The Decision Problem for Perfect Matchings in Dense Hypergraphs

Luyining Gan
Department of Mathematics and Statistics, University of Nevada, Reno, NV, USA

Jie Han[1]
School of Mathematics and Statistics and Center for Applied Math,
Beijing Institute of Technology, China

Abstract

Given $1 \leq \ell < k$ and $\delta \geq 0$, let $\mathbf{PM}(k, \ell, \delta)$ be the decision problem for the existence of perfect matchings in n-vertex k-uniform hypergraphs with minimum ℓ-degree at least $\delta\binom{n-\ell}{k-\ell}$. For $k \geq 3$, the decision problem in general k-uniform hypergraphs, equivalently $\mathbf{PM}(k, \ell, 0)$, is one of Karp's 21 NP-complete problems. Moreover, for $k \geq 3$, a reduction of Szymańska showed that $\mathbf{PM}(k, \ell, \delta)$ is NP-complete for $\delta < 1 - (1 - 1/k)^{k-\ell}$. A breakthrough by Keevash, Knox and Mycroft [STOC '13] resolved this problem for $\ell = k - 1$ by showing that $\mathbf{PM}(k, k - 1, \delta)$ is in P for $\delta > 1/k$. Based on their result for $\ell = k - 1$, Keevash, Knox and Mycroft conjectured that $\mathbf{PM}(k, \ell, \delta)$ is in P for every $\delta > 1 - (1 - 1/k)^{k-\ell}$.

In this paper it is shown that this decision problem for perfect matchings can be reduced to the study of the minimum ℓ-degree condition forcing the existence of fractional perfect matchings. That is, we hopefully solve the "computational complexity" aspect of the problem by reducing it to a well-known extremal problem in hypergraph theory. In particular, together with existing results on fractional perfect matchings, this solves the conjecture of Keevash, Knox and Mycroft for $\ell \geq 0.4k$.

2012 ACM Subject Classification Theory of computation → Complexity classes

Keywords and phrases Computational Complexity, Perfect Matching, Hypergraph

Digital Object Identifier 10.4230/LIPIcs.ICALP.2022.64

Category Track A: Algorithms, Complexity and Games

1 Introduction

As arguably the most natural extension of graph objects to hypergraphs, matchings have attracted a great deal of attention from both mathematicians and theoretical computer scientists. However, the study of hypergraph matching problems is still a challenging task. One particular reason for this is that finding the maximal matchings in k-uniform hypergraphs for $k \geq 3$ (e.g. 3-partite 3-uniform hypergraphs) is famously NP-complete [14], in contrast to the tractability in the graph case (Edmonds' blossom algorithm [5]).

Hypergraph matchings also find exciting applications in other fields, e.g. the Existence Conjecture of Block Designs [16, 8], Ryser's Conjecture on Latin Squares and Samuels' Conjecture in Probability Theory. For applications on practical problems, one prime example is that Asadpour, Feige and Saberi [2] used hypergraph perfect matchings to study the *Santa Claus problem.*

In this paper we continue the study of the decision problem of perfect matchings in dense hypergraphs, initiated by Karpiński, Ruciński and Szymańska [15]. Given $k \geq 2$, a *k-uniform hypergraph* (or k-graph) H consists of a vertex set $V(H)$ and an edge set $E(H)$, where each edge in $E(H)$ is a set of k vertices of H. A subset $M \subseteq E(H)$ is a *matching* if every two edges from M are vertex-disjoint. A matching in H is called *perfect* if it covers all vertices

[1] Corresponding author.

49th International Colloquium on Automata, Languages, and Programming (ICALP 2022).
Editors: Mikołaj Bojańczyk, Emanuela Merelli, and David P. Woodruff;
Article No. 64; pp. 64:1–64:16
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

of H. Given a k-graph H with an ℓ-element vertex set S (where $0 \leq \ell \leq k-1$) we define $\deg_H(S)$ to be the number of edges containing S. The *minimum ℓ-degree* $\delta_\ell(H)$ of H is the minimum of $\deg_H(S)$ over all ℓ-element sets of vertices in H.

The following decision problem was raised by Keevash, Knox and Mycroft [18], generalizing a problem of Karpiński, Ruciński and Szymańska [15] for the case $\ell = k-1$.

▶ **Problem 1.** *Given integers $\ell < k$ and $\delta \in [0,1]$, denote by $\mathbf{PM}(k,\ell,\delta)$ the problem of deciding whether there is a perfect matching in a given k-graph H on $n \in k\mathbb{N}$ vertices with $\delta_\ell(H) \geq \delta\binom{n-\ell}{k-\ell}$.*

The motivating fact is that for $k \geq 3$, $\mathbf{PM}(k,\ell,0)$ is equivalent to the problem for general k-graphs, so is NP-complete; on the other hand $\mathbf{PM}(k,\ell,\delta)$ is trivially in P when δ is large (e.g., when $\delta > 1-1/k$ by the result of [9]) because all such k-graphs contain perfect matchings. Therefore, it is natural to ask for the point where the behavior changes. A reduction of Szymańska [22] showed that $\mathbf{PM}(k,\ell,\delta)$ is NP-complete for $k \geq 3$ and $\delta < 1-(1-1/k)^{k-\ell}$. In a breakthrough paper, Keevash, Knox and Mycroft [18] conjectured that $1-(1-1/k)^{k-\ell}$ is the turning point and verified the case $\ell = k-1$.

▶ **Conjecture 2** (Keevash, Knox and Mycroft [18]). *For $1 \leq \ell < k$, $\mathbf{PM}(k,\ell,\delta)$ is in P for every $\delta > 1-(1-1/k)^{k-\ell}$.*

Recently, Han and Treglown [13] showed that the conjecture holds for $0.5k \leq \ell \leq (1+\ln(2/3))k \approx 0.59k$. In this paper we verify Conjecture 2 for all $\ell \geq 0.4k$. In fact, our main result reduces the conjecture to the study of the minimum-degree-type threshold for the existence of a perfect fractional matching in k-graphs. To illustrate this, we introduce the following definitions.

Given a k-graph $H = (V,E)$, a *fractional matching* in H is a function $\omega : E \to [0,1]$ such that for each $v \in V(H)$ we have that $\sum_{e \ni v} w(e) \leq 1$. Then $\sum_{e\in E(H)} w(e)$ is the *size* of w. If the size of w in H is n/k then we say that w is a *perfect fractional matching*. Given $k,\ell \in \mathbb{N}$ such that $\ell \leq k-1$, define $c^*_{k,\ell}$ to be the smallest number c such that every k-graph H on n vertices with $\delta_\ell(H) \geq (c+o(1))\binom{n-\ell}{k-\ell}$ contains a perfect fractional matching. The following is our main result.

▶ **Theorem 3.** *Suppose $k,\ell \in \mathbb{N}$ such that $1 \leq \ell \leq k-1$. Then for any $\delta \in (c^*_{k,\ell},1]$, $\mathbf{PM}(k,\ell,\delta)$ is in P. That is, for any $\delta \in (c^*_{k,\ell},1]$, there exists a constant $c = c(k)$ such that there is an algorithm with running time $O(n^c)$ which given any n-vertex k-graph H with $\delta_\ell(H) \geq \delta\binom{n-\ell}{k-\ell}$, determines whether H contains a perfect matching.*

In fact, in [13] a similar result was proved for $\delta \in (\delta^*,1]$ where $\delta^* = \max\{c^*_{k,\ell},1/3\}$. Comparing with their result, Theorem 3 drops the extra $1/3$ and thus extends the result to large values of ℓ, e.g., for $\ell > (1+\ln(2/3))k$.

For the parameter $c^*_{k,\ell}$, Alon, Frankl, Huang, Rödl, Ruciński, and Sudakov [1] in 2012 made the following conjecture.

▶ **Conjecture 4** ([1]). *For all $\ell,k \in \mathbb{N}$, $c^*_{k,\ell} = 1-(1-1/k)^{k-\ell}$.*

They [1] verified the case $k-\ell \leq 4$. The conjecture was further validated by Kühn, Osthus and Townsend [19, Theorem 1.7] for $\ell \geq k/2$ and by Han [10, Theorem 1.5] for $\ell = (k-1)/2$. In a recent work, Frankl and Kupavskii [6] verified this conjecture for $\ell \geq 0.4k$. Unfortunately, despite the efforts from experts in the field, Conjecture 4 is still open and appears to be very challenging for small values of ℓ. In fact, Conjecture 4 is also closely related to an old conjecture of Erdős on the size of the largest matching in hypergraphs (in particular, results of [6, 10] are corollaries of the corresponding progresses on the conjecture of Erdős).

Combining Theorem 3 with the current status on $c^*_{k,\ell}$ we get the following corollary.

▶ **Corollary 5.** *Conjecture 2 holds for $\ell \geq 0.4k$ and for $k - \ell \leq 4$.*

Thus, by Theorem 3, Conjecture 2 holds for all cases when $c^*_{k,\ell} = 1-(1-1/k)^{k-\ell}$, that is, whenever Conjecture 4 holds. Indeed, it is not hard to show that if $\delta > c^*_{k,\ell}$, then the k-graph contains a matching that covers all but exactly k vertices (see Theorem 11). Given this, our result can be viewed as the efficient detection of a certain class of *divisibility constructions* that prevents the existence of perfect matchings. As a consequence, we reduce the decision problem to an extremal problem on the existential problem of a perfect fractional matching, which can be recognized as a resolution on the "computational complexity" aspect of this problem.

