H_{\mathbf{i}}$ is the induced k-graph on the vertex set $\bigcup_{X\in\mathcal{P}:i_X=1} X$ whose edge set consists of all edges of H with index $\mathbf{i}$, and $\mathcal{P}_{\mathbf{i}}$ is the restriction of $\mathcal{P}$ to $V(H_{\mathbf{i}})$ (so $H_{\mathbf{i}}$ is $\mathcal{P}_{\mathbf{i}}$-partite). The canonical subgraphs $H_{\mathbf{i}}$ will act as prototypes for the k-graphs on which we use the inductive hypothesis, which will be induced subgraphs of the canonical subgraphs on randomly chosen sets of vertices. Note that by Proposition 7.5(ii), at most $\mu'_0 n^{k-1}$ $\mathcal{P}_{\mathbf{i}}$-partite $(k-1)$-sets $S \subseteq V(H_{\mathbf{i}})$ have $|H_{\mathbf{i}}(S)| < n/\ell + 3\gamma n/4$: this inheritance of the codegree condition of H is fundamental to our inductive approach. It also provides the analogue of condition $(\alpha 1)$ when we apply Proposition 7.5 again to each $H_{\mathbf{i}}$, which we will do to 'boost' the part sizes and edge-detection parameter.

Since $\mathcal{P}_{\mathbf{i}}$ may not be robustly maximal with respect to $H_{\mathbf{i}}$, we now need to identify robustly maximal partitions for each canonical subgraph. Let $S = s^{k^k}$. We claim that there exist $i \in [S]$, and partitions $\mathcal{Q}_{\mathbf{i}}$ of $V(H_{\mathbf{i}})$ for each $\mathbf{i} \in I$ such that $\mathcal{Q}_{\mathbf{i}}$ refines $\mathcal{P}_{\mathbf{i}}$ and is $(c, c, \mu'_{i-1}, \mu'_i)$-robustly maximal with respect to $H_{\mathbf{i}}$. To see this, we repeatedly apply Proposition 5.2 to each $\mathbf{i} \in I$. First, we choose some $\mathbf{i} \in I$ and apply Proposition 5.2 with parameters $\mu'_{\lambda S/s}$ for $0 \leq \lambda \leq s$ to obtain some $b \in [s]$ and a partition $\mathcal{Q}_{\mathbf{i}}$ of $V(H_{\mathbf{i}})$ such that $\mathcal{Q}_{\mathbf{i}}$ refines $\mathcal{P}_{\mathbf{i}}$ and is $(c, c, \mu'_{(b-1)S/s}, \mu'_{bS/s})$-robustly maximal with respect to $H_{\mathbf{i}}$. We then repeat the following step. After j steps, we will have chosen j of these partitions to be $(c, c, \mu'_{(b-1)S/s^j}, \mu'_{bS/s^j})$-robustly maximal for some $b \in [s^j]$. We then choose some $\mathbf{i}$ we have not yet considered and apply Proposition 5.2 with parameters $\mu'_{\lambda S/s^{j+1}}$ for $(b-1)s \leq \lambda \leq bs$ to obtain some $b' \in [s^{j+1}]$ and a partition $\mathcal{Q}_{\mathbf{i}}$ of $V(H_{\mathbf{i}})$ such that $\mathcal{Q}_{\mathbf{i}}$ refines $\mathcal{P}_{\mathbf{i}}$ and is $(c, c, \mu'_{(b'-1)S/s^{j+1}}, \mu'_{b'S/s^{j+1}})$-robustly maximal with respect to $H_{\mathbf{i}}$. Since $|I| \leq k^k$ we can repeat this process for every $\mathbf{i}$, which proves that the claim holds.

For simplicity of notation we now relabel μ'_{i-1} as μ_0 and μ'_i as μ_6, and introduce new constants $\mu_1, \dots, \mu_5$ and μ'' such that

$$\mu \ll \mu_0 \ll \mu_1 \ll \cdots \ll \mu_5 \ll \mu_6 \ll \mu'' \ll \mu'.$$

For each $\mathbf{i} \in I$, we write

$$I_{\mathbf{i}} := I^{\mu_0}_{\mathcal{Q}_{\mathbf{i}}}(H_{\mathbf{i}}) \quad \text{and} \quad L_{\mathbf{i}} := L^{\mu_0}_{\mathcal{Q}_{\mathbf{i}}}(H_{\mathbf{i}}).$$

We will apply Proposition 7.5 to $H_{\mathbf{i}}$ with $\mathcal{P}_{\mathbf{i}}$ in place of $\mathcal{P}'$, $\mathcal{Q}_{\mathbf{i}}$ in place of $\mathcal{P}$, μ_0 in place of both μ and ε, and $1/\ell+3\gamma/4$ and $\gamma/12$ in place of d and ψ respectively (n, c and k are

unchanged). Indeed, condition ($\alpha 1$) of Proposition 7.5 holds by our earlier observation that at most $\mu_0' n^{k-1}$ $\mathcal{P}_{\mathbf{i}}$-partite $(k-1)$-sets $S \subseteq V(H_{\mathbf{i}})$ have $|H_{\mathbf{i}}(S)| < n/\ell + 3\gamma n/4$, whilst condition ($\alpha 2$) holds since $\mathcal{Q}_{\mathbf{i}}$ is (c, c, μ_0, μ_6)-robustly maximal with respect to $H_{\mathbf{i}}$. So by Proposition 7.5(i) and (iii) we deduce that $I_{\mathbf{i}}$ is full with respect to $\mathcal{P}_{\mathbf{i}}$, and every part of $\mathcal{Q}_{\mathbf{i}}$ has size at least $n/\ell + 2\gamma n/3$. The latter conclusion implies that $\mathcal{Q}_{\mathbf{i}}$ is in fact $(1/\ell + 2\gamma/3, c, \mu_0, \mu_6)$-robustly maximal, and also that we can apply Proposition 7.5 again with $1/\ell$ in place of c (and all other variables as before); Proposition 7.5(iv) then implies that

$$I_{\mathcal{Q}_{\mathbf{i}}}^{1/2(k\ell)^k}(H_{\mathbf{i}}) = I_{\mathbf{i}}.$$

Indeed, we have the boosting property that $I_{\mathcal{Q}_{\mathbf{i}}}^{\mu_0}(H_{\mathbf{i}})$ and $L_{\mathcal{Q}_{\mathbf{i}}}^{\mu_0}(H_{\mathbf{i}})$ are essentially independent of the edge-detection parameter, in that they remain unchanged when μ_0 is replaced by any constant between μ_0 and $1/2(k\ell)^k$.

Finally, say that a vertex of $H_{\mathbf{i}}$ is *bad for* $\mathbf{i}$ if it lies in fewer than dn^{k-1} edges $e \in H_{\mathbf{i}}$ with $\mathbf{i}_{\mathcal{Q}_{\mathbf{i}}}(e) \in I_{\mathbf{i}}$. We say that a vertex is *bad* if it is bad for some $\mathbf{i} \in I$. Since $|I| = r^{k-1}$, by Proposition 7.5(v) (with μ_1/r^{k-1} in place of ψ), the number of bad vertices is at most $\mu_1 n$. Using assumption ($\beta 3$), i.e. that every vertex is contained in at least dn^{k-1} edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) \in L$, we may greedily choose a matching M of such edges with size at most $\mu_1 n$ which covers all of the bad vertices. We now restrict attention to $V' := V \setminus V(M)$.

Step 3: The meet and join. The principal difficulty we must overcome in order to apply the inductive hypothesis is to ensure that an analogue of ($\beta 4$) holds, that is, we require subgraphs in which the index vector of the entire vertex set lies in the appropriate robust edge-lattice. In this step we introduce two additional partitions of V that will be used to achieve this.

We begin by noting that for the subgraph chosen in $H_{\mathbf{i}}$, the index vector of the vertex set with respect to $\mathcal{Q}_{\mathbf{i}}$ simply lists the number of vertices chosen from each part of $\mathcal{Q}_{\mathbf{i}}$. Hence if two vertices x and y are contained in the same part of $\mathcal{Q}_{\mathbf{i}}$ for *every* $\mathcal{Q}_{\mathbf{i}}$ such that $x, y \in V(H_{\mathbf{i}})$, then they are interchangeable for the current purpose. With this in mind, we first let $\mathcal{Q}^{\cap}$ be the 'meet' of the partitions $\mathcal{Q}_{\mathbf{i}}$ for $\mathbf{i} \in I$: we say that two vertices in the same part W of $\mathcal{P}$ are in the same part of $\mathcal{Q}^{\cap}$ if and only if they lie in the same part of $\mathcal{Q}_{\mathbf{i}}$ for every $\mathbf{i} \in I$ with $i_W = 1$. This is not our final definition, as we will also require that every part of $\mathcal{Q}^{\cap}$ has size at least D.

To achieve this, note that $\mathcal{P}$ has at most k^2 parts, and each part X of $\mathcal{P}$ is partitioned into at most k parts by each $\mathcal{Q}_{\mathbf{i}}$ such that $i_X = 1$, so the number of parts of $\mathcal{Q}^{\cap}$ is at most $k^2 \cdot k^{|I|} = k^{2+r^{k-1}} < K$. Thus there is some $z \in [K+1]$ such that no part of $\mathcal{Q}^{\cap}$ has size between $(z-1)(kK)^{z-1}D$ and $z(kK)^z D$. Let B be the union of all parts of size at most $z(kK)^z D$. Then $|B| \leq K \cdot (z-1)(kK)^{z-1}D$. Similarly to the end of Step 2, we greedily choose a matching M' in H of size at most $|B|$ that covers every vertex of B so that every edge $e \in M'$ has $\mathbf{i}_{\mathcal{P}}(e) \in L$. We will now restrict attention to $V' \setminus V(M')$.

Next we introduce the 'join' $\mathcal{Q}^\cup$ of the partitions $\mathcal{Q}_\mathbf{i}$, which will allow us to avoid the problem of 'hidden' approximate divisibility barriers (recall Construction 3.1). Let G be the graph whose vertices are the parts of $\mathcal{Q}^\cap$ not contained in B, where $Z, Z' \in \mathcal{Q}^\cap$ are adjacent in G if they are contained in the same part of some $\mathcal{Q}_\mathbf{i}$. Then the parts of $\mathcal{Q}^\cup$ are formed by taking the union of the parts of $\mathcal{Q}^\cap$ in each component of G.

As an aid to memory, in the remainder of the proof we tend to use the consistent notation W, X, Y, Z for general parts of $\mathcal{P}$, $\mathcal{Q}^\cup$, $\mathcal{Q}_\mathbf{i}$ (for $\mathbf{i} \in I$), $\mathcal{Q}^\cap$ respectively, i.e. earlier letters in the alphabet denote (potentially) larger parts.

Step 4: Restricting to good vertices. We now delete all of the at most $2k\mu_1 n$ vertices covered by $M \cup M'$. To avoid introducing more complicated notation, all of our notation is to be now understood as referring to the undeleted vertices, e.g. V now refers to $V \setminus (M \cup M')$. However, all variables (including n) remain unchanged in value. Note that any part Z of $\mathcal{Q}^\cap$ had size at least $z(kK)^z D$ before deleting the vertices covered by M' (otherwise we would have deleted it). Since $|M'| \le |B| \le K \cdot (z-1)(kK)^{z-1} D$, we now have

$$|Z| \ge z(kK)^z D - k\big|M'\big| \ge D. \tag{1}$$

Since we have deleted a number of vertices from H, we now show that the properties of H from the statement of the lemma and from step 2 are preserved (albeit with weaker constants) after these vertex deletions. Note that (A1)–(A3) in the following claim are the respective analogues of (β1)–(β3) for $H_\mathbf{i}$ with respect to $\mathcal{P}_\mathbf{i}$ and $\mathcal{Q}_\mathbf{i}$.

Claim 7.7. *The following hold for every* $\mathbf{i} \in I$.

(A1) *At most* $\mu_0 n^{k-1}$ $\mathcal{P}_\mathbf{i}$*-partite* $(k-1)$*-sets* $S \subseteq V(H_\mathbf{i})$ *have* $|H_\mathbf{i}(S)| < (1/\ell + \gamma/2)n$.
(A2) $\mathcal{Q}_\mathbf{i}$ *is* $(1/\ell, 2c, \mu_2, \mu_5)$*-robustly maximal with respect to* $H_\mathbf{i}$ *for each* $\mathbf{i} \in I$.
(A3) *Every vertex of* $H_\mathbf{i}$ *lies in at least* $\frac{1}{2}dn^{k-1}$ *edges* $e \in H_\mathbf{i}$ *with* $\mathbf{i}_{\mathcal{Q}_\mathbf{i}}(e) \in I_\mathbf{i}$.
(A4) $\mathbf{i}_\mathcal{P}(V) \in L$.
(A5) *Every part of* $\mathcal{P}$, $\mathcal{Q}^\cup$ *and each* $\mathcal{Q}_\mathbf{i}$ *has size at least* $(1/\ell + \gamma/2)n$.
(A6) $I^{\mu'}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i}) = I^{\mu_1}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i}) = I_\mathbf{i}$.
(A7) $\mathcal{P}$ *is* $(1/\ell, 2c, \mu_2, \mu'')$*-robustly maximal with respect to* H, *and* $I = I^{\mu'}_\mathcal{P}(H)$.

For (A1), recall from step 2 that it was true with $3\gamma/4$ in place of $\gamma/2$. Lemma 5.3 implies (A2), since $\mathcal{Q}_\mathbf{i}$ was $(1/\ell + 2\gamma/3, c, \mu_0, \mu_6)$-robustly maximal with respect to $H_\mathbf{i}$ before any vertices were deleted. The first part of (A7) follows by the same argument. For (A3), recall that M covered all vertices which were in fewer than dn^{k-1} edges of $H_\mathbf{i}$ with $\mathbf{i} \in I_\mathbf{i}$ and at most $2k\mu_1 n \le dn/2$ vertices were deleted. For (A4), note that it was true before any vertices were deleted by (β4), and $\mathbf{i}_\mathcal{P}(V(M \cup M')) \in L$ since $\mathbf{i}_\mathcal{P}(e) \in L$ for every $e \in M \cup M'$.

For (A5), recall that each part of any $\mathcal{Q}_\mathbf{i}$ had size at least $(1/\ell + 2\gamma/3)n$ before the deletions, and at most $2k\mu_1 n \le \gamma n/6$ vertices were deleted. The bounds for $\mathcal{Q}^\cup$ and $\mathcal{P}$

follow, since every part of $\mathcal{Q}^\cup$ contains a part of $\mathcal{Q}_\mathbf{i}$ for some $\mathbf{i} \in I$, and $\mathcal{Q}^\cup$ refines $\mathcal{P}$. For (A6), we observed before any deletions that $I^{1/2(k\ell)^k}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i}) = I^{\mu_0}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i}) = I_\mathbf{i}$, and $|V(H_\mathbf{i})|$ had size at least n both before and after the deletions. So if $\mathbf{v} \in I_\mathbf{i}$ then there were at least $n^k/2(k\ell)^k$ edges of $H_\mathbf{i}$ of index $\mathbf{v}$, at most $2k\mu_1 n^k$ of which were deleted, so $\mathbf{v} \in I^{\mu'}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i})$. On the other hand, if $\mathbf{v} \notin I_\mathbf{i}$ then there were at most $\mu_0(kn)^k < \mu_1 n^k$ edges of $H_\mathbf{i}$ of index $\mathbf{v}$, so $\mathbf{v} \notin I^{\mu_1}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i})$. Since $I^{\mu'}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i}) \subseteq I^{\mu_1}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i})$, this proves (A6). The same argument proves the second part of (A7), completing the proof of the claim.

Step 5: Choosing sizes. In this step we determine how many vertices each of our final random subgraphs should choose from each part of $\mathcal{Q}^\cap$ to ensure that the analogue of ($\beta 4$) holds. We accomplish this in three stages. Firstly, we choose rough targets for the number of vertices to be contained in each subgraph. Secondly, we determine how many vertices are to be chosen from each part of $\mathcal{Q}^\cup$. Thirdly, we determine how many vertices are to be chosen from each part of $\mathcal{Q}^\cap$.

Claim 7.8. *There exist integers $\rho_\mathbf{i}$ for each $\mathbf{i} \in I$ which satisfy*

(B1) $\sum_{\mathbf{i}\in I} \rho_\mathbf{i}\mathbf{i} = \mathbf{i}_\mathcal{P}(V)$,
(B2) $\rho_\mathbf{i} \geq \eta n$ *for each* $\mathbf{i} \in I$.

To prove Claim 7.8, for each $X \in \mathcal{P}$ we start by reserving $N := \lceil 2\eta n \rceil$ vertices of X for each $\mathbf{i} \in I$ with $i_X = 1$. Recall that for any part X of $\mathcal{P}$ there are exactly r^{k-2} vectors $\mathbf{i} \in I$ such that $i_X = 1$, so exactly $r^{k-2}N$ vertices are reserved from each part of $\mathcal{P}$. Let V' be the set of unreserved vertices and write $V'_j = V' \cap V_j$; recalling that each V_j now has size $n - |M \cup M'|$ after the deletions, each V'_j has size $n' := n - |M \cup M'| - r^{k-2}N$. Also, by (A5) at least n'/k vertices remain unreserved in each part of $\mathcal{P}$. Since I is full with respect to $\mathcal{P}'$, by Proposition 6.13 we deduce that $\mathbf{i}_\mathcal{P}(V') \in PC(I)$. That is, we may fix reals $\lambda_\mathbf{i} \geq 0$ for each $\mathbf{i} \in I$ so that $\sum_{\mathbf{i}\in I} \lambda_\mathbf{i}\mathbf{i} = \mathbf{i}_\mathcal{P}(V')$.

Let $a_\mathbf{i} = \lfloor \lambda_\mathbf{i} \rfloor$ for each $\mathbf{i} \in I$. Then $\mathbf{v} = \mathbf{i}_\mathcal{P}(V') - \sum_{\mathbf{i}\in I} a_\mathbf{i}\mathbf{i}$ is a sum of $|I| \leq r^{k-1}$ vectors of length at most k, so $\mathbf{v} \in B(\mathbf{0}, kr^{k-1})$. We also have $\mathbf{v} \in L$, since $\mathbf{i}_\mathcal{P}(V')$ was obtained by subtracting some vectors in I from $\mathbf{i}_\mathcal{P}(V)$, and $\mathbf{i}_\mathcal{P}(V) \in L$ by (A4). So by Proposition 4.8 we may choose integers $b_\mathbf{i}$ with $|b_\mathbf{i}| \leq K$ for each $\mathbf{i} \in I$ so that $\sum_{\mathbf{i}\in I} b_\mathbf{i}\mathbf{i} = \mathbf{v}$.

Let $\rho_\mathbf{i} = a_\mathbf{i} + b_\mathbf{i} + N$ for each $\mathbf{i} \in I$. Then (B2) holds as $a_\mathbf{i} \geq 0$ and $|b_\mathbf{i}| \leq K$ for each $\mathbf{i} \in I$, and (B1) holds as

$$\sum_{\mathbf{i}\in I} \rho_\mathbf{i}\mathbf{i} = \sum_{\mathbf{i}\in I} a_\mathbf{i}\mathbf{i} + \sum_{\mathbf{i}\in I} b_\mathbf{i}\mathbf{i} + \sum_{\mathbf{i}\in I} N\mathbf{i} = \mathbf{i}_\mathcal{P}(V') + \sum_{\mathbf{i}\in I} N\mathbf{i} = \mathbf{i}_\mathcal{P}(V).$$

This completes the proof of Claim 7.8.

The vector $\mathbf{n}^\cup_\mathbf{i}$ with respect to $\mathcal{Q}^\cup$ obtained in the next claim determines how many vertices the random subgraph for $\mathbf{i}$ will take from each part of $\mathcal{Q}^\cup$. Note that (C1)

ensures that the sizes are correct to be a partition of V, (C2) that there is no divisibility obstruction to our later choice of $\mathcal{Q}^\cap$, and (C3) that the sizes are in roughly equal proportion from each part of $\mathcal{Q}^\cup$ which is a subset of a part of $\mathcal{P}_\mathbf{i}$, and zero for any other part of $\mathcal{Q}^\cup$.