We now give an overview of the minimum-degree-type conditions as well as the divisibility constructions.

1.1 Minimum degree conditions and divisibility barriers

The minimum degree conditions forcing a perfect matching have been studied extensively over the last two decades. Focusing on the asymptotical thresholds, all known results support the following conjecture raised by Hàn, Person and Schacht [9]. Note that this corresponds to the case when the decision problem is trivially in P (a trivial algorithm that always outputs yes).

▶ **Conjecture 6** (Hàn–Person–Schacht, [9])**.** *Given $1 \leq \ell < k$, if a k-graph H on n vertices satisfies $\delta_\ell(H) \geq (\max\{1/2, c^*_{k,\ell}\} + o(1))\binom{n-\ell}{k-\ell}$, then H contains a perfect matching.*

The conjecture has attracted a great deal of attention and so far has been verified for $\ell \geq 3k/8$ by Frankl and Kupavskii [6] and a handful of pairs (k, ℓ). Note that this conjecture is slightly weaker than our problem, as e.g. for certain values of ℓ, it *suffices* to show that $c^*_{k,\ell} \leq 1/2$, rather than determining the precise value of $c^*_{k,\ell}$ (and this is the reason that the known record on Conjecture 6 by [6] is slightly wider than that for the conjecture on $c^*_{k,\ell}$).

In fact under the assumption $\delta_\ell(H) \geq (c^*_{k,\ell} + o(1))\binom{n-\ell}{k-\ell}$, Chang, Ge, Han and Wang [3] recently proved that one can find a matching in H of size $n/k-1$ (see Theorem 11). However, such H may or may not have a perfect matching, and, prior to this work, it is not clear how to characterize these two types of k-graphs. To understand this, what is interesting to our problem is the divisibility constructions that achieve the bound 1/2 in the above conjecture. Consider an n-vertex set V with a bipartition $X\dot\cup Y$, where X and Y have almost equal size subject to that $|Y|$ is odd. Now define a k-graph H_0 on V with the edge set consisting of all k-tuples that contain an even number of vertices in Y. It is not hard to see that $\delta_\ell(H) \approx \frac{1}{2}\binom{n-\ell}{k-\ell}$ and H_0 has no perfect matching. To see this, note that any matching in H_0 covers an even number of vertices in Y, so not the entire Y.

One can actually construct such partitions for an arbitrary number of parts. For certain sizes of parts, divisibility conditions similar to the parity issue in the above example prevent the existence of perfect matchings. Thus, our result and algorithm can be viewed as efficient detection of such constructions. Indeed, in the Keevash–Knox–Mycroft proof of Conjecture 2 for $\ell = k-1$, they designed efficient algorithms to exhibit a number of $(O(n^{k+1}))$ such partitions and tested the divisibility (solubility) for each of them. In contrast, we show that one can focus on one partition and prove a sufficient and necessary condition for the existence of a perfect matching solely on that partition. This will be made clear in Section 2.

1.2 Related work

The decision problem for perfect matchings in dense hypergraphs was first raised by Karpiński, Ruciński and Szymańska [15] for the case $\ell = k-1$, where they formulated the problem as $\mathbf{PM}(k,\delta)$ which is equivalent to $\mathbf{PM}(k,k-1,\delta)$ in this paper. They showed that $\mathbf{PM}(k,1/2-\varepsilon)$ is in P for some absolute $\varepsilon > 0$, thus showing that $1/2$ is not the turning point for the change of behavior, while Szymańska's [22] reduction showed that $\mathbf{PM}(k,\delta)$ is NP-complete when $\delta < 1/k$. This leaves a hardness gap for $\delta \in [1/k, 1/2)$. Significant progress was made by Keevash–Knox–Mycroft [17, 18] who showed that $\mathbf{PM}(k,\delta)$ is in P for $\delta > 1/k$. This hardness problem was fully settled by Han [11] who proved that $\mathbf{PM}(k,1/k)$ is in P. Very recently, this result was strengthened by Han and Keevash [12], who showed that the minimum $(k-1)$-degree condition can be weakened to $n/k - c$ for any constant $c > 0$ and their algorithm can actually output the perfect matching provided that one exists.

The similar decision problem for Hamilton cycles (spanning cycles) has also been studied. First, it is well-known that it is NP-complete to determine if a (2-)graph has a Hamilton cycle. A k-graph C is called a tight cycle if its vertices can be listed in a cyclic order so that the edges are all consecutive k-tuples. For tight Hamilton cycles in dense k-graphs, it is shown that there is no such hardness gap as for perfect matchings. That is, it is showed that by Rödl, Ruciński and Szemerédi [21] for n-vertex k-graph H, if $\delta_{k-1}(H) \geq (1/2+o(1))n$ then H contains a tight Hamilton cycle, i.e., the decision problem is trivially in P; on the other hand, Garbe and Mycroft [7] showed that there exists a constant C such that if $\delta_{k-1}(H) \geq n/2 - C$, then the decision problem of tight Hamilton cycles is NP-complete. However, such hardness gap is shown to exist for looser cycles [7].

Han and Treglown [13] considered the similar decision problem for F-factors[2] in graphs and k-graphs. In particular, they determined the turning point for the F-factor problem for graphs and thus disproved a conjecture of Yuster [23].

2 A partition lemma and a structural theorem

To prove Theorem 3, we shall establish a structural theorem (Theorem 9) for perfect matchings, namely, we exhibit a sufficient and necessary condition for the existence of perfect matchings, which, in addition, can be checked in polynomial time. In turn, the heart of the proof of the structural theorem is the lattice-based absorption method developed by Han [11], which features a vertex partition of the given k-graph (Lemma 8).

The first key definition is the *reachability* introduced by Lo and Markström [20] for the absorption property we need for building the perfect matching.

2.1 Reachability

Let H be an n-vertex k-graph. For $i \in \mathbb{N}$ and $\beta \in (0,1)$, we say that two vertices u and v in $V(H)$ are *(β, i)-reachable in H* if there are at least βn^{ik-1} $(ik-1)$-sets S such that both $H[S \cup \{u\}]$ and $H[S \cup \{v\}]$ have perfect matchings. We refer to such a set S as a *reachable $(ik-1)$-set for u and v*. We say a vertex set $U \subseteq V(H)$ is *(β,i)-closed in H* if any two vertices $u, v \in U$ are (β,i)-reachable in H. Given any $v \in V(H)$, define $\tilde{N}_{\beta,i}(v,H)$ to be the set of vertices in $V(H)$ that are (β,i)-reachable to v in H.

[2] Given k-graphs F and H, an F-factor in H is a set of vertex-disjoint copies of F whose union covers $V(H)$.

2.2 Index vector and robust vector

Given an n-vertex k-graph H and integer $r \ge 0$, let $\mathcal{P} = \{V_0, V_1, \dots, V_r\}$ be a partition of $V(H)$ into disjoint vertex sets, namely, $\dot\bigcup_{0 \le i \le r} V_i = V(H)$. In this paper, every partition has an implicit ordering of its parts.

Next we introduce the index vectors and edge-lattices. Given a k-graph H and a partition $\mathcal{P} = \{V_0, V_1, \dots, V_s, V_{s+1}, \dots, V_r\}$ of $V(H)$, the *index vector* $\mathbf{i}_{\mathcal{P}}(e) \in \mathbb{Z}^r$ of an edge $e \in E(H)$ with respect to $\mathcal{P}$ is the vector whose coordinates are the sizes of the intersections of e with each part of $\mathcal{P}$ *except* V_0, namely, $\mathbf{i}_{\mathcal{P}}(e)|_i = |e \cap V_i|$ for $i \in [r]$, where $\mathbf{v}|_i$ is defined as the ith digit of $\mathbf{v}$. For any $\mathbf{v} = \{v_1, \dots, v_r\} \in \mathbb{Z}^r$, let $|\mathbf{v}| := \sum_{i=1}^r v_i$. Here we say that $\mathbf{v} \in \mathbb{Z}^r$ is a *k-vector* if it has non-negative coordinates and $|\mathbf{v}| = k$. In previous works, for $\mu > 0$, the set of μ-robust vectors (denoted by $I^\mu_{\mathcal{P}}(H)$) is defined as the vectors $\mathbf{i} \in \mathbb{Z}^r$ such that H contains at least μn^k edges whose index vectors are equal to $\mathbf{i}$. In this paper we need a more detailed description of robust vectors – where we need to distinguish the roles of two different groups of V_i.