The proof of the claim proceeds through three stages. In the first stage, we use the constants $\rho_\mathbf{i}$ and the part sizes of each $\mathcal{Q}_\mathbf{i}$ to determine provisional values for each $\mathbf{n}_\mathbf{i}^\cup$ which satisfy (C1) and (C3). In the second stage we 'snap' each $\mathbf{n}_\mathbf{i}^\cup$ onto a nearby lattice point so that (C2) is satisfied, preserving (C3) (although (C1) may no longer hold). In the final stage we make further adjustments to restore (C1) while preserving (C2) and (C3).

Claim 7.9. *There exist vectors $\mathbf{n}_\mathbf{i}^\cup$ with respect to $\mathcal{Q}^\cup$ for $\mathbf{i} \in I$ which satisfy*

(C1) $\sum_{\mathbf{i}\in I} \mathbf{n}_\mathbf{i}^\cup = \mathbf{i}_{\mathcal{Q}^\cup}(V)$,
(C2) $\mathbf{n}_\mathbf{i}^\cup \in L^{\mu_6}_{\mathcal{Q}^\cup}(H_\mathbf{i})$ *for every* $\mathbf{i} \in I$, *and*
(C3) *for any* $\mathbf{i} \in I$, *any part* W *of* $\mathcal{P}$ *and any part* $X \subseteq W$ *of* $\mathcal{Q}^\cup$ *we have*

$$\left(n_\mathbf{i}^\cup\right)_X = \begin{cases} \rho_\mathbf{i}\frac{|X|}{|W|} \pm K & \text{if } i_W = 1, \\ 0 & \text{otherwise.} \end{cases}$$

To prove the claim, for each $\mathbf{i} \in I$ we choose a vector $\mathbf{n}_\mathbf{i}$ with respect to $\mathcal{Q}_\mathbf{i}$ by taking $(n_\mathbf{i})_Y$ to be either $\lfloor \rho_\mathbf{i}|Y|/|W| \rfloor$ or $\lceil \rho_\mathbf{i}|Y|/|W| \rceil$ for each $W \in \mathcal{P}_\mathbf{i}$ and $Y \in \mathcal{Q}_\mathbf{i}$ with $Y \subseteq W$. We make these choices so that $\sum_{Y\subseteq W}(n_\mathbf{i})_Y = \rho_\mathbf{i}$ for each $W \in \mathcal{P}$ with $i_W = 1$; this is possible since

$$\sum_{Y\in\mathcal{Q}_\mathbf{i}:Y\subseteq W} \frac{\rho_\mathbf{i}|Y|}{|W|} = \frac{\rho_\mathbf{i}|W|}{|W|} = \rho_\mathbf{i}.$$

Now, for any $\mathbf{i} \in I$, observe that this requirement implies that $\mathbf{n}_\mathbf{i} \in L^{\mathcal{Q}_\mathbf{i}\mathcal{P}_\mathbf{i}}_{\max}$. So by Proposition 6.12(iii) (with $\mathcal{Q}_\mathbf{i}$ and $\mathcal{P}_\mathbf{i}$ in place of $\mathcal{P}$ and $\mathcal{P}'$ respectively) we may choose parts Y, Y' of $\mathcal{Q}_\mathbf{i}$ so that $\mathbf{n}_\mathbf{i}^1 := \mathbf{n}_\mathbf{i} - \mathbf{u}_Y + \mathbf{u}_{Y'} \in L_\mathbf{i} = L^{\mu'}_{\mathcal{Q}_\mathbf{i}}(H_\mathbf{i})$, where the final equality follows from (A6). Let $\mathcal{Q}_\mathbf{i}^0$ be the partition of V whose parts are those of $\mathcal{Q}_\mathbf{i}$ and those of $\mathcal{Q}^\cup$ that are disjoint from $V(H_\mathbf{i})$. Similarly, let $\mathbf{n}_\mathbf{i}^0$ be the vector with respect to $\mathcal{Q}_\mathbf{i}^0$ corresponding to $\mathbf{n}_\mathbf{i}^1$, that is, with additional zero co-ordinates corresponding to the parts of $\mathcal{Q}^\cup$ that are disjoint from $V(H_\mathbf{i})$. Then $\mathbf{n}_\mathbf{i}^0 \in L^{\mu_6}_{\mathcal{Q}_\mathbf{i}^0}(H_\mathbf{i})$. Finally, let $\mathbf{n}_\mathbf{i}' := (\mathbf{n}_\mathbf{i}^0 \mid \mathcal{Q}^\cup)$, so $\mathbf{n}_\mathbf{i}' \in L^{\mu_6}_{\mathcal{Q}^\cup}(H_\mathbf{i})$ by Proposition 4.5(i).

For any part W of $\mathcal{P}$ and any part $X \subseteq W$ of $\mathcal{Q}^\cup$, since by (A5) there are at most k parts $Y \subseteq X$ of $\mathcal{Q}_\mathbf{i}$ we have

$$\left(n_\mathbf{i}'\right)_X = \sum_{Y\in\mathcal{Q}_\mathbf{i}:Y\subseteq X} \left(\rho_\mathbf{i}\frac{|Y|}{|W|} \pm 1\right) = \begin{cases} \rho_\mathbf{i}\frac{|X|}{|W|} \pm 2k & \text{if } i_W = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

Now, crucially, we have $\mathbf{i}_{\mathcal{Q}^\cup}(V) \in L^{\mu_6}_{\mathcal{Q}^\cup}(H)$. Indeed, $\mathcal{P}$ is $(1/\ell, 2c, \mu_6, \mu'')$-robustly maximal with respect to H by (A7), $\mathcal{Q}^\cup$ is a refinement of $\mathcal{P}$ with parts of size at least $n/\ell \geq 2c|V(H)|$ by (A5), and $(\mathbf{i}_{\mathcal{Q}^\cup}(V) \mid \mathcal{P}) = \mathbf{i}_{\mathcal{P}}(V) \in L = L^{\mu''}_{\mathcal{P}}(H)$ by (A4) and (A7). So Proposition 5.4 implies that $\mathbf{i}_{\mathcal{Q}^\cup}(V) \in L^{\mu_6}_{\mathcal{Q}^\cup}(H)$, as desired. Since $\mathbf{n}'_{\mathbf{i}} \in L^{\mu_6}_{\mathcal{Q}^\cup}(H_{\mathbf{i}}) \subseteq L^{\mu_6}_{\mathcal{Q}^\cup}(H)$ for each $\mathbf{i} \in I$, we deduce that

$$\mathbf{v} := \mathbf{i}_{\mathcal{Q}^\cup}(V) - \sum_{\mathbf{i}\in I} \mathbf{n}'_{\mathbf{i}} \in L^{\mu_6}_{\mathcal{Q}^\cup}(H).$$

Furthermore, for any $W \in \mathcal{P}$ by (B1) we have $\sum_{\mathbf{i}\in I: i_W=1} \rho_{\mathbf{i}} = |W|$, so for any $X \in \mathcal{Q}^\cup$ with $X \subseteq W$ by (2) we have

$$v_X = |X| - \sum_{\mathbf{i}\in I: i_W=1} \left(\rho_{\mathbf{i}} \frac{|X|}{|W|} \pm 2k\right) = 0 \pm 2k|I|.$$

Since $|\mathcal{Q}^\cup| \leq k^2$ and $|I| \leq r^{k-1}$ we have $\mathbf{v} \in B(\mathbf{0}, 2k^3 r^{k-1})$, so we may apply Proposition 4.8 to obtain integers $a_{\mathbf{i}'}$ with $|a_{\mathbf{i}'}| \leq k^{-k}K$ for each $\mathbf{i}' \in I^{\mu_6}_{\mathcal{Q}^\cup}(H)$ such that $\sum_{\mathbf{i}'\in I^{\mu_6}_{\mathcal{Q}^\cup}(H)} a_{\mathbf{i}'}\mathbf{i}' = \mathbf{v}$.

For each $\mathbf{i} \in I$, we define $\mathbf{n}^\cup_{\mathbf{i}} := \mathbf{n}'_{\mathbf{i}} + \sum a_{\mathbf{i}'}\mathbf{i}'$, where the sum is taken over all $\mathbf{i}' \in I^{\mu_6}_{\mathcal{Q}^\cup}(H)$ with $(\mathbf{i}' \mid \mathcal{P}) = \mathbf{i}$. Then $\mathbf{n}^\cup_{\mathbf{i}}$ is a linear combination of vectors in $L^{\mu_6}_{\mathcal{Q}^\cup}(H_{\mathbf{i}})$, so (C2) holds. Also, (C1) holds as

$$\mathbf{i}_{\mathcal{Q}^\cup}(V) - \sum_{\mathbf{i}\in I} \mathbf{n}^\cup_{\mathbf{i}} = \mathbf{i}_{\mathcal{Q}^\cup}(V) - \sum_{\mathbf{i}\in I} \mathbf{n}'_{\mathbf{i}} - \mathbf{v} = \mathbf{0}.$$

Finally, consider any $\mathbf{i} \in I$, any part W of $\mathcal{P}$ with $i_W = 1$ and any part $X \subseteq W$ of $\mathcal{Q}^\cup$. Since there are at most $(k-1)^k$ vectors $\mathbf{i}' \in I^{\mu_6}_{\mathcal{Q}^\cup}(H)$ with $i'_X = 1$, by definition of $\mathbf{n}^\cup_{\mathbf{i}}$ we have

$$\left|\left(n^\cup_{\mathbf{i}}\right)_X - \left(n'_{\mathbf{i}}\right)_X\right| \leq (k-1)^k \max_{\mathbf{i}'} |a_{\mathbf{i}'}| \leq (1-1/k)^k K;$$

together with (2) we have (C3). This completes the proof of Claim 7.9.

In the final claim of this step we determine the required part sizes for the random subgraphs of the next step. For proof, we start by choosing provisional values for $\mathbf{n}^\cap_{\mathbf{i}}$ to satisfy (D1) and (D3). If (D2) fails then (C2) will imply that it can be remedied by transferrals between pairs of parts of $\mathcal{Q}$ that lie in the same part of $\mathcal{Q}^\cup$. By definition of $\mathcal{Q}^\cup$ there is a path in the auxiliary graph G between these parts. By repeated 'swaps' we can effectively move the required adjustment along the edges of the path, all the while preserving (D1) and (D3), until (D2) holds. Throughout the proof, we identify the vectors $\mathbf{n}^\cap_{\mathbf{i}}$ and $\mathbf{n}^\cup_{\mathbf{i}}$ with their restrictions to the parts of $\mathcal{Q}^\cap$ and $\mathcal{Q}^\cup$ that are contained within parts of $\mathcal{P}_{\mathbf{i}}$ (they are zero on all other parts).

Claim 7.10. *There exist vectors* $\mathbf{n}_{\mathbf{i}}^{\cap}$ *with respect to* $\mathcal{Q}^{\cap}$ *for* $\mathbf{i} \in I$ *which satisfy*

(D1) $\sum_{\mathbf{i}\in I} \mathbf{n}_{\mathbf{i}}^{\cap} = \mathbf{i}_{\mathcal{Q}^{\cap}}(V)$,
(D2) $(\mathbf{n}_{\mathbf{i}}^{\cap} \mid \mathcal{Q}_{\mathbf{i}}) \in L_{\mathbf{i}}$ *for every* $\mathbf{i} \in I$, *and*
(D3) *For any part* W *of* $\mathcal{P}$ *and any part* $Z \subseteq W$ *of* $\mathcal{Q}^{\cap}$ *we have*

$$\left(n_{\mathbf{i}}^{\cap}\right)_Z = \begin{cases} \rho_{\mathbf{i}} \frac{|Z|}{|W|} \pm 2K & \text{if } i_W = 1 \\ 0 & \text{otherwise.} \end{cases}$$

To prove the claim, we start by choosing provisional values for $\mathbf{n}_{\mathbf{i}}^{\cap}$, where for each $\mathbf{i} \in I$, each part X of $\mathcal{Q}^{\cup}$ contained within a part of $\mathcal{P}_{\mathbf{i}}$, and each part $Z \subseteq X$ of $\mathcal{Q}^{\cap}$, we take $(n_{\mathbf{i}}^{\cap})_Z$ to be either $\lfloor (n_{\mathbf{i}}^{\cup})_X |Z|/|X| \rfloor$ or $\lceil (n_{\mathbf{i}}^{\cup})_X |Z|/|X| \rceil$, with choices made so that

$$\sum_{\mathbf{i}\in I} \left(n_{\mathbf{i}}^{\cap}\right)_Z = |Z| \quad \text{for each } Z \in \mathcal{Q}^{\cap}, \quad \text{and} \tag{3}$$

$$\sum_{Z\subseteq X} \left(n_{\mathbf{i}}^{\cap}\right)_Z = \left(n_{\mathbf{i}}^{\cup}\right)_X \quad \text{for each } X \in \mathcal{Q}^{\cup}. \tag{4}$$

To see that we can make such choices, fix $X \in \mathcal{Q}^{\cup}$ and let A be the matrix with rows indexed by the index vectors $\mathbf{i} \in I$ such that X is contained within a part of $\mathcal{P}_{\mathbf{i}}$, and columns indexed by the parts $Z \subseteq X$ of $\mathcal{Q}^{\cap}$, where the $(\mathbf{i}, Z)$ entry of A is $(n_{\mathbf{i}}^{\cup})_X |Z|/|X|$. Then the sum of column Z is $|Z|$ by (C1), and the sum of row $\mathbf{i}$ is $(n_{\mathbf{i}}^{\cup})_X$. Thus Theorem 4.9 implies that we can choose such values for $\mathbf{n}_{\mathbf{i}}^{\cap}$. Note that Eq. (4) can be reformulated as

$$\left(\mathbf{n}_{\mathbf{i}}^{\cap} \mid \mathcal{Q}^{\cup}\right) = \mathbf{n}_{\mathbf{i}}^{\cup}.$$

Also, for any part W of $\mathcal{P}$ and any part $Z \subseteq W$ of $\mathcal{Q}^{\cap}$, it follows from (C3) that $(n_{\mathbf{i}}^{\cap})_Z$ is equal to $\rho_{\mathbf{i}}|Z|/|W| \pm (K+1)$ if $i_W = 1$ and is zero otherwise.

Eq. (3) implies that the current values for $\mathbf{n}_{\mathbf{i}}^{\cap}$ satisfy (D1) and (D3); we now modify them to satisfy (D2). Consider $\mathbf{i}_1, \mathbf{i}_2 \in I$, and suppose that Z and Z' are parts of $\mathcal{Q}^{\cap}$ which are subsets of distinct parts Y_1 and Y_1' respectively of $\mathcal{Q}_{\mathbf{i}_1}$, but that Z and Z' are subsets of the same part Y_2 of $\mathcal{Q}_{\mathbf{i}_2}$. We define an $(\mathbf{i}_1, \mathbf{i}_2, Z, Z')$-*swap* as the operation of increasing $(n_{\mathbf{i}_1}^{\cap})_Z$ and $(n_{\mathbf{i}_2}^{\cap})_{Z'}$ each by one, and decreasing $(n_{\mathbf{i}_2}^{\cap})_Z$ and $(n_{\mathbf{i}_1}^{\cap})_{Z'}$ by one (all other co-ordinates of $\mathbf{n}_{\mathbf{i}_1}^{\cap}$ and $\mathbf{n}_{\mathbf{i}_2}^{\cap}$ remain unchanged, as do all other vectors $\mathbf{n}_{\mathbf{i}}^{\cap}$). Clearly $\sum_{\mathbf{i}\in I} \mathbf{n}_{\mathbf{i}}^{\cap}$ is unchanged by any $(\mathbf{i}_1, \mathbf{i}_2, Z, Z')$-swap, as is $(\mathbf{n}_{\mathbf{i}}^{\cap} \mid \mathcal{Q}_{\mathbf{i}})$ for any $\mathbf{i} \neq \mathbf{i}_1, \mathbf{i}_2$. Further, the operation has no effect on $(\mathbf{n}_{\mathbf{i}_2}^{\cap} \mid \mathcal{Q}_{\mathbf{i}_2})$, since Z and Z' are contained in the same part of $\mathcal{Q}_{\mathbf{i}_2}$. However, $(\mathbf{n}_{\mathbf{i}_1}^{\cap} \mid \mathcal{Q}_{\mathbf{i}_1})$ *is* affected; specifically, an $(\mathbf{i}_1, \mathbf{i}_2, Z, Z')$-swap adds $\mathbf{u}_{Y_1} - \mathbf{u}_{Y_1'}$ to $(\mathbf{n}_{\mathbf{i}_1}^{\cap} \mid \mathcal{Q}_{\mathbf{i}_1})$. By performing several sequences of swaps, we shall ensure that (D2) is satisfied for each $\mathbf{i} \in I$ in turn.

Fix some $\mathbf{i} \in I$. Recall from (C2) that $\mathbf{n}_{\mathbf{i}}^{\cup} \in L_{\mathcal{Q}^{\cup}}^{\mu_6}(H_{\mathbf{i}})$. Since $\mu_0 \ll \mu_6$, Proposition 4.5(ii) applied with $\mathcal{Q}_{\mathbf{i}}$ and $\mathcal{Q}^{\cup}$ in place of $\mathcal{Q}$ and $\mathcal{P}$ respectively implies that there exists $\mathbf{n}_{\mathbf{i}}^{*} \in L_{\mathbf{i}}$ such that $(\mathbf{n}_{\mathbf{i}}^{*} \mid \mathcal{Q}^{\cup}) = \mathbf{n}_{\mathbf{i}}^{\cup}$, where we identify $\mathcal{Q}^{\cup}$ with its restriction to $V(H_{\mathbf{i}})$ and $\mathbf{n}_{\mathbf{i}}^{\cup}$ with its restriction to parts of $\mathcal{Q}^{\cup}[V(H_{\mathbf{i}})]$. We let $\mathbf{d}_{\mathbf{i}} = \mathbf{n}_{\mathbf{i}}^{*} - (\mathbf{n}_{\mathbf{i}}^{\cap} \mid \mathcal{Q}_{\mathbf{i}})$ and note

that $(\mathbf{d_i} \mid \mathcal{Q}^\cup) = \mathbf{0}$, since $(\mathbf{n_i^\cap} \mid \mathcal{Q}^\cup) = \mathbf{n_i^\cup}$. Hence we may write $\mathbf{d_i} = \sum_{\mathbf{v}\in\mathcal{D}_\mathbf{i}} \mathbf{v}$ for some sequence $\mathcal{D}_\mathbf{i}$ of transferrals, such that Y and Y' are parts of $\mathcal{Q}_\mathbf{i}$ which are contained in the same part of $\mathcal{Q}^\cup$ for each transferral $\mathbf{u}_Y - \mathbf{u}_{Y'} \in \mathcal{D}_\mathbf{i}$. Then $(\mathbf{n_i^\cap} \mid \mathcal{Q}_\mathbf{i}) + \sum_{\mathbf{v}\in\mathcal{D}_\mathbf{i}} \mathbf{v} = \mathbf{n_i^*} \in L_\mathbf{i}$. We now reduce $\mathcal{D}_\mathbf{i}$ to a sequence of at most $k-2$ transferrals, maintaining the property that $(\mathbf{n_i^\cap} \mid \mathcal{Q}_\mathbf{i}) + \sum_{\mathbf{v}\in\mathcal{D}_\mathbf{i}} \mathbf{v} \in L_\mathbf{i}$. To see that this is possible, recall that $I_\mathbf{i}$ is full with respect to $\mathcal{P}_\mathbf{i}$, and by (A5) every part of $\mathcal{Q}_\mathbf{i}$ has size at least $(1/\ell + \gamma/2)n$, so $|L_{\max}^{\mathcal{Q}_i\mathcal{P}_i}/L_\mathbf{i}| \le k-1$ by Proposition 6.12(i) and (iv). Thus the required reduction of $\mathcal{D}_\mathbf{i}$ follows from Proposition 6.10.

For each $\mathbf{u}_Y - \mathbf{u}_{Y'} \in \mathcal{D}_\mathbf{i}$, we carry out the following procedure. Choose parts Z and Z' of $\mathcal{Q}^\cap$ with $Z \subseteq Y$ and $Z' \subseteq Y'$. Since Y and Y' are contained in the same part of $\mathcal{Q}^\cup$, by definition of $\mathcal{Q}^\cup$ there is a path in the auxiliary graph G from Z to Z', i.e. we have $Z = Z_0, \dots, Z_p = Z'$ and $Y_1, \dots, Y_p$ for some $p \le |\mathcal{Q}^\cap| \le k^{2+k^k}$, such that $Z_j \in \mathcal{Q}^\cap$ for $0 \le j \le p$, and for each $j \in [p]$ there exists $\mathbf{i}_j \in I$ such that $Y_j \in \mathcal{Q}_{\mathbf{i}_j}$ and $Z_{j-1}, Z_j \subseteq Y_j$.

Note that each Z_j is contained in the same $W \in \mathcal{P}$ as Y and Y', and that by the definition of $\mathcal{D}_\mathbf{i}$ we have $i_W = 1$. Hence for each $0 \le j \le p$ there exists a part $Y_j^* \in \mathcal{Q}_\mathbf{i}$ which contains Z_j. Thus $Y_0^* = Y$ and $Y_p^* = Y'$.

For each $j \in [p]$ in turn we apply swaps as follows. If $Y_{j-1}^* = Y_j^*$ then there is no swap. Otherwise, we perform an $(\mathbf{i}, \mathbf{i}_j, Z_{j-1}, Z_j)$-swap; as noted above, the only effect of this is to add $\mathbf{u}_{Y_{j-1}^*} - \mathbf{u}_{Y_j^*}$ to $(\mathbf{n_i^\cap} \mid \mathcal{Q}_\mathbf{i})$. So the net effect of performing these swaps for every member of $\mathcal{D}_\mathbf{i}$ is to add $\mathbf{u}_Y - \mathbf{u}_{Y'}$ to $(\mathbf{n_i^\cap} \mid \mathcal{Q}_\mathbf{i})$. By choice of Y and Y', after these modifications we have $(\mathbf{n_i^\cap} \mid \mathcal{Q}_\mathbf{i}) \in L_\mathbf{i}$, and crucially, $(\mathbf{n_{i'}^\cap} \mid \mathcal{Q}_{\mathbf{i}'})$ is unchanged for any $\mathbf{i}' \ne \mathbf{i}$ and $\sum_{\mathbf{i}\in I} \mathbf{n_i^\cap}$ is unchanged.

We proceed in this manner for every $\mathbf{i} \in I$; then the vectors $\mathbf{n_i^\cap}$ obtained at the end of this process must satisfy (D2). Since (D1) held before we made any modifications, and $\sum_{\mathbf{i}\in I} \mathbf{n_i^\cap}$ is preserved by each modification, we conclude that (D1) still holds. Finally recall that for any $\mathbf{i} \in I$, any part W of $\mathcal{P}_\mathbf{i}$ and any part $Z \subseteq W$ of $\mathcal{Q}^\cap$ our provisional values satisfied $(n_\mathbf{i}^\cap)_Z = \rho_\mathbf{i}|Z|/|W| \pm (K+1)$. Recall that $|I| = r^{k-1}$; now (D3) follows, as we performed at most $p \le k^{2+k^k}$ swaps for each $\mathbf{v} \in \bigcup_{\mathbf{i}\in I} \mathcal{D}_\mathbf{i}$, so there were at most $p(k-2)r^{k-1} < K$ swaps in total, and no swap changed any co-ordinate of $\mathbf{n_i^\cap}$ by more than one. This completes the proof of Claim 7.10.

Step 6: The random selection. In this final step, we now partition V into disjoint sets $T_\mathbf{i}$ for $\mathbf{i} \in I$, where for each $Z \in \mathcal{Q}^\cap$ the number of vertices of $T_\mathbf{i}$ taken from Z is $(n_\mathbf{i}^\cap)_Z$. To ensure that this is possible, we require that $\sum_{\mathbf{i}\in I} \mathbf{n_i^\cap} = \mathbf{i}_{\mathcal{Q}^\cap}(V)$ and that each co-ordinate of each $\mathbf{n_i^\cap}$ is non-negative. The first of these conditions holds by (D1). For the second, observe that for any $\mathbf{i} \in I$, any part W of $\mathcal{P}_\mathbf{i}$ and any part $Z \subseteq W$ of $\mathcal{Q}^\cap$ we have $\rho_\mathbf{i} \ge \eta n$ by (B2), $|Z| \ge D$ by (1) and $|W| \le n$. So (D3) implies that

$$\left(n_\mathbf{i}^\cap\right)_Z \ge \rho_\mathbf{i}|Z|/|W| - 2K \ge \eta|Z|/2 \ge 0. \tag{5}$$

We choose such a partition uniformly at random. That is, for each part Z of $\mathcal{Q}^\cap$, we choose uniformly at random a partition of Z into sets $Z_\mathbf{i}$ for $\mathbf{i} \in I$ so that $|Z_\mathbf{i}| = (n_\mathbf{i}^\cap)_Z$

for each $\mathbf{i} \in I$. Then for each $\mathbf{i} \in I$ we let

$$T_{\mathbf{i}} := \bigcup_{Z \in \mathcal{Q}^\cap} Z_{\mathbf{i}}, \qquad \hat{H}_{\mathbf{i}} = H[T_{\mathbf{i}}], \qquad \hat{\mathcal{P}}_{\mathbf{i}} = \mathcal{P}_{\mathbf{i}}[T_{\mathbf{i}}] \quad \text{and} \quad \hat{\mathcal{Q}}_{\mathbf{i}} = \mathcal{Q}_{\mathbf{i}}[T_{\mathbf{i}}].$$

Note that $T_{\mathbf{i}} \subseteq V(H_{\mathbf{i}})$, as by (D3) $\mathbf{n}^\cap_{\mathbf{i}}$ is zero on parts of $\mathcal{Q}^\cap$ not contained in $V(H_{\mathbf{i}})$, so $\hat{H}_{\mathbf{i}}$ is a $\hat{\mathcal{P}}_{\mathbf{i}}$-partite k-graph on the vertex set $T_{\mathbf{i}}$. Also, since $(\mathbf{n}^\cap_{\mathbf{i}} \mid \mathcal{Q}_{\mathbf{i}}) \in L_{\mathbf{i}}$ by (D2), and every $\mathbf{i}' \in I_{\mathbf{i}}$ has $(\mathbf{i}' \mid \mathcal{P}) = \mathbf{i}$, we have $(\mathbf{n}^\cap_{\mathbf{i}} \mid \mathcal{P}) = t_{\mathbf{i}}\mathbf{i}$ for some integers $t_{\mathbf{i}}$. For any $W \in \mathcal{P}_{\mathbf{i}}$ with $i_W = 1$ we have

$$t_{\mathbf{i}} = \big(\mathbf{n}^\cap_{\mathbf{i}} \mid \mathcal{P}\big)_W = \sum_Z \big(n^\cap_{\mathbf{i}}\big)_Z \overset{\text{(D3)}}{=} \sum_Z \left(\frac{\rho_{\mathbf{i}}|Z|}{|W|} \pm 2K\right) = \rho_{\mathbf{i}} \pm 2K^2, \tag{6}$$

where both sums are taken over all parts Z of $\mathcal{Q}^\cap$ with $Z \subseteq W$, and we recall that $\mathcal{Q}^\cap$ has at most K parts. So in particular we have $\eta n/2 \le t_{\mathbf{i}} \le n$ for any $\mathbf{i} \in I$ by (B2).

Let $\varepsilon_* = \mu_2$, $\mu_* = \mu_3$, $\mu'_* = \mu_4$, $c_* = 6c/\eta$, $d_* = d/4$, and $\gamma_* = \gamma/4$.

Claim 7.11. *For any $\mathbf{i} \in I$ the following properties each hold with high probability.*

(E1) *All but at most $\varepsilon_* t_{\mathbf{i}}^{k-1}$ $\mathcal{P}_{\mathbf{i}}$-partite $(k-1)$-sets $S \subseteq T_{\mathbf{i}}$ have $|\hat{H}_{\mathbf{i}}(S)| \ge (1/(\ell-1)+\gamma_*)t_{\mathbf{i}}$.*
(E2) *$\hat{\mathcal{Q}}_{\mathbf{i}}$ is $(c_*, c_*, \mu_*, \mu'_*)$-robustly maximal with respect to $\hat{H}_{\mathbf{i}}$.*
(E3) *$L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}}) = L_{\mathbf{i}}$, and any vertex of $T_{\mathbf{i}}$ is in at least $d_* t_{\mathbf{i}}^{k-1}$ edges $e \in \hat{H}_{\mathbf{i}}$ with $\mathbf{i}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(e) \in L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}})$.*

We first recall from (A2) that $\mathcal{Q}_{\mathbf{i}}$ is $(1/\ell, 2c, \mu_2, \mu_5)$-robustly maximal with respect to $H_{\mathbf{i}}$, and from (5) that $(n^\cap_{\mathbf{i}})_Z \ge \eta|Z|/2$ for any $Z \in \mathcal{Q}^\cap$ contained in a part of $\mathcal{P}_{\mathbf{i}}$, so Lemma 5.5 implies that (E2) holds with high probability. For (E1), recall from (A1) that all but at most $\mu_0 n^{k-1}$ $\mathcal{P}_{\mathbf{i}}$-partite $(k-1)$-sets $S \subseteq V(H_{\mathbf{i}})$ have $|H_{\mathbf{i}}(S)| \ge (1/\ell + \gamma/2)n$. So it suffices to show that for any such S we have $|H_{\mathbf{i}}(S) \cap T_{\mathbf{i}}| \ge (1/(\ell-1)+\gamma_*)t_{\mathbf{i}}$ with high probability. Let W be the part of $\mathcal{P}$ such that $H_{\mathbf{i}}(S) \subseteq W$. Note that $|W| \le (1 - 1/\ell - \gamma/2)n$ as each part of $\mathcal{P}'$ is partitioned into $r \ge 2$ parts by $\mathcal{P}$, and each part of $\mathcal{P}$ has size at least $(1/\ell + \gamma/2)n$ by (A5). Now

$$\begin{aligned}
\mathbb{E}\big[\big|H_{\mathbf{i}}(S) \cap T_{\mathbf{i}}\big|\big] &= \sum_{Z \subseteq W} \mathbb{E}\big[\big|H_{\mathbf{i}}(S) \cap Z_{\mathbf{i}}\big|\big] = \sum_{Z \subseteq W} \frac{(n^\cap_{\mathbf{i}})_Z |H_{\mathbf{i}}(S) \cap Z|}{|Z|} \\
&\overset{\text{(D3)}}{\ge} \frac{\rho_{\mathbf{i}}|H_{\mathbf{i}}(S)|}{|W|} - 2K^2 \\
&\overset{(6)}{\ge} \frac{t_{\mathbf{i}}(1/\ell + \gamma/2)}{1 - 1/\ell - \gamma/2} - 4K^2 \\
&\ge (1+\varepsilon)\left(\frac{1}{\ell-1} + \gamma_*\right) t_{\mathbf{i}}.
\end{aligned}$$

Since $|H_{\mathbf{i}}(S) \cap T_{\mathbf{i}}|$ is a sum of independent hypergeometric random variables, with high probability $|H_{\mathbf{i}}(S) \cap T_{\mathbf{i}}| \geq (1/(\ell-1)+\gamma_*)t_{\mathbf{i}}$ by Corollary 4.11.

For (E3), recall from (A6) that $L_{\mathbf{i}} = L^{\mu_1}_{\mathcal{Q}_{\mathbf{i}}}(H_{\mathbf{i}}) = L^{\mu'}_{\mathcal{Q}_{\mathbf{i}}}(H_{\mathbf{i}})$. For any $\mathbf{i}' \in L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}})$ there are at least $\mu_*(kt_{\mathbf{i}})^k > \mu_1|V(H_{\mathbf{i}})|^k$ edges $e \in \hat{H}_{\mathbf{i}} \subseteq H_{\mathbf{i}}$ with $\mathbf{i}_{\mathcal{Q}_{\mathbf{i}}}(e) = \mathbf{i}'$, so we have $L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}}) \subseteq L^{\mu_1}_{\mathcal{Q}_{\mathbf{i}}}(H_{\mathbf{i}})$. Now suppose that $\mathbf{i}' \in L^{\mu'}_{\mathcal{Q}_{\mathbf{i}}}(H_{\mathbf{i}})$, so there are at least $\mu'|V(H_{\mathbf{i}})|^k \geq \mu' n^k$ edges $e \in H_{\mathbf{i}}$ with $\mathbf{i}_{\mathcal{Q}_{\mathbf{i}}}(e) = \mathbf{i}'$. Recall that $\mathcal{Q}^\cap$ has at most K parts, so at most $K\varepsilon(kn)^k$ of these edges contain a vertex from a part of $\mathcal{Q}^\cap$ of size at most εn. Fix one of the remaining edges $e = \{x_1, \ldots, x_k\}$ and for each $j \in [k]$ let Z_j and W_j be the parts of $\mathcal{Q}^\cap$ and $\mathcal{P}$ respectively which contain x_j. Since $|Z_j| \geq \varepsilon n$ for each $j \in [k]$, the probability that $e \in \hat{H}_{\mathbf{i}}$ equals

$$\prod_{j\in[k]} \frac{(n^\cap_{\mathbf{i}})_{Z_j}}{|Z_j|} \overset{\text{(D3)}}{=} \prod_{j\in[k]} \left(\frac{\rho_{\mathbf{i}}}{|W_j|} \pm \frac{2K}{\varepsilon n}\right) \overset{(6)}{>} \frac{t^k_{\mathbf{i}}}{n^k} - \frac{3kK}{\varepsilon n}.$$

Therefore the expected number of edges $e \in \hat{H}_{\mathbf{i}}$ with $\mathbf{i}_{\mathcal{Q}_{\mathbf{i}}}(e) = \mathbf{i}'$ is at least $(\mu' n^k - K\varepsilon(kn)^k)((t_{\mathbf{i}}/n)^k - 3kK/\varepsilon n) > \mu' t^k_{\mathbf{i}}/2$. By Corollary 4.13, with high probability there are at least $\mu' t^k_{\mathbf{i}}/3 \geq \mu_*|V(\hat{H}_{\mathbf{i}})|^k$ such edges, so we have $\mathbf{i}' \in L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}})$, proving that $L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}}) = L_{\mathbf{i}}$ as claimed.

Now consider any vertex $x \in T_{\mathbf{i}}$ and let $E(x)$ denote the number of edges $e \in \hat{H}_{\mathbf{i}}$ with $\mathbf{i}_{\mathcal{Q}_{\mathbf{i}}}(e) \in L_{\mathbf{i}}$ which contain x. To estimate $E(x)$, recall from (A3) that x is in at least $\frac{1}{2}dn^{k-1}$ edges $e \in H_{\mathbf{i}}$ with $\mathbf{i}_{\mathcal{Q}_{\mathbf{i}}}(e) \in L_{\mathbf{i}}$. The same argument as above shows that $\mathbb{E}(E(x)) \geq \frac{1}{3}dt^{k-1}_{\mathbf{i}}$ and so with high probability $E(x) \geq \frac{1}{4}dt^{k-1}_{\mathbf{i}}$. This implies (E3), as $L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}}) = L_{\mathbf{i}}$, so completes the proof of the claim.

To summarise, for each $\mathbf{i} \in I$ we have the following:

- $\hat{H}_{\mathbf{i}}$ is a $\hat{\mathcal{P}}_{\mathbf{i}}$-partite k-graph on $T_{\mathbf{i}}$ with parts of size $t_{\mathbf{i}}$.
- At most $\varepsilon_* t^{k-1}_{\mathbf{i}}$ $\hat{\mathcal{P}}_{\mathbf{i}}$-partite $(k-1)$-sets $S \subseteq T_{\mathbf{i}}$ have $|\hat{H}_{\mathbf{i}}(S)| < (1/(\ell-1)+\gamma_*)t_{\mathbf{i}}$.
- $\hat{\mathcal{Q}}_{\mathbf{i}}$ is a partition of $T_{\mathbf{i}}$ which refines $\hat{\mathcal{P}}_{\mathbf{i}}$ and is $(c_*, c_*, \mu_*, \mu'_*)$-robustly maximal with respect to $\hat{H}_{\mathbf{i}}$.
- Any vertex of $T_{\mathbf{i}}$ is in at least $d_* t^{k-1}_{\mathbf{i}}$ edges $e \in \hat{H}_{\mathbf{i}}$ with $\mathbf{i}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(e) \in L_{\mathbf{i}} = L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}})$.
- $\mathbf{i}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(T_{\mathbf{i}}) = (\mathbf{n}^\cap_{\mathbf{i}} \mid \mathcal{Q}_{\mathbf{i}}) \in L_{\mathbf{i}} = L^{\mu_*}_{\hat{\mathcal{Q}}_{\mathbf{i}}}(\hat{H}_{\mathbf{i}})$ by (D2).

Since

$$1/t_{\mathbf{i}} \ll \varepsilon_* \ll \mu_* \ll \mu'_* \ll c_*, d_* \ll \gamma_*, 1/k,$$

we conclude that $\hat{H}_{\mathbf{i}}$ contains a perfect matching $M_{\mathbf{i}}$ by our inductive hypothesis. Therefore $M \cup M' \cup \bigcup_{\mathbf{i}\in I} M_{\mathbf{i}}$ is a perfect matching in (the original) H. This completes the proof of Lemma 7.6. □

8. Proof of the structure theorem

In this section we use Lemma 7.1 to prove Theorem 1.10. We start by proving an easier version in the first subsection, which is not algorithmic, but has a cleaner statement, and will be used in the proof of the full theorem. In the second subsection we prove a weaker version of the algorithmic result, which provides some details omitted from the extended abstract of this paper [16]. The third and fourth subsections contain some technical preliminaries for the main result: some analysis of how full lattices behave under merging of parts, and some results on subsequence sums in abelian groups. The final subsection contains the proof of Theorem 1.10.

8.1. An easier version

In this subsection we prove the following easier version of Theorem 1.10.

Theorem 8.1. *Under Setup 1.7, H has a perfect matching if and only if every full pair $(\mathcal{P}, L)$ for H is soluble.*

While the statement of this version is more appealing than that of Theorem 1.10, it is impractical to use directly for our algorithm, as the number of full pairs for H may be exponentially large. The forward implication of Theorem 8.1 is easy to prove: if H has a perfect matching M then $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M)) = \mathbf{0} \in L$, and Lemma 6.9 implies that $(\mathcal{P}, L)$ is soluble. In fact the forward implication of Theorem 1.10 now follows immediately; we simply observe from the definition that there is no C-certificate for H for any $C \geq 0$. We now consider the backward implication of Theorem 8.1. We need the following non-partite analogue of Proposition 7.5 (we omit the similar proof).

Proposition 8.2. *Suppose that $1/n \ll \mu, \varepsilon \ll \psi, d, c, 1/k$, and let V be a set of size n. Let H be a k-graph on V in which at most εn^{k-1} $(k-1)$-sets $S \subseteq V$ have $d_H(S) < dn$, and let $\mathcal{P}$ be a partition of V with parts of size at least cn such that $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free. Then the following properties hold.*

(i) *$I^{\mu}_{\mathcal{P}}(H)$ is full.*
(ii) *For any $\mathbf{i} \in I^{\mu}_{\mathcal{P}}(H)$ at most ψn^{k-1} $(k-1)$-sets $S \subseteq V$ with $\mathbf{i}_{\mathcal{P}}(S) = \mathbf{i} - \mathbf{u}_X$ for some $X \in \mathcal{P}$ have $|H_{\mathbf{i}}(S)| < (d-\psi)n$.*
(iii) *Each part of $\mathcal{P}$ has size at least $(d-\psi)n$.*
(iv) *$I^{dc^{k-1}/2k!}_{\mathcal{P}}(H) = I^{\mu}_{\mathcal{P}}(H)$.*
(v) *For any $\mathbf{i} \in I^{\mu}_{\mathcal{P}}(H)$ at most ψn vertices of $V(H_{\mathbf{i}})$ lie in fewer than $c^{k-2}dn^{k-1}/2(k-1)!$ edges of $H_{\mathbf{i}}$.*

Next we show that we can modify a robustly maximal partition to arrange that every vertex belongs to many edges with index on the robust edge-lattice. Note that it would

not help to use the technique employed in previous sections of removing a small matching covering the bad vertices, as we will require the full vertex set to remain intact in order to use the fact that every full pair is soluble.