▶ **Definition 7** (μ-robust vectors). *Let $\mathcal{P} = \{V_0, V_1, \dots, V_s, V_{s+1}, \dots, V_r\}$ be a partition of $V(H)$. Given $\mu > 0$, define $I^\mu_{\mathcal{P}}(H) := I^\mu_{\mathcal{P},1}(H) \cup I^\mu_{\mathcal{P},2}(H)$ as the union of the following two sets:*

1. *the set $I^\mu_{\mathcal{P},1}(H)$ consists of all k-vectors $\mathbf{i} \in \mathbb{Z}^r$ such that $\mathbf{i}|_i = 0$ for $i \in \{0, 1, \dots, s\}$ and H contains at least μn^k edges e with $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$;*
2. *the set $I^\mu_{\mathcal{P},2}(H)$ consists of all k-vectors $\mathbf{i} \in \mathbb{Z}^r$ such that $\mathbf{i}|_i = 1$ for some $i \in [s]$, $\mathbf{i}|_j = 0$ for $j \in \{0, 1, \dots, s\} \setminus \{i\}$ and every vertex $v \in V_i$ is in at least μn^{k-1} edges e with $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$.*

The new ingredient of this definition is the assumption (2), which helps us to classify the vertices which do not enjoy the reachability information.

Now we are ready to state our partition lemma, which outputs a refined partition compared to the partition lemmas in [11, 13]. Throughout the paper, we write $\alpha \ll \beta \ll \gamma$ to mean that it is possible to choose the positive constants α, β, γ from right to left. More precisely, there are increasing functions f and g such that, given γ, whenever we choose some $\beta \le f(\gamma)$ and $\alpha \le g(\beta)$, the subsequent statement holds. Hierarchies of other lengths are defined analogously.

▶ **Lemma 8.** *Given integers $k \ge 3$, $C > 0$ and real $\delta > 0$, suppose we have $1/n_0 \ll \mu \ll \beta \ll \delta' \ll \delta, 1/k, 1/C$. Given an n-vertex k-graph H with $n \ge n_0$ and $\delta_\ell(H) \ge \delta\binom{n-\ell}{k-\ell}$, there is a partition $\mathcal{P}$ of $V(H)$ as*

$$\mathcal{P} = \{V_0, V_1, \dots, V_s, V_{s+1}, \dots, V_r\}$$

such that with $c := \lfloor 1/\delta \rfloor$

1. *$s \le 2^{\binom{c+k-2}{k-1}}$ and $r - s \le c$,*
2. *$|V_0| \le k^{2^{\binom{c+k-2}{k-1}}} \left(k\binom{k+2^{\binom{c+k-2}{k-1}}+c-1}{k} + \binom{c+k-2}{k-1}C \right)$ and $|\bigcup_{0 \le i \le s} V_i| \le c\delta' n$,*
3. *for $1 \le i \le s$, $|V_i| \ge (k-1)|V_0| + k\binom{k+2^{\binom{c+k-2}{k-1}}+c-1}{k} + \binom{c+k-2}{k-1}C$,*
4. *for $1 \le i \le s$, there exists $\mathbf{i} \in I^\mu_{\mathcal{P},2}(H)$ such that $\mathbf{i}|_i = 1$,*
5. *for $s+1 \le i \le r$, $|V_i| \ge \delta' n/2$ and V_i is $(\beta, 2^c)$-closed in $H[\bigcup_{s+1 \le i \le r} V_i]$.*

In particular, such a partition $\mathcal{P}$ of H can be found in time $O(n^{2^{c-1}k+1})$.

2.3 Lattices, solubility and the structural theorem

Keevash, Knox and Mycroft [18] introduced the following notions, which help us to transfer the divisibility problem to an algebraic setting as follows.

Given a partition $\mathcal{P}$ of m parts, denote by $L^\mu_{\mathcal{P}}(H)$ the *lattice* (additive subgroup) in $\mathbb{Z}^m$ generated by $I^\mu_{\mathcal{P}}(H)$. We write $L^m_{\max}$ for the lattice generated by all k-vectors, that is, $L^m_{\max} := \{\mathbf{v} \in \mathbb{Z}^m : k \text{ divides } |\mathbf{v}|\}$.

Suppose $L \subset L^{|\mathcal{P}|}_{\max}$ is a lattice in $\mathbb{Z}^{|\mathcal{P}|}$, where $\mathcal{P}$ is a partition of a set V. The *coset group* of $(\mathcal{P}, L)$ is $Q = Q(\mathcal{P}, L) := L^{|\mathcal{P}|}_{\max}/L$. For any $\mathbf{i} \in L^{|\mathcal{P}|}_{\max}$, the *residue* of $\mathbf{i}$ in Q is $R_Q(\mathbf{i}) := \mathbf{i} + L$. For any $A \subseteq V$ of size divisible by k, the *residue* of A in Q is $R_Q(A) := R_Q(\mathbf{i}_{\mathcal{P}}(A))$.

Let $q \in \mathbb{N}$. A (possibly empty) matching M in H of size at most q is a *q-solution* for $(\mathcal{P}, L^\mu_{\mathcal{P}}(H))$ (in H) if $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M)) \in L^\mu_{\mathcal{P}}(H)$; we say that $(\mathcal{P}, L^\mu_{\mathcal{P}}(H))$ is *q-soluble* if it has a q-solution. We also need a strengthening of this definition as follows. Given a set $U \subset V(H)$, we define that $(\mathcal{P}, L^\mu_{\mathcal{P}}(H))$ is *(U, q)-soluble* if there is a matching M in H such that M covers U and M is a $(|U| + q)$-solution.

In our proof, we shall pick a suitable $\mu > 0$ and let q be an upper bound of the order of the coset group $Q = L^{|\mathcal{P}|}_{\max}/L^\mu_{\mathcal{P}}(H)$ (then a trivial bound is $q = \binom{r+k-1}{k}$, the number of ee) and U be the part V_0. Then we show that H has a perfect matching if and only if $(\mathcal{P}, L^\mu_{\mathcal{P}}(H))$ is (V_0, q)-soluble.

▶ **Theorem 9** (Structural Theorem). *Let $k, \ell, q \in \mathbb{N}$ where $\ell \le k-1$ and let $\gamma > 0$ be given. There exist $n_0, C := C(k, q) \in \mathbb{N}$ and $\beta, \mu > 0$ such that*

$$1/n_0 \ll \beta, \mu \ll \delta' \ll \gamma, c^*_{k,\ell}, 1/q, 1/C, 1/k. \qquad (1)$$

Let H be an n-vertex k-graph with $\delta_\ell(H) \ge (c^*_{k,\ell} + \gamma)\binom{n-\ell}{k-\ell}$, where $n \ge n_0$ and k divides n. Suppose $\mathcal{P}$ is a partition of $V(H)$ satisfying Lemma 8 (1)-(5) with $\delta = c^*_{k,\ell}$. Moreover, suppose $|Q(\mathcal{P}, L^\mu_{\mathcal{P}}(H))| \le q$. Then H contains a perfect matching if and only if $(\mathcal{P}, L^\mu_{\mathcal{P}}(H))$ is (V_0, q)-soluble.

3 Highlights of the proof: a comparison with the Han–Treglown proof

The basic idea for establishing the structural theorem is to distinguish the roles of *robust* and *non-robust* edges: to avoid the divisibility barriers, we may have to use edges with certain (combination of) index vectors. For some index vectors $\mathbf{v}$ there are many edges e with $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{v}$, namely, there are many "replacements" even when we are forbidden from using, say, a small number of such edges. For other index vectors $\mathbf{v}$ there are few edges e with $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{v}$, so we have to be careful when using such edges. In fact, the algebraic setting allows us to show that one can restrict the attention to only a *constant* number of such non-robust edges (using the lattice and coset group arguments), and thus this can be tested by brute force. Then the rest of the proof follows from the lattice-based absorption argument. Roughly speaking, it reserves a small matching which can be used to turn an almost perfect matching to a perfect matching given certain divisibility condition on the leftover vertices.

In [13] Han and Treglown proved our Theorem 3 under the additional assumption that $\delta > 1/3$, which gives a resolution of Conjecture 2 for $0.5k \le \ell \le 0.59k$. Embarrassingly, this does not solve the conjecture for $\ell = k-2$, which might be considered as the easiest case after the resolution of the case $\ell = k-1$. Below we shall first outline the proof in [13], and then explain our innovation compared with their approach and how such an improvement is achieved.

The partition lemma used in [13] is Lemma 12 in this paper (which we use as a building block to establish our partition). The key problem is that when $\ell < k-1$, one can not apply Lemma 12 directly to the k-graph H, as in H there might be a set W of vertices v which are not reachable to many vertices, namely, $|\tilde{N}_{\beta,i}(v, H)|$ is small for any proper choice of $\beta > 0$ and $i \in \mathbb{N}$. However, it is straightforward to show that $|W|$ is small, and (after some work) we can apply Lemma 12 with $S = V(H) \setminus W$ and get a partition of $V(H) \setminus W$. Now we face the following challenge.

▶ **Problem.** *Suppose $|W| = \Omega(n)$. How do we find a matching M covering W so that $H - V(M)$ has a perfect matching (or conclude that none exists)?*

The problem is trivial if $|W|$ is a constant, for which we can do brute force search for a matching M of constant size, which involves $O(n^{|W|})$ possibilities; otherwise it is hopeless without further assumptions.

Furthermore, it was not clear how to deal with the vertices of W by absorption, as $|W|$ might be smaller than the threshold for μ-robustness but still a small linear size, i.e., $\varepsilon n \le |W| < \mu n$, so that every vector touching W will not be recorded as a μ-robust vector. The proof in [13] avoided the "decision" part of the problem by assuming $\delta > 1/3$, so that when W is non-empty $V(H) \setminus W$ is closed, in which case H always contains a perfect matching (so any matching M covering W will work). Therefore, the problem is left open for $\delta < 1/3$ (i.e., for $\ell > (1 + \ln(2/3))k \approx 0.595k$).

We also note that the existence of W is not a problem in the existential results in the literature. For previous works on sufficient (minimum-degree-type) conditions for perfect matchings, those vertices can be put into a small matching of small linear size, whose removal does not affect much the minimum-degree conditions, guaranteeing that the absorption can proceed after the removal of this small matching.

Our new proof can be seen as a considerable refinement to the previous approach, where we strengthen our control on both the partition and the μ-robust vectors. As mentioned earlier, our new proof features a finer partition lemma (Lemma 8) than previous ones, where we classify vertices in W as well. More precisely, we first partition $S := V(H) \setminus W$, the set of vertices which are 1-reachable to $\Omega(n)$ other vertices, by Lemma 12 and denote the partition by $\mathcal{P}_1 = \{W_1, \dots, W_d\}$. Then we classify vertices of W according to their edge distributions in $\mathcal{P}_1$, that is, we obtain a partition of W by collecting vertices with common robust edge vectors together, so that the partition satisfies Definition 7 (2). Next we put the clusters that are too small (smaller than a certain constant) to a trash set V_0 in a recursive manner. This results a trash set V_0 of constant order, and because we have no control on V_0 at all, we will check how to match V_0 by brute force in time $O(n^{|V_0|})$. Now the good point is that all clusters survived from this greedy process have a good (though still constant) size (Lemma 8 (3)), which is enough (and crucial) for a (refined) absorption argument to work in later proofs. Since all the above procedures can be done in polynomial time, we get the desired polynomial-time algorithm for the decision problem $\mathbf{PM}(k, \ell, \delta)$.

4 Proof of Theorem 3

Now we prove Theorem 3. Recall that $c^*_{k,\ell} \ge c^*_{k,k-1} = 1/k$. Then $\lfloor 1/c^*_{k,\ell} \rfloor \le k$. Let $C := C(k, q)$ be given by Theorem 9 and let

$$q := \binom{k + 2^{\binom{2k-2}{k-1}} + k - 1}{k}.$$

Suppose we have constants satisfying the hierarchy (1).

Both Lemma 8 and Theorem 9 require that n is larger than a constant n_0, and by custom k-graphs with less than n_0 vertices can be tested by brute force. By Lemma 8, in time $O(n^{2^{k-1}k+1})$ we can find a partition $\mathcal{P}$ satisfying Lemma 8 (1)-(5). Because of Lemma 8 (1), we know $r \le k + 2^{\binom{2k-2}{k-1}}$ and obtain that $|Q(\mathcal{P}, L^{\mu}_{\mathcal{P}}(H))| \le \binom{r+k-1}{k} \le q$ (no matter what $L^{\mu}_{\mathcal{P}}(H)$ actually is). Then by Theorem 9, to determine if H contains a perfect matching it suffices to test if $(\mathcal{P}, L^{\mu}_{\mathcal{P}}(H))$ is (V_0, q)-soluble. This can be done by testing whether any matching M of size at most $|V_0| + q$ covering V_0 is a solution of $(\mathcal{P}, L^{\mu}_{\mathcal{P}}(H))$, in time $O(n^{|V_0|+q})$. The overall time is polynomial in n because

$$q = \binom{k + 2^{\binom{2k-2}{k-1}} + k - 1}{k}$$

and

$$|V_0| \le k 2^{\binom{2k-2}{k-1}} \left(k \binom{2^{\binom{2k-2}{k-1}} + 2k - 1}{k} + \binom{2k-2}{k-1} C \right),$$

where we recall that $C := C(k, q)$ only depends on k.

Organization. The rest of this paper is organized as follows. Note that it remains to prove Lemma 8 and Theorem 9. We collect and prove a number of auxiliary results in Section 5, and give a proof of Lemma 8 in Section 6. In Section 7, we prove an absorbing lemma, which is an important component of the proof of Theorem 9. The proof of Theorem 9 is presented in Section 8.

5 Useful tools

In this section we collect together some results that will be used in our proof of Theorem 9. When considering ℓ-degree together with ℓ'-degree for some $\ell' \neq \ell$, the following proposition is very useful.

▶ **Proposition 10.** *Let $0 \le \ell \le \ell' < k$ and H be a k-graph. If $\delta_{\ell'}(H) \ge x\binom{n-\ell'}{k-\ell'}$ for some $0 \le x \le 1$, then $\delta_{\ell}(H) \ge x\binom{n-\ell}{k-\ell}$.*

Proof. Since $\ell \le \ell'$, we count $\delta_\ell(H)$ by

$$\begin{aligned}
\delta_\ell(H) &\ge \delta_{\ell'}(H)\binom{n-\ell}{\ell'-\ell}\frac{1}{\binom{k-\ell}{\ell'-\ell}} \\
&\ge x\binom{n-\ell'}{k-\ell'}\binom{n-\ell}{\ell'-\ell}\frac{1}{\binom{k-\ell}{\ell'-\ell}} \\
&\ge x\binom{n-\ell}{k-\ell}\binom{k-\ell}{\ell'-\ell}\frac{1}{\binom{k-\ell}{\ell'-\ell}} = x\binom{n-\ell}{k-\ell}
\end{aligned}$$

where the last inequality is from $\binom{a}{b}\binom{b}{c} = \binom{a}{c}\binom{a-c}{b-c}$. ◀

5.1 Almost perfect matchings

Let $k, \ell \in \mathbb{N}$ where $\ell \le k-1$. Given $D \in \mathbb{N}$, define $\delta(k, \ell, D)$ as the smallest number δ such that every k-graph H on $n \in k\mathbb{N}$ vertices with $\delta_\ell(H) \ge (\delta + o(1))\binom{n-\ell}{k-\ell}$ contains a matching covering all but at most D vertices. It is proved in [13] that $\delta(k, \ell, k) \le \max\{1/3, c^*_{k,\ell}\}$. We need the extra term $1/3$ removed, which was very recently proved by Chang, Ge, Han and Wang [3].

▶ **Theorem 11** ([3]). *Let k, ℓ be integers such that $1 \leq \ell \leq k-1$ and $\gamma > 0$, then there exists $n_0 \in \mathbb{N}$ such that the following holds for $n \geq n_0$. Suppose H is an n-vertex k-graph with $\delta_\ell(H) \geq (c^*_{k,\ell} + \gamma)\binom{n-\ell}{k-\ell}$, then H contains a matching M that covers all but at most $2k - \ell - 1$ vertices. In particular, when $n \in k\mathbb{N}$, M is a perfect matching or covers all but exactly k vertices, namely, $\delta(k, \ell, k) \leq c^*_{k,\ell}$.*

To build the partition, we need the following partition lemma from [13].

▶ **Lemma 12** ([13, Lemma 6.3]). *Given $\delta' > 0$, integers $c, k \geq 2$ and $0 < \alpha \ll 1/c, \delta'$, there exists a constant $\beta > 0$ such that the following holds for all sufficiently large n. Assume H is an n-vertex k-graph and $S \subseteq V(H)$ is such that $|\tilde{N}_{\alpha,1}(v, H) \cap S| \geq \delta' n$ for any $v \in S$. Further, suppose every set of $c+1$ vertices in S contains two vertices that are $(\alpha, 1)$-reachable in H. Then in time $O(n^{2^{c-1}k+1})$ we can find a partition $\mathcal{P}$ of S into $V_1, \dots, V_r$ with $r \leq \min\{c, 1/\delta'\}$ such that for any $i \in [r]$, $|V_i| \geq (\delta' - \alpha)n$ and V_i is $(\beta, 2^{c-1})$-closed in H.*

To deal with the vertices that are reachable to few other vertices, we collect them by the following greedy process. Note that a similar lemma was used in [4].

▶ **Lemma 13.** *Let integers $c, k \geq 2$ be given and suppose $1/n \ll \delta' \ll \alpha, 1/k, 1/c$. Assume that H is a k-graph on n vertices satisfying that every set of $c+1$ vertices contains two vertices that are $(2\alpha, 1)$-reachable in H. Then in time $O(cn^{k+1})$ we can find a set of vertices $S \subseteq V(H)$ with $|S| \geq (1 - c\delta')n$ such that $|\tilde{N}_{\alpha,1}(v, H[S])| \geq \delta' n$ for any $v \in S$.*

We remark that in the above lemma it is important to obtain the conclusion on $\tilde{N}_{\alpha,1}(v, H[S])$ rather than $\tilde{N}_{\alpha,1}(v) \cap S$. Indeed, in the latter one the reachable sets are still defined in H, so may contain vertices in $V(H) \setminus S$. This is not strong enough in our later proof (see Lemma 14 and its proof).