Lemma 8.3. *Suppose that $k \geq 3$ and $1/n \ll \varepsilon, \mu \ll \mu' \ll c \ll \gamma \ll 1/k$. Let H be a k-graph on n vertices such that at most εn^{k-1} $(k-1)$-sets $A \subseteq V(H)$ have $d_H(A) < (1+\gamma)n/k$ and every vertex is in at least γn^{k-1} edges of H. Also let $\mathcal{P}$ be a partition of $V(H)$ that is (c, c, μ, μ')-robustly maximal with respect to H. Then there exists a partition $\mathcal{P}'$ of $V(H)$ with at most $k-1$ parts such that*

(i) *$\mathcal{P}'$ is $(4c, 2c, 4\sqrt{\mu}, \mu'/2)$-robustly maximal with respect to H,*
(ii) *$L^{6\sqrt{\mu}}_{\mathcal{P}'}(H) = L^{4\sqrt{\mu}}_{\mathcal{P}'}(H) = L^{\mu}_{\mathcal{P}}(H)$, and*
(iii) *every vertex is in at least $\gamma n^{k-1}/3k$ edges $e \in H$ with $\mathbf{i}_{\mathcal{P}'}(e) \in L^{4\sqrt{\mu}}_{\mathcal{P}'}(H)$.*

Proof. First we note by Proposition 8.2(iii) that every part of $\mathcal{P}$ has size at least $n/k + \gamma n - cn > n/k$. So $|\mathcal{P}| \leq k-1$, and $\mathcal{P}$ is $(1/k, c, \mu, \mu')$-robustly maximal with respect to H. Next let B be the set of vertices that belong to fewer than $\gamma n^{k-1}/2$ edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) \in L^{\mu}_{\mathcal{P}}(H)$. We claim that $|B| < \sqrt{\mu}n/2$. For otherwise, by the inclusion–exclusion principle the number of edges e that intersect B and have $\mathbf{i}_{\mathcal{P}}(e) \notin L^{\mu}_{\mathcal{P}}(H)$ is at least $(\sqrt{\mu}n/2) \cdot \gamma n^{k-1}/2 - \binom{\sqrt{\mu}n/2}{2}\binom{n}{k-2} \geq \sqrt{\mu}\gamma n^k/5$. But this is a contradiction, as there are fewer than $k^k \mu n^k$ such edges in total, so we must have $|B| < \sqrt{\mu}n/2$.

Now consider any vertex $v \in B$ and let X be the part of $\mathcal{P}$ containing v. For each edge $e \in H$ we let $g_X(e)$ be the unique part X' such that $\mathbf{i}_{\mathcal{P}}(e) - \mathbf{u}_X + \mathbf{u}_{X'} \in L^{\mu}_{\mathcal{P}}(H)$; this is well-defined as $I^{\mu}_{\mathcal{P}}(H)$ is full by Proposition 8.2(i), so we can apply Lemma 6.1(ii). By the pigeonhole principle there exists $X(v) \in \mathcal{P}$ such that $g_X(e) = X(v)$ for at least $\gamma n^{k-1}/2k$ edges $e \in H$ containing v. Let $\mathcal{P}'$ be the partition obtained from $\mathcal{P}$ by moving v into $X(v)$ for each $v \in B$. Then $\mathcal{P}'$ is $\sqrt{\mu}$-close to $\mathcal{P}$, so by Lemma 5.3, $\mathcal{P}'$ is $(4c, 2c, 4\sqrt{\mu}, \mu'/2)$-robustly maximal with respect to H. Also, by Proposition 4.3 we have $L^{\mu^{1/3}}_{\mathcal{P}}(H) \subseteq L^{6\sqrt{\mu}}_{\mathcal{P}'}(H) \subseteq L^{4\sqrt{\mu}}_{\mathcal{P}'}(H) \subseteq L^{\mu}_{\mathcal{P}}(H)$; since $L^{\mu^{1/3}}_{\mathcal{P}}(H) = L^{\mu}_{\mathcal{P}}(H)$ by Proposition 8.2(iv) we have $L^{\mu}_{\mathcal{P}}(H) = L^{4\sqrt{\mu}}_{\mathcal{P}'}(H) = L^{6\sqrt{\mu}}_{\mathcal{P}'}(H)$. Finally, for any vertex $v \in V(H)$ the number of edges $e \in H$ containing v with $\mathbf{i}_{\mathcal{P}'}(e) \in L^{4\sqrt{\mu}}_{\mathcal{P}'}(H)$ is at least $\gamma n^{k-1}/2k - \sqrt{\mu}n^{k-1} \geq \gamma n^{k-1}/3k$. □

Proof of Theorem 8.1. We have already proved the forward implication, so it remains to prove the backward implication. Consider H as in Setup 1.7 and suppose that every full pair $(\mathcal{P}, L)$ is soluble. Introduce a new constant c with $\varepsilon \ll c \ll \gamma$, fix $s = \lfloor 1/c \rfloor$ and introduce further new constants $\mu_1, \ldots, \mu_{s+1}$ such that $\varepsilon \ll \mu_1 \ll \cdots \ll \mu_{s+1} \ll c$. By Proposition 5.2 there exists $t \in [s]$ and a partition $\mathcal{P}$ of $V(H)$ that is (c, c, μ_t, μ_{t+1})-robustly maximal with respect to H; we write $\mu = \mu_t$ and $\mu' = \mu_{t+1}$. Applying Lemma 8.3 we obtain a partition $\mathcal{P}'$ of $V(H)$ with at most $k-1$ parts which is $(4c, 2c, 4\sqrt{\mu}, \mu'/2)$-robustly maximal with respect to H, such that $L := L^{6\sqrt{\mu}}_{\mathcal{P}'}(H) = L^{4\sqrt{\mu}}_{\mathcal{P}'}(H) = L^{\mu}_{\mathcal{P}}(H)$ and ev-

ery vertex $v \in V(H)$ is in at least $\gamma n^{k-1}/3k$ edges $e \in H$ with $\mathbf{i}_{\mathcal{P}'}(e) \in L$. Note that $(\mathcal{P}', L)$ is a full pair for H by Proposition 8.2(i), so by assumption it has a solution M. Taking $V' := V \setminus V(M)$ we then have $\mathbf{i}_{\mathcal{P}'}(V') \in L$. Also, by Proposition 4.3 we have $L = L^{6\sqrt{\mu}}_{\mathcal{P}'}(H) \subseteq L^{5\sqrt{\mu}}_{\mathcal{P}'[V']}(H[V']) \subseteq L^{4\sqrt{\mu}}_{\mathcal{P}'}(H) = L$, so $L = L^{5\sqrt{\mu}}_{\mathcal{P}'[V']}(H[V'])$, and by Lemma 5.3 $\mathcal{P}'[V']$ is $(3c, 3c, 5\sqrt{\mu}, \mu'/3)$-robustly maximal with respect to H'. We may therefore apply Lemma 7.1 to $H \setminus V(M)$ with $\gamma/2$, $3c$, $5\sqrt{\mu}$, $\mu'/3$, $\gamma/4k$ and $2k^{k-1}\varepsilon$ in place of γ, c, μ, μ', d and ε respectively to obtain a perfect matching in $H[V']$. Together with M this gives a perfect matching in H. □

8.2. A slower algorithm

As a warmup for Theorem 1.10, and to provide the details for the result given in the extended abstract of this paper [16], we first prove a weaker result in which $2k(k-3)$ is replaced by $2k^{k+3}$; this is sufficient for our algorithmic result, but the running time is of course significantly worse. (Actually, in the extended abstract we assumed instead that any full pair $(\mathcal{P}^*, L^*)$ for which any matching of edges $e \in H$ with $\mathbf{i}_{\mathcal{P}^*}(e) \notin L^*$ has size less than $2k^{k+2}$ is soluble, but the proof shows that this gives the desired conclusion.) This weaker result follows from Theorem 8.1 and the following result.

Lemma 8.4. *Suppose that $k \geq 3$ and H is a k-graph such that every $2k^{k+3}$-far full pair for H is soluble. Then every full pair for H is soluble.*

Proof. Consider any full pair $(\mathcal{P}, L)$ for H. Let I be a full set in L, and let S be the set of k-vectors $\mathbf{i}$ such that H contains at least $2k^2$ disjoint edges $e \in H$ with $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$. Let L' be the lattice generated by $I \cup S$ (so $L \subseteq L'$). Consider the relation $\sim$ on $\mathcal{P}$ defined by $X \sim Y$ if and only if $\mathbf{u}_X - \mathbf{u}_Y \in L'$; it is clear that $\sim$ is an equivalence relation. Let $\mathcal{P}^*$ be the partition of V formed by taking unions of equivalence classes under $\sim$ and let $L^* = (L' \mid \mathcal{P}^*)$.

We claim that $(\mathcal{P}^*, L^*)$ is a $2k^{k+3}$-far full pair for H. To see this, we first show that L^* is transferral-free. Suppose for a contradiction that $\mathbf{u}_W - \mathbf{u}_Z \in L^*$ for some distinct parts $W, Z \in \mathcal{P}^*$. By definition there exists $\mathbf{v} \in L'$ such that $(\mathbf{v} \mid \mathcal{P}^*) = \mathbf{u}_W - \mathbf{u}_Z$. Consider such $\mathbf{v}$ that minimises $\sum_{X \in \mathcal{P}} |v_X|$. Then we cannot have $v_X > 0$ and $v_Y < 0$ for distinct parts $X, Y \in \mathcal{P}$ that are contained in the same part of $\mathcal{P}^*$. This implies that $\mathbf{v} = \mathbf{u}_X - \mathbf{u}_Y$ for some $X, Y \in \mathcal{P}$ with $X \subseteq W$ and $Y \subseteq Z$, so $X \sim Y$, which contradicts the definition of $\mathcal{P}^*$, so L^* is transferral-free. Next, the projections of vectors in I show that L^* is full. Now, since there are at most k^k possible values of $\mathbf{i}_{\mathcal{P}}(e)$ for an edge $e \in H$, by choice of S, any matching of edges $e \in H$ such that $\mathbf{i}_{\mathcal{P}^*}(e) \notin L^*$ has size less than $2k^{k+2}$. Therefore $(\mathcal{P}^*, L^*)$ is a $2k^{k+3}$-far full pair for H, as claimed.

By assumption $(\mathcal{P}^*, L^*)$ is soluble, so H has a matching M of size at most $k-2$ such that $\mathbf{i}_{\mathcal{P}^*}(V(H) \setminus V(M)) \in L^*$. We claim that in fact $\mathbf{i} := \mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M)) \in L'$. To see this, we apply Lemma 6.1(ii) to get $\mathbf{i} - \mathbf{u}_X + \mathbf{u}_{X'} \in L$ for some parts X, X' of $\mathcal{P}$. It follows that $\mathbf{i}_{\mathcal{P}^*}(V(H) \setminus V(M)) - (\mathbf{u}_X - \mathbf{u}_{X'} \mid \mathcal{P}^*) \in L^*$, so $(\mathbf{u}_X - \mathbf{u}_{X'} \mid \mathcal{P}^*) \in L^*$. Since

L^* is transferral-free, $(\mathbf{u}_X - \mathbf{u}_{X'} \mid \mathcal{P}^*) = \mathbf{0}$, so X and X' are contained in the same part of $\mathcal{P}^*$. This implies $\mathbf{u}_X - \mathbf{u}_{X'} \in L'$, so $\mathbf{i} \in L'$, as claimed.

Next let $G = L'/L$ and note that $|G| \leq |G(\mathcal{P}, L)| \leq k-1$ by Lemma 6.4. There exists $r \in \mathbb{N}$ such that we can write $\mathbf{i} = \mathbf{v} + \sum_{j\in[r]} \mathbf{i}_j$ for some $\mathbf{v} \in L$, where either $\mathbf{i}_j \in S$ or $-\mathbf{i}_j \in S$ for $j \in [r]$. Then $\mathbf{i} + L = \sum_{j\in[r]}(\mathbf{i}_j + L) \in G$, where without loss of generality $\mathbf{i}_j \in S$ for each $j \in [r]$, as for any $\mathbf{i}' \in -S$ we can replace $\mathbf{i}' + L$ by $|G| - 1$ copies of $-\mathbf{i}' + L$. By Proposition 6.10, without loss of generality $r \leq k-2$. Now by definition of S, we can greedily extend M to a matching M' where for each $j \in [r]$ we add an edge e with $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}_j$. Then $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M')) \in L$, so $(\mathcal{P}, L)$ is soluble by Lemma 6.9. □

8.3. Merging

In Lemma 8.16 we will improve Lemma 8.4 by replacing $2k^{k+3}$ by $2k(k-3)$. This will require a more careful analysis of the lattices that can be obtained by merging parts.

Definition 8.5. Suppose $(\mathcal{P}, L)$ is a full pair for a k-graph H and write $G = G(\mathcal{P}, L)$. Fix an identification of $\mathcal{P}$ with G and an element $g_0 \in G$ so that $L = L(G, g_0)$ (using Theorem 6.7). For any subgroup K of G, let $\mathcal{P}_K$ be the partition obtained from $\mathcal{P}$ by merging any two parts identified with elements in the same coset of K. Let $L_K = (L \mid \mathcal{P}_K)$.

For example $\mathcal{P}_0 = \mathcal{P}$, $L_0 = L$, $\mathcal{P}_G = \{V(H)\}$ and $L_G = k\mathbb{Z}$.

Lemma 8.6. *Under the setup of Definition 8.5, $\mathcal{P}_K$ is well-defined, $(\mathcal{P}_K, L_K)$ is a full pair, and*

$$L_K = L(G/K, g_0 + K).$$

Furthermore, for $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ we have $(\mathbf{i} \mid \mathcal{P}_K) \in L_K$ if and only if $R_G(\mathbf{i}) \in K$.

Proof. To see that $\mathcal{P}_K$ is well-defined, recall from Remark 6.8 that the identification $\pi : \mathcal{P} \to G$ is determined up to translation by G; we merge X and Y if and only if $\pi(X) - \pi(Y) \in K$, and this property is invariant under any translation of π. Note that there is an induced identification of $\mathcal{P}_K$ with G/K. Now we show that $L_K = L(G/K, g_0 + K)$; then $(\mathcal{P}_K, L_K)$ is a full pair by Lemma 6.6.

Recall from Lemma 6.6 that $L(G, g_0)$ is the set of index vectors $\mathbf{i} \in L^{\mathcal{P}}_{\max}$ with $\sum_{g\in G} i_g g = (k^{-1}\sum_{g\in G} i_g)g_0$. For any such $\mathbf{i}$ we have

$$\sum_{h\in G/K} (\mathbf{i} \mid \mathcal{P}_K)_h h = \sum_{h\in G/K} \left(\sum_{g\in h} i_g\right) h = \sum_{g\in G} i_g(g + K) = \left(k^{-1}\sum_{g\in G} i_g\right)(g_0 + K)$$
$$= \left(k^{-1}\sum_{h\in G/K} (\mathbf{i} \mid \mathcal{P}_K)_h\right)(g_0 + K).$$

Thus $(\mathbf{i} \mid \mathcal{P}_K) \in L(G/K, g_0 + K)$. Since $R_G(\mathbf{i}) = \sum_{g\in G} i_g g - (k^{-1}\sum_{g\in G} i_g)g_0$ for any $\mathbf{i} \in L^{\mathcal{P}}_{\max}$, this calculation also establishes the 'Furthermore' statement.

Conversely, consider any $\mathbf{j} \in L(G/K, g_0 + K)$, so that $\sum_{h\in G/K} j_h h = (k^{-1}\sum_{h\in G/K} j_h)(g_0 + K)$. For each $h \in G/K$ fix a representative $\hat{h} \in h$ and consider the index vector $\mathbf{i} = \sum_{h\in G/K} j_h \mathbf{u}_{\hat{h}}$ with respect to $\mathcal{P}$. Then $(\mathbf{i} \mid \mathcal{P}_K) = \mathbf{j}$ and

$$\sum_{g\in G} i_g g = \sum_{h\in G/K} j_h \hat{h} \in \Bigl(k^{-1}\sum_{h\in G/K} j_h\Bigr)(g_0 + K)$$
$$= \Bigl(k^{-1}\sum_{g\in G} i_g\Bigr)(g_0 + K).$$

Let $g' = \sum_{g\in G} i_g g - (k^{-1}\sum_{g\in G} i_g)g_0$, so $g' \in K$, and let $\mathbf{i}' = \mathbf{i} - \mathbf{u}_{g'} + \mathbf{u}_{0_G}$. Then

$$\sum_{g\in G} i'_g g = -g' + \sum_{g\in G} i_g g = \Bigl(k^{-1}\sum_{g\in G} i_g\Bigr)g_0 = \Bigl(k^{-1}\sum_{g\in G} i'_g\Bigr)g_0,$$

so $\mathbf{i}' \in L(G, g_0)$, and $(\mathbf{i}' \mid \mathcal{P}_K) = (\mathbf{i} \mid \mathcal{P}_K) = \mathbf{j}$. □

8.4. Subsequence sums in finite abelian groups

Under the setup of Definition 8.5, given a matching M in H, we define $\mathcal{S}_G(M)$ to be the set of residues of submatchings of M; that is,

$$\mathcal{S}_G(M) = \bigl\{R_G\bigl(V\bigl(M'\bigr)\bigr) \mid M' \subseteq M\bigr\}.$$

Note that $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M')) \in L$ if and only if $R_G(V(H)) = R_G(V(M'))$, and if $R_G(V(H)) \in \mathcal{S}_G(M)$ then $(\mathcal{P}, L)$ is soluble by Lemma 6.9. Thus we have the following result.

Proposition 8.7. *Under the setup of Definition 8.5, if there is some matching M such that $\mathcal{S}_G(M) = G$ then $(\mathcal{P}, L)$ is soluble.*

To analyse $\mathcal{S}_G(M)$ we require some lemmas on subsequence sums in abelian groups. Roughly speaking, we can obtain sufficient lower bounds provided that $|\mathcal{S}_G(M)|$ increases by at least 1 when we add an edge to M. Our first lemma gives structural information when this is not the case. First we introduce some notation for sequences.

For the remainder of this subsection we fix any finite abelian group G.

Definition 8.8. Let a be a sequence in G.

(i) We write $\mathcal{S}(a)$ for the set of sums of all subsequences of a.
(ii) For any subgroup K of G we write $a + K$ for the sequence of K-cosets of terms in a.

(iii) For $j \geq 1$ we write a_j for the jth term of a (when it exists).
(iv) For $j \geq 0$ we write $a^j = (a_1, \dots, a_j)$.
(v) We write $a \circ b$ for the concatenation of two sequences a and b.

Lemma 8.9. *Suppose a is a sequence in G, and $x \in G$ is such that $\mathcal{S}(a \circ (x)) = \mathcal{S}(a)$. Then $\mathcal{S}(a)$ is a union of cosets of $\langle x \rangle$.*

Proof. Consider $y \in \mathcal{S}(a)$. It suffices to show that $y + rx \in \mathcal{S}(a)$ for every $r \in \mathbb{N}$. We prove this by induction on r. The base case $r = 0$ is given. For the induction step, if $y + rx \in \mathcal{S}(a)$ then $y + (r+1)x = (y + rx) + x \in \mathcal{S}(a \circ (x)) = \mathcal{S}(a)$. □

Remark 8.10. Lemma 8.16 follows directly from Lemma 8.9 for full pairs $(\mathcal{P}, L)$ such that $|G(\mathcal{P}, L)| = p$ is prime. Indeed, consider a greedy construction of a matching M, where at each step we add an edge that increases the size of $\mathcal{S}_G(M)$, if possible. The construction terminates in at most $|G| - 1$ steps, because either $\mathcal{S}_G(M) = G$ or no edge increases the size of $\mathcal{S}_G(M)$. If $\mathcal{S}_G(M) = G$ then $(\mathcal{P}, L)$ is soluble by Proposition 8.7. On the other hand, if $\mathcal{S}_G(M) \neq G$ we claim that for any edge e disjoint from M we have $R_G(e) = 0_G$, i.e. $\mathbf{i}_{\mathcal{P}}(e) \in L$. For otherwise, since p is prime we have $\langle R_G(e) \rangle = G$, so by Lemma 8.9 $\mathcal{S}_G(M)$ is a union of cosets of G, i.e. $\mathcal{S}_G(M) = G$, which contradicts our assumption. Note also that $|M| \leq |G| - 2 \leq k - 3$, so $|V(M)| \leq k(k-3)$. Thus $(\mathcal{P}, L)$ is a $k(k-3)$-far full pair, so is soluble by the assumption of Lemma 8.16.

Our next lemma gives a monotonicity property that will allow us to extend matchings, while maintaining the coset structure for subsequence sums.

Lemma 8.11. *Suppose a is a sequence in G and $\mathcal{S}(a)$ is a union of cosets of some subgroup K. Then $\mathcal{S}(a \circ b)$ is a union of cosets of K for any sequence b in G.*

Proof. Consider $y \in \mathcal{S}(a \circ b)$, and write $y = y_1 + y_2$, such that $y_1 \in \mathcal{S}(a)$ and $y_2 \in \mathcal{S}(b)$. Let $z \in y + K$, and write $z = y + h$ for $h \in K$. Then $y_1 + h \in \mathcal{S}(a)$, and hence $z = (y_1 + h) + y_2 \in \mathcal{S}(a \circ b)$. □

We deduce a property of sequences with the following minimality property.

Definition 8.12. Suppose a is a sequence in G. We say that a is *minimal* if $\mathcal{S}(a') \neq \mathcal{S}(a)$ for any proper subsequence a' of a.

Lemma 8.13. *Suppose a is a minimal sequence in G and K is a subgroup of G. Then a has at most $|K| - 1$ elements in K.*

Proof. Let b and c be the subsequences of a consisting of its elements that are in K and not in K respectively. It suffices to prove that $|\mathcal{S}(b^{j+1})| > |\mathcal{S}(b^j)|$ for $j \geq 0$, as then b has length at most $|\mathcal{S}(b)| - 1 \leq |K| - 1$. Suppose for a contradiction that $|\mathcal{S}(b^{j+1})| = |\mathcal{S}(b^j)|$

for some $j \geq 0$. Let a' and b' be obtained from a and b by deleting b_{j+1}. We will show that $\mathcal{S}(a) = \mathcal{S}(a')$, contradicting minimality of a. Consider any $y \in \mathcal{S}(a)$. We claim that $y \in \mathcal{S}(a')$. We can assume y is a subsequence sum using b_{j+1}. Then $y = b_{j+1} + y_1 + y_2$ with $y_1 \in \mathcal{S}(b')$ and $y_2 \in \mathcal{S}(c)$. Note that $\mathcal{S}(b^j)$ is a union of cosets of $\langle b_{j+1} \rangle$ by Lemma 8.9, and the same is true of $\mathcal{S}(b')$ by Lemma 8.11. Thus $b_{j+1} + y_1 \in \mathcal{S}(b')$, so $y \in \mathcal{S}(a')$, as required. □

Now we associate the following key subgroup with a sequence (Lemma 8.15 will show that it is well-defined).

Definition 8.14. Suppose a is a sequence in G. The *key subgroup* $K(a)$ for a is the unique subgroup K such that $\mathcal{S}(a)$ is a union of cosets of K and K is maximal with this property.

Note that the key subgroup always exists since every subset of G is a union of cosets of $\{0\}$. The following lemma shows that it is unique, and derives some of its properties.

Lemma 8.15. *Suppose a is a sequence in G. Then the key subgroup $K(a)$ is well-defined, and*

(i) *if $|G| > 1$ and a has at least $|G| - 1$ nonzero entries then $|K(a)| > 1$,*
(ii) *$K(a + K(a))$ is the trivial subgroup of $G/K(a)$,*
(iii) *if $|G/K(a)| > 1$ then a has at most $|G/K(a)| - 2$ elements not in $K(a)$.*

Proof. To show that $K(a)$ is well-defined, it suffices to show that if $\mathcal{S}(a)$ is a union of cosets of K and of K', then $\mathcal{S}(a)$ is also a union of cosets of $K + K'$. To see this, let y be any element of $\mathcal{S}(a)$ and consider $z \in y + K + K'$. Write $z = y + h + h'$, where $h \in K$ and $h' \in K'$. Now $y + h \in \mathcal{S}(a)$ as $\mathcal{S}(a)$ is a union of cosets of K, so $z = (y + h) + h' \in \mathcal{S}(a)$ as $\mathcal{S}(a)$ is a union of cosets of K'.

To prove (i), we argue similarly to Remark 8.10. Either $|\mathcal{S}(a)| = |G|$ (so $K(a) = G$), or there exists $j \geq 0$ such that $a_{j+1} \neq 0$ and $\mathcal{S}(a^j) = \mathcal{S}(a^{j+1})$. In the latter case Lemma 8.9 implies that $\mathcal{S}(a^j)$ is a union of cosets of $\langle a_{j+1} \rangle$, and by Lemma 8.11 the same holds for $\mathcal{S}(a)$.

For (ii), suppose for a contradiction that $K' = K(a + K(a))$ is a non-trivial subgroup of $G/K(a)$. Then $\{x + K(a) \mid x \in \mathcal{S}(a)\} = \bigcup_{t \in T}(t + K')$ for some $T \subseteq G/K(a)$. Let K^* be the set of $g \in G$ such that g is contained in some coset $z + K(a)$ with $z \in K'$. Then K^* is a subgroup of G that strictly contains $K(a)$. Now fix representatives $\hat{t} \in t + K(a)$ for $t \in T$ and note that $\mathcal{S}(a) = \bigcup_{t \in T}(\hat{t} + K^*)$, which contradicts the definition of $K(a)$.

To show (iii), suppose for a contradiction that a has at least $|G/K(a)| - 1$ elements not in $K(a)$. Applying (i) in $G/K(a)$ we have $|K(a + K(a))| > 1$, which contradicts (ii), so we are done. □

8.5. The full result

Theorem 1.10 follows from Theorem 8.1 and the following result.

Lemma 8.16. *Suppose that $k \geq 3$ and H is a k-graph such that every $2k(k-3)$-far full pair for H is soluble. Then every full pair for H is soluble.*

First we require the following calculation.

Lemma 8.17. *Suppose $t \geq 1$ and $\mathbf{x} \in \mathbb{Z}^{t+1}$ with $x_i \geq 2$ for $i \in [t+1]$. Let*

$$f(\mathbf{x}) = \sum_{j\in[t]} x_1 \dots x_j + \sum_{j\in[t+1]} x_j + \sum_{j\in[t]} x_{j+1} \dots x_{t+1} - 4t - 2.$$

Then $f(\mathbf{x}) \leq 2(x_1 \dots x_{t+1} - 2)$.

Proof. First we consider the case $t = 1$. Then $f(\mathbf{x}) = 2x_1 + 2x_2 - 6$, so the required inequality is equivalent to $x_1 + x_2 - 1 \leq x_1x_2$; this holds as $(x_1 - 1)(x_2 - 1) \geq 0$. Now suppose $t > 1$. Note that the inequality holds if $x_i = 2$ for all $i \in [t+1]$, as then $f(\mathbf{x}) = (2^{t+1}-2)+2(t+1)+(2^{t+1}-2)-4t-2 < 2(2^{t+1}-2)$. Also note that the inequality holds if $t = 2$ and $\mathbf{x} = (2, s, 2)$ for some $s > 2$, as then $f(\mathbf{x}) = 5s + 8 - 10 < 2(4s - 2)$. For $r \in [t+1]$ write $\mathbf{x}^{r\to 2}$ for the vector obtained from x by setting x_r equal to 2. We will show that

$$\frac{f(\mathbf{x})}{f(\mathbf{x}^{r\to 2})} \leq \frac{x_r}{2} \leq \frac{x_1 \dots x_{t+1} - 2}{2x_1 \dots x_{r-1}x_{r+1} \dots x_{t+1} - 2},$$

except in the case where $t = r = 2$ and $\mathbf{x} = (2, s, 2)$ for some $s > 2$. This suffices to prove the lemma, since we can replace each x_r by 2 in turn until we reach one of the cases considered above. The second inequality is immediate from $x_r \geq 2$. For the first, we rewrite it as

$$\begin{aligned} 0 &\leq x_r f\big(\mathbf{x}^{r\to 2}\big) - 2f(\mathbf{x}) \\ &= (x_r - 2)\Bigg(\sum_{j\in[r-1]} x_1 \dots x_j + \sum_{j\in[t+1], j\neq r} x_j + \sum_{r\leq j\leq t} x_{j+1} \dots x_{t+1} - 4t - 2 \Bigg). \end{aligned}$$

If $t \geq 3$, or if $t = 2$ and $r \neq 2$, this holds since there are $2t$ positive terms in the last bracket, each of which is at least 2 and at least one of which is at least 4. Alternatively, if $t = r = 2$ and either $x_1 \geq 3$ or $x_3 \geq 3$ we note that at least two terms in the bracket are at least 3. In any case we have the desired inequality. □

Proof of Lemma 8.16. We proceed by induction on $|\mathcal{P}|$. Let $(\mathcal{P}, L)$ be a full pair and assume that every full pair $(\mathcal{P}', L')$ with $|\mathcal{P}'| < |\mathcal{P}|$ is soluble. We will show that $(\mathcal{P}, L)$

is soluble. Write $G = G(\mathcal{P}, L)$. We will construct a sequence $G = K_0 \supseteq K_1 \supseteq K_2 \supseteq \ldots \supseteq K_{t_0} \supseteq 0$ of distinct subgroups of G, together with a sequence of vertex sets $\emptyset = S_0 \subseteq S_1 \subseteq S_2 \subseteq \ldots \subseteq S_{t_0}$, such that the following conditions hold:

(i) $|S_t|/k \leq \sum_{j\in[t]} |G/K_j| + \sum_{j\in[t]} |K_{j-1}/K_j| + \sum_{j\in[t]} |K_j| - 4t$.
(ii) Every edge $e \in H$ such that $R_G(e) \notin K_t$ intersects S_t.
(iii) There exists a solution for $(\mathcal{P}_{K_t}, L_{K_t})$ using only vertices from S_t.

Recalling that $\mathcal{P}_G$ is the trivial partition of $V(H)$ into a single part, it is clear that $S_0 = \emptyset$ satisfies these conditions. Given K_t and S_t, we construct K_{t+1} and S_{t+1} as follows.

For any matching M, let $K(M)$ be the key subgroup of the sequence of residues of edges in M (in any order). Consider the set of all matchings M in $H - S_t$ such that $\mathcal{S}_G(M \cup \{e\}) = \mathcal{S}_G(M)$ for every edge $e \in H - S_t$ which does not intersect $V(M)$. This is a non-empty set, as it contains all maximal matchings in $H - S_t$. We choose such an M so that $\mathcal{S}_G(M)$ is maximal, and subject to this M is minimal. Having chosen M, we set $K_{t+1} := K(M)$. Note that the choice of M implies that $\mathcal{S}_G(M') \neq \mathcal{S}_G(M)$ for any strict subset $M' \subset M$. Indeed, if $\mathcal{S}_G(M') = \mathcal{S}_G(M)$ then minimality of M implies that some edge $e \in H - S_t$ which does not intersect $V(M')$ has $\mathcal{S}_G(M' \cup \{e\}) \neq \mathcal{S}_G(M') = \mathcal{S}_G(M)$, but then a maximal matching containing $M' \cup \{e\}$ contradicts the maximality of $\mathcal{S}_G(M)$. Also observe that any edge e with $R_G(e) \notin K_{t+1}$ intersects $S_t \cup V(M)$; otherwise $\mathcal{S}_G(M)$ is a union of cosets of $\langle R_G(e)\rangle$ by Lemma 8.9, which contradicts the definition of $K(M) = K_{t+1}$ as being the unique maximal subgroup of G such that $\mathcal{S}_G(M)$ is a union of cosets of $K(M)$. Thus $(\mathcal{P}_{K_{t+1}}, L_{K_{t+1}})$ is a $|S_t \cup V(M)|$-far full pair for H by Lemma 8.6.

First we consider the case that $K_{t+1} = K_t$. By (iii) there is a solution M_{sol} for $(\mathcal{P}_{K_t}, L_{K_t})$ using only vertices from S_t. Then $R_G(V(H) \setminus V(M_{sol})) \in K_t$ by Lemma 8.6. Since $K(M) = K_{t+1} = K_t$ there is $M' \subseteq M$ with $R_G(V(M')) = R_G(V(H) \setminus V(M_{sol}))$. Now $M' \cup M_{sol}$ is a matching, as M is disjoint from S_t, and $R_G(V(H) \setminus V(M' \cup M_{sol})) = 0_G$, so $(\mathcal{P}, L)$ is soluble by Lemma 6.9.

Now we can assume that K_{t+1} is a strict subgroup of K_t. We will show that $|S_t \cup V(M)| \leq 2k(k-3)$, so that we can apply the hypothesis of the lemma. Define $\mathbf{x} \in \mathbb{Z}^{t+1}$ by $x_j = |K_{j-1}/K_j|$ for $j \in [t]$ and $x_{t+1} = |K_t|$. Then $x_j \geq 2$ for all $j \in [t+1]$ and $x_1 \ldots x_{t+1} = |G| \leq k-1$ by Lemma 6.4. Note that M has at most $|K_{t+1}| - 1$ elements e with $R_G(e) \in K_{t+1}$ by Lemma 8.13. Also, since $R_G(e) \in K_t$ for all $e \in M$ by (ii), M has at most $|K_t/K_{t+1}| - 2$ elements e with $R_G(e) \notin K_{t+1}$ by Lemma 8.15(iii). Therefore $|M| \leq |K_{t+1}| + |K_t/K_{t+1}| - 3$.

By (i) and Lemma 8.17 we have

$$\begin{aligned}
\big|S_t \cup V(M)\big|/k &\leq \sum_{j\in[t]} |G/K_j| + \sum_{j\in[t+1]} |K_{j-1}/K_j| + \sum_{j\in[t+1]} |K_j| - 4t - 3 \\
&= f(\mathbf{x}) - \big(|K_t/K_{t+1}| - 1\big)\big(|K_{t+1}| - 1\big) \\
&\leq 2\big(|G| - 2\big) \leq 2(k-3).
\end{aligned}$$

Thus $(\mathcal{P}_{K_{t+1}}, L_{K_{t+1}})$ is a $2k(k-3)$-far full pair for H, so by assumption has a solution M_{sol}, which by definition has size at most $|G/K_{t+1}| - 1$.

If $|K_{t+1}| = 1$ then M_{sol} is a solution for $(\mathcal{P}, L)$. Otherwise, we define $S_{t+1} = S_t \cup V(M) \cup V(M_{sol})$ and proceed to the next step. The required conditions on S_{t+1} hold as

$$|S_{t+1}|/k \le |S_t|/k + |M| + |M_{sol}| \le |S_t|/k + |G/K_{t+1}| + |K_t/K_{t+1}| + |K_{t+1}| - 4,$$

every edge with $R_G(e) \notin K_{t+1}$ intersects $S_t \cup V(M)$, and there is a solution for $(\mathcal{P}_{K_{t+1}}, L_{K_{t+1}})$ using only vertices from S_{t+1} (namely M_{sol}). □

9. Deferred proofs

In this section we present the proofs of Lemmas 2.4, 5.5 and 7.1.

9.1. Proof of Lemma 2.4

Recall that we are given a k-graph H on n vertices that has a perfect matching, and for each set A of $k-1$ vertices of H we let $t_A = \max(0, (1/k+\gamma)n - d_H(A))$. Write $V = V(H)$,

$$\chi_1 = \sum_{A \in \binom{V}{k-1}} t_A^2 \quad \text{and} \quad \chi_2 = \max\left(0, \frac{n^{k-1}}{3k!} - \delta_1(H)\right).$$

We are given

(i) $\chi_1 < \varepsilon\gamma^2 n^{k+1}/4 + 3kn^k$, and
(ii) $\chi_2 + \frac{\chi_1}{\sqrt{\varepsilon}\gamma^2 n^2} < \sqrt{\varepsilon} n^{k-1}$.

(We have slightly rephrased (ii) from the hypothesis of the lemma by including the maximum in the definition of χ_2; to see that it is equivalent, note that if $\chi_2 = 0$ then by (i) we have $\frac{\chi_1}{\sqrt{\varepsilon}\gamma^2 n^2} < \sqrt{\varepsilon} n^{k-1}$.)

Let e be any edge of H. By (i) there are at most $2\varepsilon n^{k-1}$ $(k-1)$-sets $A \subseteq V \setminus e$ such that $d(A) < (1/k + \gamma/2)n$, and by (ii) we have $\chi_2 < \sqrt{\varepsilon} n^{k-1}$, so $\delta_1(H-e) \ge n^{k-1}/6k!$. Thus we can apply Theorem 2.1 to $H - e$ for every $e \in H$, with $\gamma/2$ and 3ε in place of γ and ε, to construct the set E' of edges e such that $H - e$ contains a perfect matching; in total this takes time $O(n^{3k^2-7k})$.

We will show that we can find an edge $e \in E'$ in time $O(n^{2k})$ such that $H - e$ satisfies the analogues of (i) and (ii). More precisely, for any $(k-1)$-set $A \subseteq V$ we define t_A^e by

$$t_A^e = \begin{cases} \max(0, (1/k+\gamma)(n-k) - d_{H-e}(A)) & \text{if } A \subseteq V \setminus e, \\ 0 & \text{otherwise.} \end{cases}$$

We also define

$$\chi_1^e = \sum_{A\in\binom{V\setminus e}{k-1}} \left(t_A^e\right)^2 \quad \text{and} \quad \chi_2^e = \max\left(0, \frac{(n-k)^{k-1}}{3k!} - \delta_1(H-e)\right).$$

Then we need to find $e \in E'$ so that

(i)^e $\chi_1^e < \varepsilon\gamma^2(n-k)^{k+1}/4 + 3k(n-k)^k$, and
(ii)^e $\chi_2^e + \frac{\chi_1^e}{\sqrt{\varepsilon}\gamma^2(n-k)^2} < \sqrt{\varepsilon}(n-k)^{k-1}$.

First we claim that for each $A \in \binom{V}{k-1}$,

$$\left(t_A^e\right)^2 \le t_A^2 + 2t_A\left(\left|H(A)\cap e\right| - 1 - k\gamma\right) + k^2. \tag{7}$$

Note that if $t_A = 0$ then (7) follows from $t_A^e \le k$. On the other hand, if $t_A > 0$, then

$$\begin{aligned}\left(t_A^e\right)^2 &\le \left(t_A + \left|H(A)\cap e\right| - 1 - k\gamma\right)^2\\ &= t_A^2 + 2t_A\left(\left|H(A)\cap e\right| - 1 - k\gamma\right) + \left(\left|H(A)\cap e\right| - 1 - k\gamma\right)^2\\ &\le t_A^2 + 2t_A\left(\left|H(A)\cap e\right| - 1 - k\gamma\right) + k^2,\end{aligned}$$

so (7) holds. Now to find the desired edge e, we split into two cases.

Case 1: $\chi_2 > 0$. In this case, by definition there must exist a vertex x of degree at most $n^{k-1}/3k!$, and we can find such a vertex in time $O(n^k)$. Since H contains a perfect matching, there is some $e \in E'$ containing x. We now delete e and show that conditions (i)^e and (ii)^e hold. First we define $\mathcal{A} = \{A \in \binom{V}{k-1} : x \in A\}$ and note that $\sum_{A\in\mathcal{A}} d_H(A) = (k-1)d_H(x) \le n^{k-1}/3(k-1)!$. Since $|\mathcal{A}| = \binom{n-1}{k-2}$ we have

$$\sum_{A\in\mathcal{A}} t_A \ge \left(\frac{1}{k} + \gamma\right)n\binom{n-1}{k-2} - \frac{n^{k-1}}{3(k-1)!} \ge \frac{n}{2k}\binom{n-1}{k-2}.$$

Now by the Cauchy–Schwartz inequality, we obtain $\sum_{A\in\mathcal{A}} t_A^2 \ge \frac{n^2}{4k^2}\binom{n-1}{k-2}$. Moreover, by (7) and then Cauchy–Schwartz we have

$$\begin{aligned}\sum_{A\in\binom{V\setminus e}{k-1}} \left(\left(t_A^e\right)^2 - t_A^2\right) &\le \sum_{A\in\binom{V\setminus e}{k-1}} \left(2t_A(k-1) + k^2\right)\\ &\le 2(k-1)\sqrt{\binom{n}{k-1}\chi_1} + k^2\binom{n-k}{k-1}\\ &\le \sqrt{\varepsilon}n^k,\end{aligned}$$

where the final inequality holds by (i). Since $\mathcal{A} \subseteq \binom{V}{k-1} \setminus \binom{V\setminus e}{k-1}$, we have

$$\chi_1^e \le \chi_1 - \frac{n^2}{4k^2}\binom{n-1}{k-2} + \sqrt{\varepsilon}n^k \le \chi_1 - \frac{n^k}{7k!}, \tag{8}$$

so $\chi_1^e < \varepsilon\gamma^2 n^{k+1}/4 + 3kn^k - n^k/7k! \le \varepsilon\gamma^2(n-k)^{k+1}/4 + 3k(n-k)^k$, which proves (i)^e.