Proof. Let H be a k-graph on n vertices satisfying the condition of Lemma 13. We greedily identify vertices with few "reachable neighbors" and remove the vertex together with the vertices reachable to it from H. Set $V_0 := V(H)$. First, for every two vertices $u, v \in V(H)$, we determine if they are $(\alpha, 1)$-reachable in H, which can be done by testing if any $(k-1)$-set is a reachable set in time $O(n^{k-1})$. Summing over all pairs of vertices, this step can be done in time $O(n^{k+1})$. Then we check if there is a vertex $v_0 \in V_0$ such that $|\tilde{N}_{\alpha,1}(v_0, H)| < \delta' n$ in time $O(n^2)$. If there exists such a vertex v_0, then let $A_0 := \{v_0\} \cup \tilde{N}_{\alpha,1}(v_0, H)$ and let $V_1 := V_0 \setminus A_0$. Next, we check V_1, that is, if there exists a vertex $v_1 \in V_1$ such that $|\tilde{N}_{\alpha,1}(v_1, H[V_1])| < \delta' n$, then let $A_1 := \{v_1\} \cup \tilde{N}_{\alpha,1}(v_1, H[V_1])$ and let $V_2 := V_1 \setminus A_1$ and repeat the procedure until no such v_j exists.

Suppose we stop and obtain a set of vertices $v_0, \dots, v_s$. We claim that $s < c$ and thus $|\bigcup_{0 \leq i \leq s} A_i| \leq c\delta' n$. Indeed, otherwise consider $v_0, \dots, v_c$, the first $c+1$ of them and we shall show that every pair of them is not $(2\alpha, 1)$-reachable in H, contradicting our assumption. Given $0 \leq i < j \leq c$, as $v_j \notin \tilde{N}_{\alpha,1}(v_i, H[V_i])$, v_i and v_j have less than αn^{k-1} 1-reachable sets in $H[V_i]$. Also, because $\delta' \ll \alpha, 1/c$, there are at most $c\delta' n \cdot n^{k-2} \leq \alpha n^{k-1}$ 1-reachable sets in $H \setminus H[V_i]$. These two together yield that v_i and v_j are not $(2\alpha, 1)$-reachable in H.

This greedy procedure needs to recompute $\tilde{N}_{\alpha,1}(v, H[V_i])$ at each time and can be done in time $O(cn^{k+1})$. Set $S := V(H) \setminus (\bigcup_{0 \leq i \leq s} A_i)$. We have $|S| \geq (1 - c\delta')n$ and $|\tilde{N}_{\alpha,1}(v, H[S])| \geq \delta' n$ for every $v \in S$. ◀

6 Proof of Lemma 8

Let H be an n-vertex k-graph and define

$$1/n_0 \ll \mu \ll \beta \ll \alpha \ll \gamma, \delta' \ll \delta, 1/k, 1/C.$$

Assume $n \geq n_0$ and k divides n. Write $c := \lfloor 1/\delta \rfloor$, then by Proposition 10 we have

$$(c+1)\delta_1(H) > (c+1)\delta\binom{n-1}{k-1} > (1+\gamma)\binom{n-1}{k-1}.$$

Thus every set of $c+1$ vertices of $V(H)$ contains two vertices that are $(2\alpha, 1)$-reachable, as otherwise, by the inclusion-exclusion principle and $\alpha \ll \gamma, \delta$

$$n \geq (c+1)\delta_1(H) - \binom{c+1}{2} \cdot 2\alpha n^{k-1} \geq (1+\gamma)\binom{n-1}{k-1} - (c+1)^2\alpha n^{k-1} > n$$

a contradiction.

By Lemma 13, we find $S \subseteq V(H)$ with $|S| \geq (1-c\delta')n$ such that $|\tilde{N}_{\alpha,1}(v, H[S])| \geq \delta' n$ for any $v \in S$, in time $O(n^{k+1})$. Let $V' := V(H) \setminus S$ and thus $|V'| \leq c\delta' n$. Apply Lemma 12 to $H[S]$, and in time $O(n^{2^{c-1}k+1})$ we find a partition $\mathcal{P}_1$ of S into $W_1, \dots, W_d$ with $d \leq c$ such that for $i \in [d]$, $|W_i| \geq (\delta' - \alpha)n$ and W_i is $(\beta, 2^{c-1})$-closed in $H[S]$.

Let I_d^{k-1} be the set of all $(k-1)$-vectors on $\mathcal{P}_1$ and note that $|I_d^{k-1}| = \binom{d+k-2}{d-1}$. Let $\mathcal{I}$ be the collection of all subsets of I_d^{k-1} and clearly $|\mathcal{I}| = 2^{|I_d^{k-1}|} = 2^{\binom{d+k-2}{d-1}}$. We classify the vertices in V' by the types of the edges in which they are contained. Indeed, for $I \in \mathcal{I}$, let V_I be the collection of vertices $v \in V'$ such that the following two properties hold:

- for every $\mathbf{i} \in I$, there are at least μn^{k-1} edges e of H such that $v \in e$ and $\mathbf{i}_{\mathcal{P}_1}(e \setminus \{v\}) = \mathbf{i}$;
- for every $\mathbf{i} \notin I$, there are fewer than μn^{k-1} edges e of H such that $v \in e$ and $\mathbf{i}_{\mathcal{P}_1}(e \setminus \{v\}) = \mathbf{i}$.

Clearly this defines a partition of V'. Moreover, note that $V_\emptyset = \emptyset$ – this is because any vertex in $V_\emptyset$ has vertex degree at most $2\binom{d+k-2}{d-1}\mu n^{k-1} < \delta_1(H)$, violating the minimum degree assumption. In particular, this implies 4. Note that this partition can be built by reading the edges for each $v \in V'$, so in time $O(n^k)$. Next we collect the parts that are too small and put them into a trash set V_0 in a recursive manner.

We first sort V_I, $I \in \mathcal{I}$ such that $|V_I|$ is increasing. Next, starting from $V_0 = \emptyset$, we *recursively* check in time $O(|\mathcal{I}|n)$ if next V_I, $I \in \mathcal{I}$ in the sequence satisfies that

$$|V_I| < (k-1)|V_0| + b, \text{ where } b := k\binom{k + 2^{\binom{c+k-2}{c-1}} + c - 1}{k} + \binom{c+k-2}{c-1}C$$

and if yes, put all vertices of V_I to V_0 (note here that V_0 is dynamic). Because $|\mathcal{I}| = 2^{\binom{d+k-2}{d-1}}$, straightforward computation shows that after the process we have

$$|V_0| \leq \frac{k^{|\mathcal{I}|} - 1}{k-1} b \leq k^{2^{\binom{c+k-2}{c-1}}} \left(k\binom{k + 2^{\binom{c+k-2}{c-1}} + c - 1}{k} + \binom{c+k-2}{c-1}C \right).$$

At last, in constant time we remove the empty clusters and relabel the parts V_I's to $V_1, \dots, V_s$, and relabel the parts of $\mathcal{P}_1$ as $V_{s+1}, \dots, V_r$. The resulting partition satisfies all desired properties in the lemma and the overall running time is $O(n^{2^{c-1}k+1})$.

7 An absorption lemma

The following result guarantees our collection $\mathcal{E}_{abs}$ of absorbing sets in the proof of Theorem 9. The absorption method is by now a standard way to turn an almost spanning structure to a spanning one. Here we use a variant called *lattice-based absorption method*, developed by Han [11]. We remark that the following lemma is very similar to that [11, Lemma 3.4], and the only difference is because of our refined definition of robust vectors $I^{\mu}_{\mathcal{P}}(H)$.

▶ **Lemma 14** (Absorption Lemma). *Suppose $k \geq 3$, $\delta > 0$ and let $t := 2^{\lfloor 1/\delta \rfloor}$. Suppose that*

$$1/n \ll 1/c' \ll \beta, \mu \ll 1/t, 1/k.$$

Let H be an n-vertex k-graph with a partition $\mathcal{P}$ of $V(H)$ satisfying Lemma 8 (1)-(5), where $n \geq n_0$ and k divides n. Let $n_1 := |\bigcup_{s+1 \leq i \leq r} V_i|$ (where r, s are from the statement of Lemma 8). Then there is a family $\mathcal{E}_{abs}$ consisting of at most $c' \log n_1$ disjoint tk^2-sets such that for each $A \in \mathcal{E}_{abs}$, $H[A]$ contains a perfect matching and every k-set $S \subseteq V(H)$ with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$ has at least $\sqrt{\log n_1}$ absorbing tk^2-sets in $\mathcal{E}_{abs}$.

Proof. Roughly speaking, in the proof we first exhibit a large number of absorbing sets for each k-set S with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$, and then show that the desired family $\mathcal{E}_{abs}$ can be obtained by standard probabilistic arguments. Our first task is to prove the following claim.

▷ Claim 15. Any k-set S with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$ has at least $\mu^{t+1}\beta^{k+1}n_1^{tk^2}$ absorbing tk^2-sets which consist of vertices in $\bigcup_{s+1 \leq i \leq r} V_i$ only.

Proof. We split the proof into two cases regarding to $I^{\mu}_{\mathcal{P},1}(H)$ and $I^{\mu}_{\mathcal{P},2}(H)$. Note that all reachable sets will be constructed with vertices in $\bigcup_{s+1 \leq i \leq r} V_i$ only.