Note also that $\chi_2^e \le \chi_2 + k\binom{n}{k-2}$, so by (8) we have

$$\begin{aligned}
&\left(\frac{\chi_1^e}{\sqrt{\varepsilon}\gamma^2(n-k)^2} + \chi_2^e\right) - \left(\frac{\chi_1}{\sqrt{\varepsilon}\gamma^2 n^2} + \chi_2\right)\\
&\le \frac{\chi_1}{\sqrt{\varepsilon}\gamma^2 n^2}\cdot\frac{2kn-k^2}{(n-k)^2} - \frac{n^k}{7k!\sqrt{\varepsilon}\gamma^2(n-k)^2} + k\binom{n}{k-2}\\
&\le k\sqrt{\varepsilon}n^{k-2} - \frac{n^{k-2}}{\gamma} + k\binom{n}{k-2} < -n^{k-2}.
\end{aligned}$$

Now by (ii) we deduce that

$$\frac{\chi_1^e}{\sqrt{\varepsilon}\gamma^2(n-k)^2} + \chi_2^e < \sqrt{\varepsilon}n^{k-1} - n^{k-2} < \sqrt{\varepsilon}(n-k)^{k-1},$$

which proves (ii)^e.

Case 2: $\chi_2 = 0$. We claim that we can find in time $O(n^{2k})$ an edge $e \in E'$ such that

$$\sum_{A\in\binom{V}{k-1}} \left(2t_A\left(\left|H(A)\cap e\right| - 1 - k\gamma\right) - |A\cap e|t_A^2\right) \le -\frac{k(k+1)\chi_1}{n}. \tag{9}$$

To see this, consider a perfect matching M in H (this exists by hypothesis of the lemma, although the algorithm does not have access to it). Note that

$$\begin{aligned}
\sum_{e\in M}\sum_{A\in\binom{V}{k-1}} 2t_A\left(\left|H(A)\cap e\right| - 1 - k\gamma\right) &= \sum_{A\in\binom{V}{k-1}} 2t_A \sum_{e\in M}\left(\left|H(A)\cap e\right| - 1 - k\gamma\right)\\
&= \sum_{A\in\binom{V}{k-1}} 2t_A\left(\left(\frac{1}{k}+\gamma\right)n - t_A - \frac{(1+k\gamma)n}{k}\right)\\
&= \sum_{A\in\binom{V}{k-1}} -2t_A^2 = -2\chi_1.
\end{aligned}$$

(For the second equality, note that terms with $t_A = 0$ contribute zero to each side of the equation, whilst terms with $t_A > 0$ are equal by definition of t_A and the fact that M has size n/k.) Note also that

$$\sum_{e\in M}\sum_{A\in\binom{V}{k-1}} |A\cap e|t_A^2 = \sum_{A\in\binom{V}{k-1}} (k-1)t_A^2 = (k-1)\chi_1.$$

We deduce that

$$\sum_{e\in M}\sum_{A\in\binom{V}{k-1}}\big(2t_A\big(\big|H(A)\cap e\big|-1-k\gamma\big)-|A\cap e|t_A^2\big)=-(k+1)\chi_1,$$

so by averaging there is some $e\in M$ that satisfies (9). Note further that $e\in E'$, so we can find such an edge e in time $O(n^{2k})$ simply by checking (9) for every edge of E'.

We will show that (i)^e and (ii)^e hold for this choice of e. Let $D(e)$ be the family of $(k-1)$-sets $A\subseteq V$ which intersect e. Applying (7), we obtain

$$\begin{aligned}\chi_1^e-\chi_1&=\sum_{A\in\binom{V}{k-1}}\big(\big(t_A^e\big)^2-t_A^2\big)=\sum_{A\in\binom{V\setminus e}{k-1}}\big(\big(t_A^e\big)^2-t_A^2\big)-\sum_{A\in D(e)}t_A^2\\&\le\sum_{A\in\binom{V\setminus e}{k-1}}2t_A\big(\big|H(A)\cap e\big|-1-k\gamma\big)-\sum_{A\in D(e)}t_A^2+k^2\binom{n}{k-1}.\end{aligned}\tag{10}$$

We will show that the first and second terms of (10) are close to those of (9). For the first, note that $|D(e)|=\binom{n}{k-1}-\binom{n-k}{k-1}\le\frac{k^2}{n}\binom{n}{k-1}$, so

$$\begin{aligned}&\sum_{A\in\binom{V\setminus e}{k-1}}2t_A\big(\big|H(A)\cap e\big|-1-k\gamma\big)-\sum_{A\in\binom{V}{k-1}}2t_A\big(\big|H(A)\cap e\big|-1-k\gamma\big)\\&\le\sum_{A\in D(e)}n\le k^2\binom{n}{k-1}.\end{aligned}\tag{11}$$

For the second, observe that if $A\in\binom{V}{k-1}$ is chosen uniformly at random then $|A\cap e|$ is hypergeometric with mean $k(k-1)/n$, so

$$\sum_{A\in D(e)}|A\cap e|=\sum_{A\in\binom{V}{k-1}}|A\cap e|=\frac{k(k-1)}{n}\binom{n}{k-1}.$$

We also have

$$\big|D(e)\big|=\binom{n}{k-1}-\binom{n-k}{k-1}\ge\left(\frac{k(k-1)}{n}-\frac{k^4}{2n^2}\right)\binom{n}{k-1},$$

so $\sum_{A\in D(e)}(|A\cap e|-1)\le\frac{k^4}{2n^2}\binom{n}{k-1}$ and so $\sum_{A\in D(e)}(|A\cap e|-1)t_A^2\le k^2\binom{n}{k-1}$. It follows that

$$\sum_{A\in D(e)}t_A^2\ge\sum_{A\in\binom{V}{k-1}}|A\cap e|t_A^2-k^2\binom{n}{k-1}.\tag{12}$$

Substituting (11) and (12) into (10) and then applying (9), we obtain

$$\begin{aligned}\chi_1^e &\le \chi_1 - \frac{k(k+1)\chi_1}{n} + k^2\binom{n}{k-1} + k^2\binom{n}{k-1} + k^2\binom{n}{k-1}\\ &\le \left(1-\frac{k(k+1)}{n}\right)\left(\frac{\varepsilon\gamma^2 n^{k+1}}{4} + 3kn^k\right) + 3k^2\binom{n}{k-1}\\ &\le \varepsilon\gamma^2(n-k)^{k+1}/4 + 3k(n-k)^k,\end{aligned}$$

using (i) for the second inequality, and $n^{k+1}-(n-k)^{k+1} \le k(k+1)n^k$ and $n^k-(n-k)^k \le k^2n^{k-1}$ for the third. This proves (i)^e. Further, since $\chi_2 = 0$ we have $\chi_2^e \le k\binom{n-1}{k-2}$, so

$$\frac{\chi_1^e}{\sqrt{\varepsilon}\gamma^2(n-k)^2} + \chi_2^e < \frac{\sqrt{\varepsilon}(n-k)^{k-1}}{2} + k\binom{n-1}{k-2} < \sqrt{\varepsilon}(n-k)^{k-1},$$

which proves (ii)^e. □

9.2. Weak hypergraph regularity

The proofs of Lemmas 5.5 and 7.1 use the weak hypergraph regularity lemma. We state this after the following definitions.

Definition 9.1. Suppose that $\mathcal{P}$ partitions a set V into r parts $V_1,\dots,V_r$ and G is a $\mathcal{P}$-partite k-graph on V. For $A \in \binom{[r]}{k}$ we write G_A for the induced k-partite subgraph of G with parts V_i for $i \in A$.

(i) The *density* of G_A is $d(G_A) = \frac{|G_A|}{\prod_{i\in A}|V_i|}$.
(ii) For $\varepsilon > 0$ and $A \in \binom{[r]}{k}$, we say that the k-partite subgraph G_A is *ε-vertex-regular* if for any sets $V_i' \subseteq V_i$ with $|V_i'| \ge \varepsilon|V_i|$ for $i \in A$, writing $V' = \bigcup_{i\in A} V_i'$, we have $d(G_A[V']) = d(G_A) \pm \varepsilon$.
(iii) We say that a partition $\mathcal{P}$ of $V(G)$ is *ε-regular* if all but at most $\varepsilon|V|^k$ edges of G belong to ε-vertex-regular k-partite subgraphs.
(iv) The *reduced k-graph* $R_{\mathcal{P}}^d$ is the k-graph whose vertices are the parts of $\mathcal{P}$ and whose edges are all k-sets of parts of $\mathcal{P}$ that induce an ε-vertex-regular k-partite subgraph of G of density at least d.

Note that we use the same notation $d(\cdot)$ for density as for degree, but there should be no confusion. We shall use the following formulation of the weak hypergraph regularity lemma to obtain regular partitions.

Theorem 9.2 *(Weak hypergraph regularity lemma). Suppose that $1/n \ll 1/m_0 \ll \varepsilon, 1/r \ll 1/k$. Suppose that G is a k-graph on n vertices and that $\mathcal{P}$ is a partition of $V = V(G)$ into at most r parts of size at least εn. Then there is an ε-regular partition $\mathcal{Q}$ of V into $m \le m_0$ parts such that*

(i) *each part of $\mathcal{Q}$ has size $\lfloor n/m \rfloor$ or $\lceil n/m \rceil$, and*
(ii) *$\mathcal{Q}$ is ε-close to a refinement $\mathcal{P}'$ of $\mathcal{P}$.*

Proof. Taking $1/m_0 \ll \varepsilon' \ll \varepsilon$, the most commonly-used form of the weak hypergraph regularity lemma (see [5]) states that there exist an integer $m \leq m_0$, an exceptional set $V_0 \subseteq V$ of size $|V_0| \leq \varepsilon' n$ and a partition $\mathcal{P}'$ of $V \setminus V_0$ into parts $V_1, \ldots, V_m$ of equal size which is ε'-regular (with respect to $H[V \setminus V_0]$) and has the property that V_j is a subset of some part of $\mathcal{P}$ for any $j \in [m]$. By distributing the vertices of V_0 as equally as possible among the parts of $\mathcal{P}'$ we obtain the desired partition $\mathcal{Q}$. □

9.3. Proof of Lemma 5.5

Before giving the proof of Lemma 5.5, we give a brief sketch of the idea. A rough statement of the lemma is that robust maximality is preserved by random selection, i.e. if some partition $\mathcal{P}$ of $V(H)$ is robustly maximal with respect to H and S is a suitable random subset of $V(H)$ then with high probability $\mathcal{P}[S]$ is also robustly maximal (with slightly weaker parameters) with respect to $H[S]$. It is not hard to see that $H[S]$ has a transferral-free robust edge-lattice with respect to $\mathcal{P}[S]$. The main difficulty is to show that with high probability there is no strict refinement $\mathcal{P}^\circ$ of $\mathcal{P}[S]$ (with large parts) that has a transferral-free robust edge-lattice. Since there are many possible refinements $\mathcal{P}^\circ$ to consider, a straightforward union bound on the probability will not suffice. Instead, we use the weak hypergraph regularity lemma. We show that any such refinement $\mathcal{P}^\circ$ gives rise to a partition of the reduced k-graph, which in turn gives rise a refinement $\mathcal{P}^*$ of $\mathcal{P}$ with a transferral-free robust edge-lattice. However, this contradicts the robust maximality of $\mathcal{P}$, so no such refinement $\mathcal{P}^\circ$ can exist.

Proof of Lemma 5.5. Recall that we are given a k-graph H on n vertices and a partition $\mathcal{P}$ of $V(H)$ that is (c, c', μ, μ')-robustly maximal with respect to H. We are also given a partition $\mathcal{P}'$ of $V(H)$ that refines $\mathcal{P}$ and integers $(n_Z)_{Z \in \mathcal{P}'}$ such that $\eta|Z| \leq n_Z \leq |Z|$. We choose $S \subseteq V(H)$ uniformly at random subject to the condition that $|S \cap Z| = n_Z$ for each $Z \in \mathcal{P}'$. First we note that $|S| \geq \eta n$, and for each part $X \in \mathcal{P}$ that $|S \cap X| \geq \eta|X| \geq \eta c n$. Next we show that $L^{\mu/c}_{\mathcal{P}[S]}(H[S])$ is transferral-free. It suffices to prove that $I^{\mu/c}_{\mathcal{P}[S]}(H[S]) \subseteq I^{\mu}_{\mathcal{P}}(H)$. To see this, note that if $\mathbf{i} \in I^{\mu/c}_{\mathcal{P}[S]}(H[S])$ then there are at least $(\mu/c)(\eta n)^k \geq \mu n^k$ edges $e \in H[S] \subseteq H$ with $\mathbf{i}_{\mathcal{P}[S]}(e) = \mathbf{i}$, so $\mathbf{i} \in I^{\mu}_{\mathcal{P}}(H)$.

It remains to show that with high probability there is no refinement $\mathcal{P}^\circ$ of $\mathcal{P}[S]$ with parts of size at least $(3c'/\eta)|S|$, such that $L^{(\mu')^3}_{\mathcal{P}^\circ}(H[S])$ is transferral-free. Suppose that we have such a partition $\mathcal{P}^\circ$; then we will obtain a contradiction using events that hold with high probability. Introduce new constants ε, N_0, N_1 such that $1/n \ll 1/N_1 \ll 1/N_0 \ll \varepsilon \ll \mu$. We apply Theorem 9.2 to H to obtain an ε-regular partition $\mathcal{Q}$ of $V(H)$ which has n_R parts of almost equal size, for some $N_0 \leq n_R \leq N_1$, and which is ε-close to a refinement of $\mathcal{P}$. Let $n_0 = n/n_R$, so every cluster has size $\lfloor n_0 \rfloor$ or $\lceil n_0 \rceil$. We henceforth

omit the floor and ceiling signs as these do not affect the argument. Let $R = R_{\mathcal{Q}}^{\mu'/3}$ be the $\mu'/3$-reduced k-graph on $\mathcal{Q}$.

For any $Y \in \mathcal{Q}$, note that $|Y \cap S|$ is a sum of independent hypergeometric random variables, with

$$\mathbb{E}\big[|Y \cap S|\big] = \sum_{Z \in \mathcal{P}'} \frac{n_Z |Y \cap Z|}{|Z|} \geq \sum_{Z \in \mathcal{P}'} \eta |Y \cap Z| = \eta n_0.$$

Hence by Corollary 4.11 with high probability $|Y \cap S| \geq \eta n_0/2$ for every $Y \in \mathcal{Q}$. Next we choose a partition of $V(R)$ that is 'representative' of $\mathcal{P}^\circ$.

Claim 9.3. *There exists a partition $\mathcal{S}$ of $V(R)$, whose parts $(X_{\mathcal{S}})_{X \in \mathcal{P}^\circ}$ correspond to those of $\mathcal{P}^\circ$, such that*

(i) *each part $X_{\mathcal{S}} \in \mathcal{S}$ has size at least $3c' n_R/2$, and*
(ii) *$|X \cap Y| \geq (c')^2 n_0$ whenever $Y \in X_{\mathcal{S}}$.*

To prove the claim, we choose $\mathcal{S}$ randomly as follows. Let P be the set of pairs (X, Y) such that $X \in \mathcal{P}^\circ$, $Y \in V(R) = \mathcal{Q}$ and $|X \cap Y| \geq c'|S \cap Y|$. We independently assign each $Y \in V(R)$ to a part $X_{\mathcal{S}}$ such that $(X, Y) \in P$ with probability proportional to $|X \cap Y|$, so with probability

$$\frac{|X \cap Y|}{\sum_{X:(X,Y)\in P} |X \cap Y|} \geq \frac{|X \cap Y|}{|S \cap Y|}.$$

Note that whenever $Y \in X_{\mathcal{S}}$ we have $|X \cap Y| \geq c'(\eta n_0/2) \geq (c')^2 n_0$, so (ii) is satisfied. For (i), consider any $X \in \mathcal{P}^\circ$, and note that $\sum_{Y \in V(R)} |X \cap Y| = |X| \geq (3c'/\eta)|S| \geq 3c'n$. Then

$$\sum_{Y \in V(R)} \frac{|X \cap Y|}{|S \cap Y|} \geq \sum_{Y \in V(R)} \frac{|X \cap Y|}{n_0} \geq \frac{3c'n}{n_0} = 3c' n_R, \quad \text{and}$$

$$\begin{aligned}
\mathbb{E}\big[|X_{\mathcal{S}}|\big] &\geq \sum_{Y:(X,Y)\in P} \frac{|X \cap Y|}{|S \cap Y|} \\
&= \sum_{Y \in V(R)} \frac{|X \cap Y|}{|S \cap Y|} - \sum_{Y:(X,Y)\notin P} \frac{|X \cap Y|}{|S \cap Y|} \\
&\geq 3c' n_R - c' n_R = 2c' n_R.
\end{aligned}$$

Thus (i) holds with high probability by Lemma 4.10, which proves the claim. (Note that here we used the phrase 'with high probability' to mean with probability $1 - e^{-\Omega(n_R^c)}$ for some $c > 0$ as $n_R \to \infty$, that is, with n_R large instead of our usual definition with n large.)

We now define a partition $\mathcal{Q}^*$ of V by

$$\mathcal{Q}^* = \left\{ \bigcup_{Y \in X} Y \;\middle|\; X \in \mathcal{S} \right\}.$$

So $\mathcal{Q}$ refines $\mathcal{Q}^*$, and by Claim 9.3(i) each part of $\mathcal{Q}^*$ has size at least $3c'n_R n_0/2 = 3c'n/2$. Recall that we chose $\mathcal{Q}$ to be ε-close to a refinement of $\mathcal{P}$, so we can obtain a refinement of $\mathcal{P}$ from $\mathcal{Q}$ by changing the part of a set B of at most εn vertices of V. Furthermore, any cluster Y was assigned to $X_{\mathcal{S}}$ for some $X \in \mathcal{P}^\circ$ such that $|X \cap Y| \geq (c')^2 n_0$ by Claim 9.3(ii). Since $\mathcal{P}^\circ$ is a refinement of $\mathcal{P}$, we deduce that we may obtain a refinement of $\mathcal{P}$ from $\mathcal{Q}^*$ by possibly changing the part of

(i) the at most εn vertices in B, and
(ii) the at most $n_0 \cdot (\varepsilon n/(c')^2 n_0) < \mu' n/5$ vertices which lie in clusters Y with $|Y \cap B| \geq (c')^2 n_0$.

We conclude that $\mathcal{Q}^*$ is $(\mu'/4)$-close to a refinement $\mathcal{P}^*$ of $\mathcal{P}$; in particular, $\mathcal{P}^*$ then has parts of size at least $c'n$.

To finish the proof of the lemma, it suffices to prove the following claim. Indeed, since $L^{(\mu')^3}_{\mathcal{P}^\circ}(H[S])$ is transferral-free, it will imply that the same is true of $L^{\mu'}_{\mathcal{P}^*}(H)$, contradicting the robust maximality of $\mathcal{P}$.

Claim 9.4. $I^{\mu'}_{\mathcal{P}^*}(H) \subseteq I^{\mu'/2}_{\mathcal{Q}^*}(H) \subseteq I^{\mu'/10}_{\mathcal{S}}(R) \subseteq I^{(\mu')^3}_{\mathcal{P}^\circ}(H[S])$.

The first inequality holds by Proposition 4.3 since $\mathcal{Q}^*$ is $(\mu'/4)$-close to $\mathcal{P}^*$. For the second inequality, consider $\mathbf{i} \in I^{\mu'/2}_{\mathcal{Q}^*}(H)$. By definition of R, there are at most $\varepsilon n^k + n_R^k \cdot (\mu'/3) n_0^k \leq 2\mu' n^k/5$ edges $e \in H$ with $\mathbf{i}_{\mathcal{P}^*}(e) = \mathbf{i}$ which do not lie in k-graphs corresponding to edges of R. There are at least $\mu' n^k/2$ edges $e \in H$ with $\mathbf{i}_{\mathcal{Q}^*}(e) = \mathbf{i}$, so at least $\mu' n^k/10$ of these lie in k-graphs corresponding to edges of R. Thus there are at least $\mu' n^k/10 n_0^k = \mu' n_R^k/10$ edges $e \in R$ such that $\mathbf{i}_{\mathcal{S}}(e) = \mathbf{i}$, i.e. $\mathbf{i} \in I^{\mu'/10}_{\mathcal{S}}(R)$.

For the final inequality consider $\mathbf{i} \in I^{\mu'/10}_{\mathcal{S}}(R)$. Then there are at least $\mu' n_R^k/10$ edges $e \in R$ with $\mathbf{i}_{\mathcal{S}}(e) = \mathbf{i}$. Consider such an edge $e = \{Y_1, \ldots, Y_k\}$, let $(X_j)_{\mathcal{S}}$ be the part of $\mathcal{S}$ containing Y_j for each $j \in [k]$, and let F be the k-partite subgraph of H with vertex classes $(X_j \cap Y_j)_{j \in [k]}$. By definition of R, since each part of F has size $(c')^2 n_0$ by Claim 9.3(ii), we have $|F| \geq (\mu'/3 - \varepsilon)((c')^2 n_0)^k \geq 10(\mu')^2 (|S|/n_R)^k$. Summing over all choices of e, there are at least $(\mu')^3 |S|^k$ edges $e' \in H[S]$ with $\mathbf{i}_{\mathcal{P}^\circ}(e') = \mathbf{i}$, i.e. $\mathbf{i} \in I^{(\mu')^3}_{\mathcal{P}^\circ}(H[S])$. This completes the proof of the claim, and so of the lemma. □

9.4. Proof of Lemma 7.1

Recall that we are given a k-graph H on a set $V = V(H)$ of size kn and a partition $\mathcal{P}$ of $V(H)$ such that (i) at most εn^{k-1} $(k-1)$-sets $S \subseteq V$ have $d_H(S) < (1+\gamma)n$, (ii) $\mathcal{P}$

is (c, c, μ, μ')-robustly maximal with respect to H, (iii) any vertex is in at least dn^{k-1} edges e with $\mathbf{i}_{\mathcal{P}}(e) \in L^{\mu}_{\mathcal{P}}(H)$, and (iv) $\mathbf{i}_{\mathcal{P}}(V) \in L^{\mu}_{\mathcal{P}}(H)$.

We choose uniformly at random a partition $\mathcal{P}' = (V_1, \dots, V_k)$ of V into k parts each of size n. Let H' be the induced $\mathcal{P}'$-partite subgraph of H and let $\hat{\mathcal{P}}$ be the common refinement of $\mathcal{P}$ and $\mathcal{P}'$. We also set $c_* = c/2k$, $d_* = d/2k^k$ and $\gamma_* = \gamma/2k$, and introduce a new constant μ_* with $\mu \ll \mu_* \ll \mu'$. Note that $\mu \ll \mu_* \ll c_*, d_* \ll \gamma_*, 1/k$. We will show that the following conditions hold with high probability.

(F1) At most εn^{k-1} $\mathcal{P}'$-partite $(k-1)$-sets $S \subseteq V$ have $|H'(S)| < (1/k + \gamma_*)n$,
(F2) $\hat{\mathcal{P}}$ is (c_*, c_*, μ, μ_*)-robustly maximal with respect to H',
(F3) any vertex is in at least $d_* n^{k-1}$ edges $e \in H'$ with $\mathbf{i}_{\hat{\mathcal{P}}}(e) \in L^{\mu}_{\hat{\mathcal{P}}}(H')$, and
(F4) $\mathbf{i}_{\hat{\mathcal{P}}}(V) \in L^{\mu}_{\hat{\mathcal{P}}}(H')$.

We then apply Lemma 7.6 to H' with $\hat{\mathcal{P}}$ in place of $\mathcal{P}$ and $c_*, \mu_*, d_*, \gamma_*$ in place of c, μ', d, γ to obtain a perfect matching in H', which is also a perfect matching in H.

First we note that (F1) is immediate from (i) and Corollary 4.11, and (F3) follows from (iii) and Corollary 4.13, similarly to the proof of (E3) in Claim 7.11. Next, since each part of $\mathcal{P}$ has size at least $c|V| = ckn$, by Corollary 4.11 with high probability each part of $\hat{\mathcal{P}}$ has size at least $cn/2 = c_*|V|$ and by Proposition 4.5(i), for every $\mathbf{i} \in L^{\mu}_{\hat{\mathcal{P}}}(H') \subseteq L^{\mu}_{\hat{\mathcal{P}}}(H)$ we have $(\mathbf{i} \mid \mathcal{P}) \in L^{\mu}_{\mathcal{P}}(H)$, so $L^{\mu}_{\hat{\mathcal{P}}}(H')$ is transferral-free. This gives part of (F2); it also allows us to apply Proposition 7.5 and deduce that $L^{\mu}_{\hat{\mathcal{P}}}(H')$ is full with respect to $\mathcal{P}'$.

Now we claim that $\mathbf{i} \in L^{\mu}_{\hat{\mathcal{P}}}(H')$ for every index vector $\mathbf{i}$ with respect to $\hat{\mathcal{P}}$ such that $(\mathbf{i} \mid \mathcal{P}')$ is a multiple of $\mathbf{1}$ and $(\mathbf{i} \mid \mathcal{P}) \in L^{\mu}_{\mathcal{P}}(H)$. Note that this will imply (F4), as $(\mathbf{i}_{\hat{\mathcal{P}}}(V) \mid \mathcal{P}') = n \cdot \mathbf{1}$ and $(\mathbf{i}_{\hat{\mathcal{P}}}(V) \mid \mathcal{P}) = \mathbf{i}_{\mathcal{P}}(V) \in L^{\mu}_{\mathcal{P}}(H)$ by (iv). To prove the claim, we apply Proposition 6.12(iii) to find parts X and X' of $\hat{\mathcal{P}}$ such that $\mathbf{i} - \mathbf{u}_X + \mathbf{u}_{X'} \in L^{\mu}_{\hat{\mathcal{P}}}(H')$. Now $(\mathbf{u}_X - \mathbf{u}_{X'} \mid \mathcal{P}) = (\mathbf{i} \mid \mathcal{P}) - (\mathbf{i} - \mathbf{u}_X + \mathbf{u}_{X'} \mid \mathcal{P}) \in L^{\mu}_{\mathcal{P}}(H)$, so X and X' must be contained in the same part of $\mathcal{P}$, as $L^{\mu}_{\mathcal{P}}(H)$ is transferral-free. Also, since H' is $\mathcal{P}'$-partite, X' is contained in the same part of $\mathcal{P}'$ as X, so $X' = X$ and $\mathbf{i} \in L^{\mu}_{\hat{\mathcal{P}}}(H')$, as claimed.

To finish the proof of (F2), we need to show that with high probability there is no strict refinement $\hat{\mathcal{P}}^*$ of $\hat{\mathcal{P}}$ with parts of size at least $c_*|V|$ such that $L^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$ is transferral-free; this will occupy the remainder of the proof. Suppose that we have such a partition $\hat{\mathcal{P}}^*$; then we will obtain a contradiction using events that hold with high probability. By (F1) we can apply Proposition 7.5 to deduce that $L^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$ is full and every part of $\hat{\mathcal{P}}^*$ has size at least $(1/k + \gamma_*/2)n$. Also, by Proposition 6.12(i) there is some integer ℓ such that every part of $\mathcal{P}'$ is refined into ℓ parts by $\hat{\mathcal{P}}^*$; note that $\ell < k$.

Now we use an argument which is similar in spirit to that of Lemma 5.5 (but more complicated). Let N_0, N_1 satisfy $1/n \ll 1/N_1 \ll 1/N_0 \ll \varepsilon$. We apply Theorem 9.2 to obtain an ε-regular partition $\mathcal{Q}$ of V which is ε-close to a refinement of $\mathcal{P}$ and has n_R parts of almost equal size, for some $N_0 \leq n_R \leq N_1$. Let $n_0 = kn/n_R$, so every cluster has size $\lceil n_0 \rceil$ or $\lfloor n_0 \rfloor$; we omit the floor and ceiling signs since they do not affect the

argument. Also let $R = R_{\mathcal{Q}}^{\mu'/4}$ be the $\mu'/4$-reduced k-graph of H on $\mathcal{Q}$. For any $e \in R$ we write

$$H_e = \{e' \in H : e' \cap X \neq \emptyset \ \forall X \in e\}$$

for the k-partite subgraph of H corresponding to e. We note for future reference the following claim, which shows that R inherits a codegree condition from H.

Claim 9.5. *At most $\gamma^{-1}\varepsilon n_R^{k-1}$ $(k-1)$-sets $A \subseteq V(R)$ have $|R(A)| < n_R/k$.*

To see this, recall that at most $\varepsilon|V|^k$ edges of H do not lie in ε-vertex-regular k-partite subgraphs of H formed by parts of $\mathcal{Q}$; call these *irregular* edges. So there are at most $k^2\sqrt{\varepsilon}n_R^{k-1}$ $(k-1)$-sets $A \subseteq V(R)$ for which more than $\sqrt{\varepsilon}n_0^{k-1}n$ irregular edges intersect each member of A. Fix any other A for which $|R(A)| < n_R/k$. Then we have $\sum_{B\in V_A}|H(B)| < n_0^k n_R/k + (\mu'/4)n_0^{k-1}n + \sqrt{\varepsilon}n_0^{k-1}n$, where V_A denotes the set of $(k-1)$-sets with one vertex from each cluster of A. Let m_A be the number of $(k-1)$-sets $B \in V_A$ such that $|H(B)| < (1+\gamma)n$. Then $\sum_{B\in V_A}|H(B)| \geq (n_0^{k-1} - m_A)(1+\gamma)n$, and so

$$\begin{aligned} m_A &\geq n_0^{k-1} - \frac{n_0^k n_R/k + (\mu'/4)n_0^{k-1}n + \sqrt{\varepsilon}n_0^{k-1}n}{(1+\gamma)n} \\ &= n_0^{k-1}\frac{\gamma - \mu'/4 - \sqrt{\varepsilon}}{1+\gamma} \geq \frac{1}{2}\gamma n_0^{k-1}. \end{aligned}$$

Since by (i) there are at most εn^{k-1} $(k-1)$-sets $B \subseteq V$ with $|H(B)| < (1+\gamma)n$, there can be at most $\varepsilon n_R^{k-1}/2\gamma$ such sets A. Together with the at most $k^2\sqrt{\varepsilon}n_R^{k-1}$ sets A considered earlier we have a total of at most $\gamma^{-1}\varepsilon n_R^{k-1}$ $(k-1)$-sets $A \subseteq V(R)$ with $|R(A)| < n_R/k$, proving the claim.

In the next claim we choose for each cluster of R a representative part of $\hat{\mathcal{P}}^*$ within each part of $\mathcal{P}'$. We encode these choices by k-vectors with respect to $\hat{\mathcal{P}}^*$ which are $\mathcal{P}'$-partite, by which we mean that they correspond to $\mathcal{P}'$-partite k-sets.

Claim 9.6. *There are $\mathcal{P}'$-partite k-vectors $\mathbf{i}^*(Y) \in \mathbb{Z}^{\hat{\mathcal{P}}^*}$ for each $Y \in V(R)$ such that*

(G1) $|Y \cap Z| \geq c_* n_0$ *whenever* $i^*(Y)_Z = 1$, *and*
(G2) $R_Z := |\{Y : i^*(Y)_Z = 1\}| \geq n_R/k$ *for every* $Z \in \hat{\mathcal{P}}^*$.

The proof of the claim is similar to that of Claim 9.3. For each $i \in [k]$ we let P_i be the set of pairs (Y, Z) such that $Y \in V(R) = \mathcal{Q}$, $Z \in \hat{\mathcal{P}}^*$, $Z \subseteq V_i$ and $|Y \cap Z| \geq c_* n_0$. For each $Y \in V(R)$ and $i \in [k]$ we let $i^*(Y)_Z = 1$ for a part Z with $(Y, Z) \in P_i$ chosen with probability proportional to $|Y \cap Z|$, and let $i^*(Y)_Z = 0$ otherwise, where all random choices are made independently. Thus (G1) is satisfied by choice of P_i. For (G2), first note

that $\mathbb{P}[i^*(Y)_Z = 1] \geq \frac{|Y\cap Z|}{|Y\cap V_i|}$ for every Y and $Z \subseteq V_i$ such that $|Y \cap Z| \geq c_* n_0$. Also, since $|Y \cap V_i|$ is distributed hypergeometrically, with high probability $|Y \cap V_i| = (1 \pm \varepsilon)n_0/k$ for every $Y \in \mathcal{Q}$ and $V_i \in \mathcal{P}'$ by Corollary 4.11. Now for any $V_i \in \mathcal{P}'$ and any $Z \in \hat{\mathcal{P}}^*$ with $Z \subseteq V_i$ we have

$$\sum_{Y\in V(R)} \frac{|Y \cap Z|}{|Y \cap V_i|} \geq \sum_{Y\in V(R)} \frac{|Y \cap Z|}{(1+\varepsilon)n_0/k} = \frac{|Z|}{(1+\varepsilon)n_0/k}, \quad \text{and}$$

$$\sum_{Y:(Y,Z)\notin P_i} \frac{|Y \cap Z|}{|Y \cap V_i|} \leq n_R \cdot \frac{c_* n_0}{(1-\varepsilon)n_0/k} < 2kc_* n_R.$$

Since every part Z of $\hat{\mathcal{P}}^*$ has size at least $(1/k + \gamma_*/2)n$, we obtain

$$\mathbb{E}[R_Z] \geq \sum_{Y:(Y,Z)\in P_i} \frac{|Y \cap Z|}{|Y \cap V_i|} \geq (1 + \gamma_*/3)n_R/k.$$

Thus (G2) holds with high probability by Lemma 4.10, which completes the proof of the claim.

In the next claim we show that certain 'bad' occurrences are rare. We will consider a cluster to be bad if it is not well-represented by its choice of vector in Claim 9.6. We will consider a vertex to be bad if it belongs to a good cluster but not to the part assigned to this cluster in Claim 9.6. We will consider an edge of the reduced k-graph to be bad if it either contains a bad cluster or contains two clusters whose vectors are 'incompatible'. More precisely, we make the following definitions.

Definition 9.7.

(i) We call a cluster $Y \in V(R)$ *bad* if there exists $Z \in \hat{\mathcal{P}}^*$ such that $\mathbf{i}^*(Y)_Z = 0$, but $|Y \cap Z| \geq \mu_*^{1/4} n_0$; otherwise we call Y *good*.
(ii) We call a vertex *bad* if it is contained in $Y \cap Z$ for some good cluster Y and some $Z \in \hat{\mathcal{P}}^*$ such that $\mathbf{i}^*(Y)_Z = 0$.
(iii) We call an edge e of R *bad* if either e contains a bad cluster, or e contains two clusters Y_1, Y_2 such that $\mathbf{i}^*(Y_1)$ and $\mathbf{i}^*(Y_2)$ are neither identical nor orthogonal to each other.

Claim 9.8. *H has at most $k^3\mu_*^{1/4} n$ bad vertices and R has at most $\sqrt{\mu_*} n_R^k$ bad edges.*

The first part of the claim is immediate from the definitions, using $|\hat{\mathcal{P}}^*| \leq k^2$ and $n_0 n_R = kn$. For the second, we will show that if e is bad then the number of edges $e' \in H_e$ with $\mathbf{i}_{\hat{\mathcal{P}}^*}(e') \notin L^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$ is at least $\mu_*^{1/3} n_0^k$. The claim then follows from the fact that H' has at most $\ell^k \mu_* (kn)^k$ such edges in total.

First we consider $e \in R$ that is bad because it contains a bad cluster Y_1. Then there is $Z \in \hat{\mathcal{P}}^*$ such that $\mathbf{i}^*(Y_1)_Z = 0$ but $|Y_1 \cap Z| \geq \mu_*^{1/4} n_0$. Without loss of generality $Z \subseteq V_1$.

Let $Y_2,\dots,Y_k$ be the remaining clusters of e, and let Z_j be the parts of $\hat{\mathcal{P}}^*$ such that $Z_j \subseteq V_j$ and $\mathbf{i}^*(Y_j)_{Z_j} = 1$ for each $j \in [k]$. Now we consider two induced subgraphs of H_e: let F_1 have parts $(Y_j \cap Z_j)_{j\in[k]}$ and F_2 have parts $Y_1 \cap Z$ and $(Y_j \cap Z_j)_{2\le j\le k}$. Let $\mathbf{i}_1$ and $\mathbf{i}_2$ be the index vectors with respect to $\hat{\mathcal{P}}^*$ of edges in F_1 and F_2 respectively. Note that $\mathbf{i}_1$ and $\mathbf{i}_2$ are $\mathcal{P}'$-partite k-vectors, so the edges of F_1 and F_2 are $\mathcal{P}'$-partite. Also observe that $\mathbf{i}_1 - \mathbf{i}_2 = \mathbf{u}_{Z_1} - \mathbf{u}_Z$, and hence at least one of $\mathbf{i}_1$ and $\mathbf{i}_2$ does not lie in $L^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$, since Z and Z_1 are distinct and $L^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$ is transferral-free by our choice of $\hat{\mathcal{P}}$. Thus it suffices to show that each of F_1 and F_2 contains at least $\mu_*^{1/3}n_0^k$ edges. To see this, note that by (G1) and the choice of Z, each of F_1 and F_2 contains at least c_*n_0 vertices from $k-1$ vertex classes and at least $\mu_*^{1/4}n_0$ vertices from the remaining vertex class. Furthermore H_e is ε-regular with density at least $\mu'/4$ by definition of the reduced k-graph R. Thus each of F_1 and F_2 contains at least $(\mu'/4-\varepsilon)(c_*)^{k-1}\mu_*^{1/4}n_0^k > \mu_*^{1/3}n_0^k$ edges, as required.

The other case is that $e \in R$ is bad because it contains two clusters Y_1, Y_2 such that $\mathbf{i}^*(Y_1), \mathbf{i}^*(Y_2)$ are neither identical nor orthogonal to each other. Then without loss of generality there exist $Z_1, Z_1', Z_2 \in \hat{\mathcal{P}}^*$ with $Z_1, Z_1' \subseteq V_1$ and $Z_2 \subseteq V_2$, such that $\mathbf{i}^*(Y_1)_{Z_1} = \mathbf{i}^*(Y_2)_{Z_1'} = \mathbf{i}^*(Y_1)_{Z_2} = \mathbf{i}^*(Y_2)_{Z_2} = 1$. Let $Y_3,\dots,Y_k$ be the remaining clusters of e and select $Z_j \in \hat{\mathcal{P}}^*$ such that $Z_j \subseteq V_j$ and $\mathbf{i}^*(Y_j)_{Z_j} = 1$ for each $3 \le j \le k$. Now we consider two induced subgraphs of H_e: let F_1 have parts $(Y_j \cap Z_j)_{j\in[k]}$ and F_2 have parts $Y_2 \cap Z_1', Y_1 \cap Z_2$ and $(Y_j \cap Z_j)_{3\le j\le k}$. Similarly to the previous case, each of F_1 and F_2 contains at least $\mu_*^{1/3}n_0^k$ edges, each of which is $\hat{\mathcal{P}}$-partite, and for one of them the index vector is not in $L^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$. This completes the proof of the claim.

In the next claim we use the index vectors from Claim 9.6 that are well-represented to define a partition of V. More precisely, we let

$$R_{\mathbf{i}} = \big\{Y \in V(R) \;\big|\; \mathbf{i}^*(Y) = \mathbf{i}\big\} \quad \text{and} \quad T = \big\{\mathbf{i} \in \mathbb{Z}^{\hat{\mathcal{P}}^*} \;\big|\; |R_{\mathbf{i}}| \ge \mu_*^{1/5}n_R\big\}.$$

Then we define $\mathcal{P}^* = \{X^*_{\mathbf{i}}\}_{\mathbf{i}\in T}$, where $X^*_{\mathbf{i}} = \bigcup_{Z\in\hat{\mathcal{P}}^*, i_Z=1} Z$ for each $\mathbf{i} \in T$.

Claim 9.9. *$\mathcal{P}^*$ is a partition of V which strictly refines $\mathcal{P}$ into parts of size at least $c|V|$.*

To see this, we start by showing that any distinct vectors $\mathbf{i}_1$, $\mathbf{i}_2$ of T are orthogonal. For suppose otherwise, and note that the number of $(k-1)$-sets of clusters that intersect $R_{\mathbf{i}_1}$ and $R_{\mathbf{i}_2}$ is at least $k!^{-1}(n_R)^{k-3}|R_{\mathbf{i}_1}||R_{\mathbf{i}_2}| \ge k!^{-1}\mu_*^{2/5}(n_R)^{k-1}$. By Claim 9.5 all but at most $\gamma^{-1}\varepsilon n_R^{k-1}$ of these $(k-1)$-sets A satisfy $|R(A)| \ge n_R/k$. But then we obtain at least $(k!^{-1}\mu_*^{2/5} - \gamma^{-1}\varepsilon)n_R^{k-1}(n_R/k) > \sqrt{\mu_*}n_R^k$ bad edges, which contradicts Claim 9.8.