▶ **Case 1.** *Suppose $\mathbf{i} \in I^{\mu}_{\mathcal{P},1}(H)$.*

For a k-set $S = \{y_1, \dots, y_k\}$ with $\mathbf{i}_{\mathcal{P}}(S) = \mathbf{i}$, we construct absorbing tk^2-sets for S as follows. We first fix an edge $W = \{x_1, \dots, x_k\}$ in H such that $\mathbf{i}_{\mathcal{P}}(W) = \mathbf{i}$ and $W \cap S = \emptyset$. Note that we have at least $\mu n^k - k n_1^{k-1} > \frac{\mu}{2} n^k$ choices for such an edge. Without loss of generality, we may assume that for all $i \in [k]$, x_i, y_i are in the same part V_j of $\mathcal{P}$ for $j > s$. Recall that by Lemma 8 (5) V_j is (β, t)-closed in $H[\bigcup_{s+1 \leq i \leq r} V_i]$. Since x_i is (β, t)-reachable to y_i, there are at least βn_1^{tk-1} $(tk-1)$-sets T_i such that both $H[T_i \cup \{x_i\}]$ and $H[T_i \cup \{y_i\}]$ have perfect matchings. We pick disjoint reachable $(tk-1)$-sets for each x_i, y_i, $i \in [k]$ greedily, while avoiding the existing vertices. Since the number of existing vertices is at most $tk^2 + k$, we have at least $\frac{\beta}{2} n_1^{tk-1}$ choices for such $(tk-1)$-sets in each step. Note that $W \cup T_1 \cup \dots \cup T_k$ is an absorbing set for S. First, it contains a perfect matching because each $T_i \cup \{x_i\}$ for $i \in [k]$ spans t vertex-disjoint edges. Second, $H[W \cup T_1 \cup \dots \cup T_k \cup S]$ also contains a perfect matching and each $T_i \cup \{y_i\}$ for $i \in [k]$ spans t vertex-disjoint edges. There were at least $\frac{\mu}{2} n_1^k$ choices for W and at least $\frac{\beta}{2} n_1^{tk-1}$ choices for each T_i. Thus we find at least

$$\frac{\mu}{2} n^k \times \frac{\beta^k}{2^k} n_1^{tk^2-k} \times \frac{1}{(tk^2)!} \geq \mu \beta^{k+1} n_1^{tk^2}$$

absorbing tk^2-sets for S.

▶ **Case 2.** *Suppose $\mathbf{i} \in I^{\mu}_{\mathcal{P},2}(H)$.*

Suppose $S = \{v_1, y_2, \dots, y_k\}$ with $\mathbf{i}_{\mathcal{P}}(S) = \mathbf{i}$ and $v_1 \in V_i$ for some $i \in [s]$. We construct absorbing tk^2-sets for S as follows. We fix an edge with vertex set $W = \{v_1, x_2, \dots, x_k\}$ for $x_2, \dots, x_k \in \bigcup_{s+1 \leq j \leq r} V_j \setminus \{y_2, \dots, y_k\}$ such that $\mathbf{i}_{\mathcal{P}}(W) = \mathbf{i}_{\mathcal{P}}(S) = \mathbf{i}$ and $W \cap S = \{v_1\}$. Note that by Lemma 8 (4) we have at least $\mu n^{k-1} - (k-1) n_1^{k-2} > \frac{\mu}{2} n^{k-1}$ choices for W (and $x_2, \dots, x_k$ are in $\bigcup_{s+1 \leq i \leq r} V_i$, by the definition of $I^{\mu}_{\mathcal{P},2}(H)$). Without loss of generality, we may assume that for all $i \in \{2, \dots, k\}$, x_i, y_i are in the same part V_j of $\mathcal{P}$, $j > s$. Since x_i is (β, t)-reachable to y_i, there are at least βn_1^{tk-1} $(tk-1)$-sets T_i in $V(H) \setminus V_0$ such that both $H[T_i \cup \{x_i\}]$ and $H[T_i \cup \{y_i\}]$ have perfect matchings. We pick disjoint reachable $(tk-1)$-sets in $V(H) \setminus V_0$ for each x_i, y_i, $i \in \{2, \dots, k\}$ greedily, while avoiding the existing vertices. Since the number of existing vertices is at most $tk(k-1) + (k-1)$, we have at least

$\frac{\beta}{2}n_1^{tk-1}$ choices for such $(tk-1)$-sets in each step. At last, let us pick a matching M of size t in H that is vertex disjoint from the existing vertices (the purpose is to let the absorbing set contain exactly tk^2 vertices). For the number of choices for $V(M)$, we can sequentially choose disjoint edges satisfying any μ-robust edge vector $\mathbf{i} \in I^{\mu}_{\mathcal{P},1}(H)$ and infer that there are at least $\frac{1}{2}\mu^t n^{tk}$ choices.

Note that each choice of $(W \setminus \{v_1\}) \cup T_2 \cup \cdots \cup T_k \cup V(M)$ is an absorbing set for S. First, it contains a perfect matching because each $T_i \cup \{x_i\}$ for $i \in \{2, \ldots, k\}$ spans t vertex-disjoint edges and M is a matching. Second, $H[W \cup T_1 \cup \cdots \cup T_k \cup S]$ also contains a perfect matching as each $T_i \cup \{y_i\}$ for $i \in \{2, \ldots, k\}$ spans t vertex-disjoint edges, W is an edge and M is a matching. There were at least $\frac{\mu}{2}n_1^{k-1}$ choices for W and at least $\frac{\beta}{2}n_1^{tk-1}$ choices for each T_i and $\frac{1}{2}\mu^t n^{tk}$ choices for $V(M)$. Thus we find at least

$$\frac{\mu}{2}n^k \times \left(\frac{\beta}{2}n_1^{tk-1}\right)^{k-1} \times \frac{1}{2}\mu^t n^{tk} \times \frac{1}{(tk^2)!} \geq \mu^{t+1}\beta^k n_1^{tk^2}$$

absorbing tk^2-sets for S, with vertices from $\bigcup_{s+1 \leq i \leq r} V_i$ only. ◁

Continuing the proof of Lemma 14, we pick a family $\mathcal{E}$ of tk^2-sets by including every tk^2-subset of $\bigcup_{s+1 \leq i \leq r} V_i$ with probability $p = c' n_1^{-tk^2} \log n_1$ independently, uniformly at random. Then the expected number of elements in $\mathcal{E}$ is $p\binom{n_1}{tk^2} \leq \frac{c'}{tk^2} \log n_1$ and the expected number of intersecting pairs of tk^2-sets is at most

$$p^2 \binom{n_1}{tk^2} \times tk^2 \times \binom{n_1}{tk^2-1} \leq \frac{c'^2(\log n_1)^2}{n_1} = o(1).$$

Then by Markov's inequality, with probability at least $1 - 1/(tk^2) - o(1)$, $\mathcal{E}$ contains at most $c' \log n_1$ sets and they are pairwise vertex disjoint.

For every k-set S with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$, let X_S be the number of absorbing sets for S in $\mathcal{E}$. Then by Claim 15,

$$\mathbb{E}(X_S) \geq p\mu^{t+1}\beta^{k+1}n_1^{tk^2} = \mu^{t+1}\beta^{k+1}c' \log n_1.$$

By Chernoff's bound,

$$\mathbb{P}\left(X_S \leq \frac{1}{2}\mathbb{E}(X_S)\right) \leq \exp\left\{-\frac{1}{8}\mathbb{E}(X_S)\right\} \leq \exp\left\{-\frac{\mu^{t+1}\beta^{k+1}c' \log n_1}{8}\right\} = o(n^{-k}),$$

since $1/c' \ll \beta, \mu \ll 1/m$. Thus, with probability $1 - o(1)$, for each k-set S with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$, there are at least

$$\frac{1}{2}\mathbb{E}(X_S) \geq \frac{\mu^{t+1}\beta^{k+1}c' \log n_1}{2} > \sqrt{\log n_1}$$

absorbing sets for S in $\mathcal{E}$. We obtain $\mathcal{E}_{abs}$ by deleting the elements of $\mathcal{E}$ that are not absorbing sets for any k-set S and thus $|\mathcal{E}_{abs}| \leq |\mathcal{E}| \leq c' \log n_1$. ◀

8 Proof of Theorem 9

Now we are ready to prove Theorem 9. Let H be an n-vertex k-graph, and let $\mathcal{P}$ be a partition given by Lemma 8 satisfying (1)-(5). We first prove the forward implication.

8.1 Proof of the forward implication of Theorem 9

If H contains a perfect matching M, then $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M)) = \mathbf{0} \in L^{\mu}_{\mathcal{P}}(H)$. Let M' be the smallest submatching of M that covers V_0, so $|M'| \le |V_0|$. We shall show that there exists a matching $M'' \subset (M \setminus M')$ such that $|M''| \le q$ and $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M' \cup M'')) \in L^{\mu}_{\mathcal{P}}(H)$, implying that $(\mathcal{P}, L^{\mu}_{\mathcal{P}}(H))$ is (V_0, q)-soluble.

Indeed, suppose that $M'' \subset (M \setminus M')$ is a smallest matching such that $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M' \cup M'')) \in L^{\mu}_{\mathcal{P}}(H)$ and $|M''| = m \ge q$. Let $M'' = \{e_1, \dots, e_m\}$ and consider the $m+1$ partial sums

$$\sum_{i=1}^{j} \mathbf{i}_{\mathcal{P}}(e_i) + L^{\mu}_{\mathcal{P}}(H),$$

for $j = 0, 1, \dots, m$. Since $|Q(\mathcal{P}, L^{\mu}_{\mathcal{P}}(H))| \le q \le m$, two of the sums must be in the same coset. That is, there exist $0 \le j_1 < j_2 \le m$ such that

$$\sum_{i=j_1+1}^{j_2} \mathbf{i}_{\mathcal{P}}(e_i) \in L^{\mu}_{\mathcal{P}}(H).$$

So the matching $M^* := M' \setminus \{e_{j_1+1}, \dots, e_{j_2}\}$ satisfies that $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M^* \cup M')) \in L^{\mu}_{\mathcal{P}}(H)$ and $|M^*| < |M''|$, a contradiction.

8.2 Proof of the backward implication of Theorem 9

We first introduce the following useful constant. Given a set I of k-vectors in $\mathbb{Z}^r$, and $m \in \mathbb{N}$, consider the set J of all m'-vectors that are in the lattice in $\mathbb{Z}^r$ generated by I with $0 \le m' \le m$. That is, for any $\mathbf{v} \in J$, there exist $a_{\mathbf{i}} \in \mathbb{Z}$, $\mathbf{i} \in I$ such that

$$\mathbf{v} = \sum_{\mathbf{i} \in I} a_{\mathbf{i}} \mathbf{i}$$

Then let $C^* := C^*(r, k, I, m)$ be the maximum of $|a_{\mathbf{i}}|$ over all such $\mathbf{v}$. Furthermore, let $C_{\max} := C_{\max}(k, m)$ be the maximum of $C^* = C^*(r, k, I, m)$ over all $r \le r_0(k) := 2^{\binom{2k-1}{k-1}} + k$ and all families of k-vectors $I \subseteq \mathbb{Z}^r$.

Now we start the proof. Recall that $c^*_{k,\ell} \ge c^*_{k,k-1} = 1/k$. Then $\lfloor 1/c^*_{k,\ell} \rfloor \le k$. Define constants

$$t := 2^k \quad \text{and} \quad C := C_{\max}(k, kq + k).$$

Define an additional constant $c' > 0$ so that

$$1/n_0 \ll 1/c' \ll \beta, \mu \ll \delta' \ll 1/k, 1/q, 1/C, 1/t.$$

Let $n \ge n_0$ be a multiple of k. Let H be as in the statement of the theorem and $\mathcal{P}$ be a partition of $V(H)$ satisfying Lemma 8 (1)-(5), where the C therein is as defined above. In particular, Property (5) and the choice of t imply that for $s+1 \le i \le r$, $|V_i| \ge \delta' n/2$ and V_i is (β, t)-closed in $H[\bigcup_{s+1 \le i \le r} V_i]$. Furthermore, assume that $(\mathcal{P}, L^{\mu}_{\mathcal{P}}(H))$ is (V_0, q)-soluble, that is, there is a matching M_1 of size at most $|V_0| + q$ such that M_1 covers V_0 and it is a $(|V_0| + q)$-solution, that is,

$$\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M_1)) \in L^{\mu}_{\mathcal{P}}(H).$$

Let $n_1 := |\bigcup_{s+1\le i\le r} V_i|$. We first apply Lemma 14 to H and get a family $\mathcal{E}_{abs}$ consisting of at most $c'\log n_1$ disjoint tk^2-sets such that every k-set S of vertices with $\mathbf{i}_{\mathcal{P}}(S)\in I^{\mu}_{\mathcal{P}}(H)$ has at least $\sqrt{\log n_1}$ absorbing tk^2-sets in $\mathcal{E}_{abs}$.

Note that $V(M_1)$ may intersect $V(\mathcal{E}_{abs})$ in at most $(|V_0|+q)k$ absorbing sets of $\mathcal{E}_{abs}$. Let $\mathcal{E}_0$ be the subfamily of $\mathcal{E}_{abs}$ obtained from removing the tk^2-sets that intersect $V(M_1)$. Let M_0 be the perfect matching on $V(\mathcal{E}_0)$ that is the union of the perfect matchings on each member of $\mathcal{E}_0$. Note that every k-set S with $\mathbf{i}_{\mathcal{P}}(S)\in I^{\mu}_{\mathcal{P}}(H)$ has at least $\sqrt{\log n_1}-(|V_0|+q)k$ absorbing sets in $\mathcal{E}_0$.

Next we want to "store" some disjoint edges for each k-vector in $I^{\mu}_{\mathcal{P}}(H)$ for later steps, and at the same time we also cover the rest vertices of $\bigcup_{1\le i\le s} V_i$ (recall that V_0 is covered by M_1). More precisely, set $C' := C^*(r,k,I^{\mu}_{\mathcal{P}}(H),kq+k)\le C$. Note that Lemma 8 (3) guarantees that for each $i\in[s]$, V_i has at least $\binom{2k-2}{k-1}C$ uncovered vertices. We construct a matching M_2 in $H\setminus V(M_0\cup M_1)$ which consists of C' disjoint edges e with $\mathbf{i}_{\mathcal{P}}(e)=\mathbf{i}$ for every $\mathbf{i}\in I^{\mu}_{\mathcal{P}}(H)$ and also cover the rest vertices of $\bigcup_{1\le i\le s} V_i$ by μ-robust edges. So

$$|M_2|\le \binom{k+r-1}{k}C' + \left|\bigcup_{1\le i\le s} V_i\right|.$$

Note that the process is possible because H contains at least μn^k edges for each $\mathbf{i}\in I^{\mu}_{\mathcal{P},1}(H)$ and every vertex in $\bigcup_{1\le i\le s} V_i$ is in at least μn^{k-1} edges for $\mathbf{i}\in I^{\mu}_{\mathcal{P},2}(H)$ and

$$|V(M_0\cup M_1\cup M_2)|\le tk^2c'\log n_1+(|V_0|+q)k+\left(\binom{k+r-1}{k}C' + \left|\bigcup_{1\le i\le s} V_i\right|\right)k<\mu n_1<\mu n, \tag{2}$$

which allow us to choose desired edges in a greedy manner. Moreover, for every $i\in[s]$, the number of μ-robust index vectors $\mathbf{i}$ such that $\mathbf{i}|_i=1$ is at most $\binom{2k-2}{k-1}$, and thus the process above needs at most $\binom{2k-2}{k-1}C$ uncovered vertices from V_i, which is okay by our construction[3].

Let $H' := H\setminus V(M_0\cup M_1\cup M_2)$ and $n' := |H'|$. So $n'\ge n-\mu n$ and by $\delta(k,\ell,k)\le c^*_{k,\ell}$ due to Theorem 11,

$$\delta_\ell(H')\ge \delta_\ell(H)-\mu n^{k-\ell}\ge (\delta(k,\ell,k)+\gamma/2)\binom{n'-\ell}{k-\ell}.$$

By the definition of $\delta(k,\ell,k)$, we have a matching M_3 in H covering all but at most k vertices. Let U be the set of vertices in H' uncovered by M_3. We are done if $U=\emptyset$. Otherwise because k divides n we have $|U|=k$.

We write $Q := Q(\mathcal{P},L^{\mu}_{\mathcal{P}}(H))$ for brevity. Recall that $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M_1))\in L^{\mu}_{\mathcal{P}}(H)$. Note that by definition, the index vectors of all edges in M_2 are in $I^{\mu}_{\mathcal{P}}(H)$. So we have $\mathbf{i}_{\mathcal{P}}(V(H)\setminus V(M_1\cup M_2))\in L^{\mu}_{\mathcal{P}}(H)$, namely, $R_Q(V(H)\setminus V(M_1\cup M_2))=\mathbf{0}+L^{\mu}_{\mathcal{P}}(H)$. Thus,

$$\sum_{e\in M_0\cup M_3} R_Q(e)+R_Q(U)=\mathbf{0}+L^{\mu}_{\mathcal{P}}(H).$$

Suppose $R_Q(U)=\mathbf{v}_0+L^{\mu}_{\mathcal{P}}(H)$ for some $\mathbf{v}_0\in L^d_{\max}$; so

$$\sum_{e\in M_0\cup M_3} R_Q(e)=-\mathbf{v}_0+L^{\mu}_{\mathcal{P}}(H).$$

[3] Remark. This is where we need Lemma 8 (3), the lower bound of $|V_i|$, $i\in[s]$.

We use the following claim proved in [13] (earlier versions appeared in [11, 18]). Its proof is via the coset arguments and is very similar to the one used in the proof of the forward implication.