Thus for every $Z \in \hat{\mathcal{P}}^*$ there is at most one $\mathbf{i} \in T$ such that $i_Z = 1$, so the parts of $\mathcal{P}^*$ are pairwise disjoint. Furthermore, for each $Z \in \hat{\mathcal{P}}^*$ there must be some $\mathbf{i} \in T$ such that $i_Z = 1$. Indeed, if $i_Z = 0$ for every $\mathbf{i} \in T$ then the number of $Y \in V(R)$ with $i^*(Y)_Z = 1$ is at most $\ell^k\mu_*^{1/5}n_R$, which contradicts (G2). Thus $\mathcal{P}^*$ is a partition of V. Clearly each part of $\mathcal{P}^*$ has size at least $c|V|$, since this is true of $\hat{\mathcal{P}}^*$.

Finally, to see that $\mathcal{P}^*$ refines $\mathcal{P}$, recall that $\mathcal{Q}$ was chosen to be ε-close to a refinement of $\mathcal{P}$. So there are at most $\varepsilon|V|/c_*n_0 < \mu_*^{1/5}n_R$ clusters Y of $\mathcal{Q}$ which intersect more than one part of $\mathcal{P}$ in at least c_*n_0 vertices. Now, consider any $\mathbf{i} \in T$. By definition of T we have $|R_{\mathbf{i}}| \geq \mu_*^{1/5}n_R$, so we may choose $Y \in R_{\mathbf{i}}$ for which there is a unique part $X \in \mathcal{P}$ such that $|X \cap Y| \geq c^*n_0$. We claim that $Z \subseteq X$ for any $Z \in \hat{\mathcal{P}}^*$ with $Z \subseteq X^*_{\mathbf{i}}$. Indeed, since $\hat{\mathcal{P}}^*$ is a refinement of $\mathcal{P}$ we have $Z \subseteq X'$ for some $X' \in \mathcal{P}$. But $i_Z = 1$ by definition of $X^*_{\mathbf{i}}$, so $|Y \cap Z| \geq c_*n_0$ by (G1). By choice of Y this implies that $X = X'$, as claimed. This completes the proof of Claim 9.9.

Our final claim will complete the proof of the lemma. Indeed, it implies that $\mathbf{i} \in L^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$ for any index vector $\mathbf{i}$ with respect to $\hat{\mathcal{P}}^*$ such that $\mathbf{i}' = (\mathbf{i} \mid \mathcal{P}^*) \in L^{\mu'}_{\mathcal{P}^*}(H)$ and $(\mathbf{i} \mid \mathcal{P}')$ is a multiple of $\mathbf{1}$. We can apply this with $\mathbf{i}' \in L^{\mu'}_{\mathcal{P}^*}(H)$ equal to a transferral $\mathbf{u}_{X_1^*} - \mathbf{u}_{X_2^*}$, which must exist since $\mathcal{P}$ is (c, c, μ, μ')-robustly maximal with respect to H and $\mathcal{P}^*$ strictly refines $\mathcal{P}$ into parts of size at least $c|V|$. Letting Z_1, Z_2 be the intersections of X_1^* and X_2^* respectively with V_1 we deduce that $\mathbf{u}_{Z_1} - \mathbf{u}_{Z_2} \in L^{\mu_*}_{\hat{\mathcal{P}}^*}(H)$, contradicting our assumption that $L^{\mu_*}_{\hat{\mathcal{P}}^*}(H)$ was transferral-free. Thus it remains to prove the following claim.

Claim 9.10. *For any $\mathcal{P}'$-partite k-vector $\mathbf{i}$ with respect to $\hat{\mathcal{P}}^*$ such that $(\mathbf{i} \mid \mathcal{P}^*) \in I^{\mu'}_{\mathcal{P}^*}(H)$ we have $\mathbf{i} \in I^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$.*

To see this, let $J_{\mathbf{i}}$ be the set of good edges $e \in R$ such that $e \subseteq \bigcup_{\mathbf{i}' \in T} R_{\mathbf{i}'}$ for which H_e contains some edge e' with no bad vertex and $\mathbf{i}_{\mathcal{P}^*}(e') = (\mathbf{i} \mid \mathcal{P}^*)$. We claim that $|J_{\mathbf{i}}| \geq \mu' n_R^k/2$. To see this, note that since $(\mathbf{i} \mid \mathcal{P}^*) \in I^{\mu'}_{\mathcal{P}^*}(H)$, there are at least $\mu'(kn)^k$ edges $e' \in H$ such that $\mathbf{i}_{\mathcal{P}^*}(e') = (\mathbf{i} \mid \mathcal{P}^*)$. Of these, at most $\varepsilon(kn)^k + \mu'(kn)^k/4$ are not contained in $\bigcup_{e \in R} H_e$, and at most $k^2\mu_*^{1/4}(kn)^k$ contain a bad vertex. Thus there are at least $2\mu' n_R^k/3$ edges of R that contain at least one of the remaining edges e'. Of these edges of R, at most $\sqrt{\mu_*}n_R^k$ are bad, and at most $k^k\mu_*^{1/5}n_R^k$ intersect $\bigcup_{\mathbf{i} \notin T} R_{\mathbf{i}}$. We conclude that $|J_{\mathbf{i}}| \geq \mu' n_R^k/2$, as claimed.

Next we claim that for any $e \in J_{\mathbf{i}}$ there are at least $(\mu')^2 n_0^k$ edges $e^* \in H_e$ with $\mathbf{i}_{\hat{\mathcal{P}}^*}(e^*) = \mathbf{i}$. To see this, fix some $e' \in H_e$ with no bad vertex, and consider any $v \in e'$. Since v is not bad we have $\mathbf{i}^*(Y_v)_{Z_v} = 1$ for the $Y_v \in e$ and $Z_v \in \hat{\mathcal{P}}^*$ with $v \in Y \cap Z$, and since each cluster $Y \in e$ lies in $R_{\mathbf{i}'}$ for some $\mathbf{i}' \in T$, we have $\mathbf{i}^*(Y_v) \in T$. Recall that the vectors in T are orthogonal, so for each $Z \in \hat{\mathcal{P}}^*$ there is a unique $\mathbf{i}^Z \in T$ such that $i^Z_Z = 1$. So we must have $\mathbf{i}^*(Y_v) = \mathbf{i}^{Z_v}$, that is, $Y_v \in R_{\mathbf{i}^{Z_v}}$. Since $\mathbf{i}_{\mathcal{P}^*}(e') = (\mathbf{i} \mid \mathcal{P}^*)$, there must be exactly $(\mathbf{i} \mid \mathcal{P}^*)_{X^*_{\mathbf{i}'}}$ clusters of e in $R_{\mathbf{i}'}$ for each $\mathbf{i}' \in T$. Thus we can order the clusters of e as $(Y_1, \dots, Y_k)$ such that $\mathbf{i}^*(Y_j)_{Z_j} = 1$ for each $j \in [k]$, where $Z_j \subseteq V_j$ is such that $i_{Z_j} = 1$ for $j \in [k]$. Now the sets $Y_j \cap Z_j$ for $j \in [k]$ each have size at least c_*n_0 by (G1), and so by definition of R they induce a subgraph of H_e with at least $(\mu'/4 - \varepsilon)(c_*n_0)^k > (\mu')^2 n_0^k$ edges, as claimed.

Summing over $e \in J_{\mathbf{i}}$, we obtain at least $(\mu')^2 n_0^k \cdot \mu' n_R^k/2 \geq \mu_*(kn)^k$ edges $e^* \in H$ with $\mathbf{i}_{\hat{\mathcal{P}}^*}(e^*) = \mathbf{i}$, and each such edge lies in H' since $\mathbf{i}$ is $\hat{\mathcal{P}}$-partite. Thus $\mathbf{i} \in I^{\mu_*}_{\hat{\mathcal{P}}^*}(H')$, which completes the proof of Claim 9.10, and so of the lemma. □

10. Concluding remarks

The only case where we did not resolve the complexity status of $\mathbf{PM}(k,\delta)$ is when $\delta = 1/k$; this has been solved while this paper was under review by Han [11]: he showed that there is a polynomial time algorithm, using some theory developed in our paper and a lattice-based absorbing method.

We will conclude our paper with some remarks on multipartite versions of our results, tightness of the parameters, and other degree conditions.

10.1. Multipartite analogues

The following multipartite versions of our main results may be proved similarly to the non-partite versions.

Theorem 10.1. *Fix $k \geq 3$ and $\gamma > 0$. Then there is an algorithm with running time $O(n^{3k^2-7k+1})$, which given any k-partite k-graph with parts of size n such that every partite $(k-1)$-set has degree at least $(1/k+\gamma)n$, finds either a perfect matching or a certificate*[1] *that no perfect matching exists.*

Theorem 10.2. *Suppose that $k \geq 3$ and $1/n \ll \varepsilon \ll \gamma \ll 1/k$. Let $\mathcal{P}'$ partition a vertex set V into k parts each of size n. Suppose that H is a $\mathcal{P}'$-partite k-graph H on V such that $\delta_1(H) \geq \gamma n^{k-1}$, and at most εn^{k-1} $\mathcal{P}'$-partite $(k-1)$-sets $A \subseteq V(H)$ have $d_H(A) < (1/k+\gamma)n$. Then H has a perfect matching if and only if there is no $2k(k-3)$-certificate for H.*

Let $\mathbf{PPM}(k,\delta)$ denote the problem of deciding whether there is a perfect matching in a given k-partite k-graph with parts of size n such that every partite $(k-1)$-set has degree at least δn. Theorem 10.1 implies that $\mathbf{PPM}(k,\delta)$ can be decided in polynomial time when $\delta > 1/k$. On the other hand, a similar argument to that of Szymańska [28] shows that $\mathbf{PPM}(k,\delta)$ is NP-complete for $\delta < 1/k$.

10.2. Tightness of parameters

We believe that a stronger version of Lemma 8.16 is true, in which $2k(k-3)$ is replaced by $k(k-3)$, as in the case for full pairs $(\mathcal{P}, L)$ where $|G(\mathcal{P}, L)|$ is prime (see Remark 8.10). This would immediately imply the following improved version of Theorem 1.10.

Conjecture 10.3. *Under Setup 1.7, H has a perfect matching if and only if there is no $k(k-3)$-certificate for H.*

[1] A C-certificate for a k-partite k-graph is defined in an analogous way as for a general k-graph.

A proof of this conjecture would allow us to improve the running time in Theorem 1.1. We also conjecture that a similar strengthening of Theorem 10.2 holds. Conjecture 10.3, if true, would be best-possible. To see this, consider the following construction, which is similar to Construction 1.6.

Construction 10.4. Let $\mathcal{P} = (A_1, \dots, A_{k-1})$ be a partition of a set of n vertices with $|A_j| = n/(k-1) \pm 2$ and $\sum_{j\in[k-1]} j|A_j| = k-2$ modulo $k-1$. Let B be a subset of A_1 of size $k(k-2)-1$. Let H be the k-graph with vertex set $\bigcup_{j\in[k-1]} A_j$ whose edges are

(1) any k-set e with $\sum_{j\in[k-1]} j|e \cap A_j| = 0$ modulo $k-1$, and
(2) any k-set of vertices in B.

Consider the full pair $(\mathcal{P}, L)$ for H, where L is the lattice of index vectors $\mathbf{i}$ such that $\sum_{j\in[k-1]} i_j j = 0$ modulo $k-1$, and note that $(\mathcal{P}, L)$ is a $k(k-3)$-far full pair for H. Since any matching M contains at most $k-3$ edges contained in B, we have $\sum_{j\in[k-1]} j|V(M) \cap A_j| \neq k-2$ for any matching M in H and so H cannot contain a perfect matching. On the other hand, $(\mathcal{P}, L)$ is not a $(k(k-3)-1)$-far full pair for H since removing $k(k-3)-1$ vertices from B still leaves k vertices which form a single edge e with $\mathbf{i}_{\mathcal{P}}(e) \notin L$. It is not hard to show that indeed no $(k(k-3)-1)$-far full pair for H exists.

We also remark that the following construction shows that the constant $1/k$ in Setup 1.7 is best-possible.

Construction 10.5 *(Space barrier).* Let V be a set of size n and fix $S \subseteq V$ with $|S| < n/k$. Let H be the k-graph whose edges are all k-sets that intersect S.

Note that any matching in Construction 10.5 has at most $|S|$ edges, so is not perfect.

10.3. Other degree conditions

It is natural to ask whether similar results could be obtained for the decision problem for perfect matching in k-graphs under the weaker complex degree sequence assumption used in [17]. However, our inductive argument in the key lemma depends crucially on the inheritance of the codegree condition by subgraphs, and this inheritance need not hold for a complex degree sequence condition. Thus an alternative proof strategy would be needed for such a result.

It is also natural to consider the decision problem of determining the existence of a perfect matching in k-graphs satisfying a minimum ℓ-degree condition for any $\ell \in [k-1]$. However, one should note that the ℓ-degree threshold for the existence of a perfect matching is open in general. One can read [12] as suggesting that either a space barrier or a divisibility barrier is always extremal for such a problem, and this was formalised as Conjecture 3.6 in [26]. Let $\mathbf{PM}(k, \ell, \delta)$ denote the problem of deciding whether there is a

perfect matching in a given k-graph on n vertices with minimum ℓ-degree at least $\delta\binom{n}{k-\ell}$. Szymańska [29] showed that $\mathbf{PM}(k,\ell,\delta)$ is NP-complete when δ is less than the space barrier threshold, i.e. $\delta < 1-(1-1/k)^{k-\ell}$. It is then natural to make the following conjecture.

Conjecture 10.6. $\mathbf{PM}(k,\ell,\delta)$ *is in P for* $\delta > 1-(1-1/k)^{k-\ell}$.

Theorem 1.1 establishes the case $\ell = k-1$ of Conjecture 10.6. Beyond this, the conjecture is immediate for those cases where the space barrier threshold is (asymptotically) equal to the threshold at which a perfect matching is guaranteed. The conjecture of Hàn, Person and Schacht mentioned above would therefore imply all cases of Conjecture 10.6 with $(1-1/k)^{k-\ell} \leq 1/2$, and some cases of the latter conjecture are implied by partial results for the former. Specifically, Conjecture 10.6 holds in the case $k=3, \ell=1$ by a result of Hàn, Person and Schacht [12], in the case $k=4, \ell=1$ by a result of Lo and Markström [23], and in the cases $k=5, \ell=1$ and $k=6, \ell=2$ by results of Alon, Frankl, Huang, Rödl, Ruciński and Sudakov [2]. To our knowledge all other cases remain open.

Acknowledgments

We thank the anonymous referees for many helpful comments.

References

[1] R. Aharoni, A. Georgakopoulos, P. Sprüssel, Perfect matchings in r-partite r-graphs, European J. Combin. 30 (2009) 39–42.
[2] N. Alon, P. Frankl, H. Huang, V. Rödl, A. Ruciński, B. Sudakov, Large matchings in uniform hypergraphs and the conjectures of Erdős and Samuels, J. Combin. Theory Ser. A 119 (2012) 1200–1215.
[3] A. Asadpour, U. Feige, A. Saberi, Santa Claus meets hypergraph matchings, in: Proc. APPROX-RANDOM, 2008, pp. 10–20.
[4] Zs. Baranyai, On the factorization of the complete uniform hypergraph, in: Infinite and Finite Sets, in: Colloq. Math. Soc. János Bolyai, vol. 10, North-Holland, Amsterdam, 1975, pp. 91–108.
[5] F.R.K. Chung, Regularity lemmas for hypergraphs and quasi-randomness, Random Structures Algorithms 2 (1991) 241–252.
[6] A. Czygrinow, V. Kamat, Tight co-degree condition for perfect matchings in 4-graphs, Electron. J. Combin. 19 (2012), P20.
[7] D.E. Daykin, R. Häggkvist, Degrees giving independent edges in a hypergraph, Bull. Aust. Math. Soc. 23 (1981) 103–109.
[8] G.A. Dirac, Some theorems on abstract graphs, Proc. Lond. Math. Soc. 2 (1952) 69–81.
[9] J. Edmonds, Paths, trees, and flowers, Canad. J. Math. 17 (1965) 449–467.
[10] J. Han, Near perfect matchings in k-uniform hypergraphs, Combin. Probab. Comput. (2014), in press.
[11] J. Han, Decision problem for perfect matchings in dense k-uniform hypergraphs, arXiv:1409.5931.
[12] H. Hàn, Y. Person, M. Schacht, On perfect matchings in uniform hypergraphs with large minimum vertex degree, SIAM J. Discrete Math. 23 (2009) 732–748.
[13] S. Janson, T. Łuczak, A. Ruciński, Random Graphs, Wiley–Interscience, 2000.
[14] R.M. Karp, Reducibility among combinatorial problems, Complex. Computer Comput. (1972) 85–103.
[15] M. Karpiński, A. Ruciński, E. Szymańska, Computational complexity of the perfect matching problem in hypergraphs with subcritical density, Internat. J. Found. Comput. Sci. 21 (2010) 905–924.

[16] P. Keevash, F. Knox, R. Mycroft, Polynomial-time perfect matchings in dense hypergraphs (extended abstract), in: Proc. 45h ACM STOC, 2013, pp. 311–320.
[17] P. Keevash, R. Mycroft, A geometric theory for hypergraph matchings, Mem. Amer. Math. Soc. 233 (1098) (2014).
[18] I. Khan, Perfect matchings in 3-uniform hypergraphs with large vertex degree, SIAM J. Discrete Math. 27 (2013) 1021–1039.
[19] I. Khan, Perfect matchings in 4-uniform hypergraphs, arXiv:1101.5675.
[20] D. Kühn, D. Osthus, Matchings in hypergraphs of large minimum degree, J. Graph Theory 51 (2006) 269–280.
[21] D. Kühn, D. Osthus, A. Treglown, Matchings in 3-uniform hypergraphs, J. Combin. Theory Ser. B 103 (2013) 291–305.
[22] A. Lo, K. Markström, Perfect matchings in 3-partite 3-uniform hypergraphs, J. Combin. Theory Ser. A 127 (2014) 22–57.
[23] A. Lo, K. Markström, F-factors in hypergraphs via absorption, Graphs Combin. (2014), http://dx.doi.org/10.1007/s00373-014-1410-8.
[24] K. Markström, A. Ruciński, Perfect matchings (and Hamilton cycles) in hypergraphs with large degrees, European J. Combin. 32 (2011) 677–687.
[25] O. Pikhurko, Perfect matchings and K_4^3-tilings in hypergraphs of large codegree, Graphs Combin. 24 (2008) 391–404.
[26] V. Rödl, A. Ruciński, Dirac-type questions for hypergraphs — a survey (or more problems for Endre to solve), in: An Irregular Mind (Szemerédi is 70), vol. 21, 2010, pp. 1–30.
[27] V. Rödl, A. Ruciński, E. Szemerédi, Perfect matchings in large uniform hypergraphs with large minimum collective degree, J. Combin. Theory Ser. A 113 (2009) 613–636.
[28] E. Szymańska, The complexity of almost perfect matchings in uniform hypergraphs with high codegree, in: 20th International Workshop on Combinatorial Algorithms, IWOCA 2009, in: Lecture Notes in Comput. Sci., vol. 5874, Springer-Verlag, 2009, pp. 438–449.
[29] E. Szymańska, The complexity of almost perfect matchings and other packing problems in uniform hypergraphs with high codegree, European J. Combin. 34 (2013) 632–646.
[30] A. Treglown, Y. Zhao, Exact minimum degree thresholds for perfect matchings in uniform hypergraphs, J. Combin. Theory Ser. A 119 (2012) 1500–1522.
[31] A. Treglown, Y. Zhao, Exact minimum degree thresholds for perfect matchings in uniform hypergraphs II, J. Combin. Theory Ser. A 120 (2013) 1463–1482.
[32] W.T. Tutte, The factorisation of linear graphs, J. Lond. Math. Soc. 22 (1947) 107–111.
[33] J.H. van Lint, R.M. Wilson, A Course in Combinatorics, Cambridge University Press, 2001.
[34] V.A. Vatutin, V.G. Mikhailov, Limit theorems for the number of empty cells in an equiprobable scheme for group allocation of particles, Theory Probab. Appl. 27 (1982) 734–743.