▷ Claim 16 ([13, Claim 5.1]). There exist $e_1, \ldots, e_p \in M_0 \cup M_3$ for some $p \le q-1$ such that

$$\sum_{i\in[p]} R_Q(e_i) = -\mathbf{v}_0 + L^{\mu}_{\mathcal{P}}(H). \tag{3}$$

That is, we have $\sum_{i\in[p]} \mathbf{i}_{\mathcal{P}}(e_i) + \mathbf{i}_{\mathcal{P}}(U) \in L^{\mu}_{\mathcal{P}}(H)$. Let $Y := \bigcup_{i\in[p]} e_i \cup U$ and thus $|Y| = pk + k \le qk + k$. We now complete the perfect matching by absorption. Since $\mathbf{i}_{\mathcal{P}}(Y) \in L^{\mu}_{\mathcal{P}}(H)$, we have the following equation

$$\mathbf{i}_{\mathcal{P}}(Y) = \sum_{\mathbf{v}\in I^{\mu}_{\mathcal{P}}(H)} a_{\mathbf{v}}\mathbf{v},$$

where $a_{\mathbf{v}} \in \mathbb{Z}$ for all $\mathbf{v} \in I^{\mu}_{\mathcal{P}}(H)$. Since $|Y| \le qk+k$, by the definition of C', we have $|a_{\mathbf{v}}| \le C'$ for all $\mathbf{v} \in I^{\mu}_{\mathcal{P}}(H)$. Noticing that $a_{\mathbf{v}}$ may be negative, we can assume $a_{\mathbf{v}} = b_{\mathbf{v}} - c_{\mathbf{v}}$ such that one of $b_{\mathbf{v}}, c_{\mathbf{v}}$ is $|a_{\mathbf{v}}|$ and the other is zero for all $\mathbf{v} \in I^{\mu}_{\mathcal{P}}(H)$. So we have

$$\sum_{\mathbf{v}\in I^{\mu}_{\mathcal{P}}(H)} c_{\mathbf{v}}\mathbf{v} + \mathbf{i}_{\mathcal{P}}(Y) = \sum_{\mathbf{v}\in I^{\mu}_{\mathcal{P}}(H)} b_{\mathbf{v}}\mathbf{v}.$$

This equation means that given a family $\mathcal{E} = \{W^{\mathbf{v}}_1, \ldots, W^{\mathbf{v}}_{c_{\mathbf{v}}} : \mathbf{v} \in I^{\mu}_{\mathcal{P}}(H)\}$ of disjoint k-subsets of $V(H) \setminus Y$ such that $\mathbf{i}_{\mathcal{P}}(W^{\mathbf{v}}_i) = \mathbf{v}$ for all $i \in [c_{\mathbf{v}}]$, we can regard $V(\mathcal{E}) \cup Y$ as the union of disjoint k-sets $\{S^{\mathbf{v}}_1, \ldots, S^{\mathbf{v}}_{b_{\mathbf{v}}} : \mathbf{v} \in I^{\mu}_{\mathcal{P}}(H)\}$ such that $\mathbf{i}_{\mathcal{P}}(S^{\mathbf{v}}_j) = \mathbf{v}$, $j \in [b_{\mathbf{v}}]$ for all $\mathbf{v} \in I^{\mu}_{\mathcal{P}}(H)$. Since $c_{\mathbf{v}} \le C'$ for all $\mathbf{v}$ and $V(M_2) \cap Y = \emptyset$, we can choose the family $\mathcal{E}$ as a subset of M_2. In summary, starting with the matching $M_0 \cup M_1 \cup M_2 \cup M_3$ leaving U uncovered, we delete the edges $e_1, \ldots, e_p$ from $M_0 \cup M_3$ given by Claim 16 and then leave $Y = \bigcup_{i\in[p]} V(e_i) \cup U$ uncovered. Next we delete the family $\mathcal{E}$ of edges from M_2 and leave $V(\mathcal{E}) \cup Y$ uncovered. Finally, we regard $V(\mathcal{E}) \cup Y$ as the union of at most

$$|M_2| + qk + k \le \sqrt{\log n_1}/2$$

k-sets S each with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$.

Note that by definition, Y may intersect at most $qk + k$ absorbing sets in $\mathcal{E}_0$, which cannot be used to absorb those sets we obtained above. Since each k-set S has at least $\sqrt{\log n_1} - (|V_0| + q)k > \sqrt{\log n_1}/2 + qk + k$ absorbing tk^2-sets in $\mathcal{E}_0$, we can greedily match each S with a distinct absorbing tk^2-set $E_S \in \mathcal{E}_0$ for S. Replacing the matching on $V(E_S)$ in M_0 by the perfect matching on $H[E_S \cup S]$ for each S gives a perfect matching in H.

References

1 Noga Alon, Peter Frankl, Hao Huang, Vojtech Rödl, Andrzej Ruciński, and Benny Sudakov. Large matchings in uniform hypergraphs and the conjecture of Erdős and Samuels. *J. Combin. Theory Ser. A*, 119(6):1200–1215, 2012. doi:10.1016/j.jcta.2012.02.004.

2 Arash Asadpour, Uriel Feige, and Amin Saber. Santa claus meets hypergraph matchings. *Approximation, randomization and combinatorial optimization, Lecture Notes in Comput. Sci., vol. 5171, Springer, Berlin*, pages 10–20, 2008.

3 Yulin Chang, Huifen Ge, Jie Han, and Guanghui Wang. Matching of given sizes in hypergraphs. *ArXiv eprint: 2106.16068*, 2021.

4 Laihao Ding, Jie Han, Shumin Sun, Guanghui Wang, and Wenling Zhou. *F*-factors in quasirandom hypergraphs. *J. London Math. Soc., to appear*, 2022.

5 Jack Edmonds. Paths, trees, and flowers. *Canad. J. Math.*, 17:449–467, 1965.

6 Peter Frankl and Andrey Kupavskii. The Erdős Matching Conjecture and concentration inequalities. *ArXiv e-prints*, 2018. arXiv:1806.08855.

7 Frederik Garbe and Richard Mycroft. Hamilton cycles in hypergraphs below the Dirac threshold. *J. Combin. Theory Ser. B*, 133:153–210, 2018. doi:10.1016/j.jctb.2018.04.010.

8 Stefan Glock, Daniela Kühn, Allan Lo, and Deryk Osthus. The existence of designs via iterative absorption: hypergraph F-designs for arbitrary F. *Mem. Amer. Math. Soc., accepted*, 2020.

9 Hiep Hàn, Yury Person, and Mathias Schacht. On perfect matchings in uniform hypergraphs with large minimum vertex degree. *SIAM J. Discrete Math*, 23:732–748, 2009.

10 Jie Han. Perfect matchings in hypergraphs and the Erdős matching conjecture. *SIAM J. Discrete Math.*, 30(3):1351–1357, 2016. doi:10.1137/16M1056079.

11 Jie Han. Decision problem for perfect matchings in dense k-uniform hypergraphs. *Trans. Amer. Math. Soc.*, 369(7):5197–5218, 2017.

12 Jie Han and Peter Keevash. Finding perfect matchings in dense hypergraphs. In *Proceedings of the 2020 ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2366–2377, 2020. doi:10.1137/1.9781611975994.145.

13 Jie Han and Andrew Treglown. The complexity of perfect matchings and packings in dense hypergraphs. *J. Combin. Theory Ser. B*, 141:72–104, 2020. doi:10.1016/j.jctb.2019.06.004.

14 Richard M. Karp. Reducibility among combinatorial problems. In *Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972)*, pages 85–103. Plenum, New York, 1972.

15 Marek Karpiński, Andrzej Ruciński, and Edyta Szymańska. Computational complexity of the perfect matching problem in hypergraphs with subcritical density. *Internat. J. Found. Comput. Sci.*, 21(6):905–924, 2010. doi:10.1142/S0129054110007635.

16 Peter Keevash. The existence of designs. *preprint*, 2014.

17 Peter Keevash, Fiachra Knox, and Richard Mycroft. Polynomial-time perfect matchings in dense hypergraphs. *Proceedings of the 45th STOC (2013)*, 2013.

18 Peter Keevash, Fiachra Knox, and Richard Mycroft. Polynomial-time perfect matchings in dense hypergraphs. *Adv. Math.*, 269:265–334, 2015. doi:10.1016/j.aim.2014.10.009.

19 Daniela Kühn, Deryk Osthus, and Timothy Townsend. Fractional and integer matchings in uniform hypergraphs. *European J. Combin.*, 38:83–96, 2014. doi:10.1016/j.ejc.2013.11.006.

20 Allan Lo and Klas Markström. F-factors in hypergraphs via absorption. *Graphs Combin.*, 31(3):679–712, 2015. doi:10.1007/s00373-014-1410-8.

21 Vojtech Rödl, Andrzej Ruciński, and Endre Szemerédi. An approximate Dirac-type theorem for k-uniform hypergraphs. *Combinatorica*, 28(2):229–260, 2008.

22 Edyta Szymańska. The complexity of almost perfect matchings and other packing problems in uniform hypergraphs with high codegree. *European J. Combin.*, 34(3):632–646, 2013. doi:10.1016/j.ejc.2011.12.009.

23 Raphael Yuster. Combinatorial and computational aspects of graph packing and graph decomposition. *Computer Science Review*, 1(1):12–26, 2007. doi:10.1016/j.cosrev.2007.07.002.