Finding Perfect Matchings in Dense Hypergraphs[*]

Jie Han[†] Peter Keevash[‡]

Abstract

We show that for any integers $k \geq 3$ and $c \geq 0$ there is a polynomial-time algorithm, that given any n-vertex k-uniform hypergraph H with minimum codegree at least $n/k - c$, finds either a perfect matching in H or a certificate that no perfect matching exists.

1 Introduction

Matchings are fundamental objects in Graph Theory and have broad applications in other branches of Science and a variety of practical problems (e.g. the assignment of graduating medical students to their first hospital appointments[1]). Applications of matchings in hypergraphs include the 'Santa Claus' allocation problem [2]; they also offer a universal framework for many important combinatorial problems, e.g. the Existence Conjecture for designs (see [7, 17]) and Ryser's conjecture [29] on transversals in Latin squares.

This paper is concerned with the algorithmic question of finding a matching that is *perfect*, meaning that it covers all vertices. The graph case of this question is well understood: Tutte's Theorem [34] gives necessary and sufficient conditions for a graph to contain a perfect matching, and Edmonds' Algorithm [4] finds such a matching in polynomial time. However, for hypergraphs it is a different story: in fact, determining whether a 3-uniform hypergraph contains a perfect matching was one of Karp's celebrated 21 NP-complete problems [12]. As the general problem is intractable (assuming P $\neq$ NP), it is natural to seek conditions that guarantee a perfect matching, or at least make the existence question tractable.

1.1 Perfect matchings under minimum degree conditions We start with some definitions used throughout the paper. Given $k \geq 2$, a k-uniform hypergraph (in short, *k-graph*) $H = (V, E)$ consists of a vertex set V and an edge set $E \subseteq \binom{V}{k}$, where every edge is a k-element subset of V. A *matching* in H is a collection of vertex-disjoint edges of H. A *perfect matching* M in H is a matching that covers all vertices of H. We always assume that k divides $n := |V(H)|$, which is clearly a necessary condition for the existence of a perfect matching in H. For $S \subseteq V(H)$ the *neighbourhood* of S is $N_H(S) := \{T \subseteq V(H) \setminus S : S \cup T \in E(H)\}$, and the *degree* of S is $\deg_H(S) = |N_H(S)|$; the subscript H is omitted if it is clear from the context. The *minimum d-degree* $\delta_d(H)$ of H is the minimum of $\deg_H(S)$ over all d-vertex sets S in H. We refer to $\delta_{k-1}(H)$ as the *minimum codegree* of H.

Rödl, Ruciński and Szemerédi [28] determined the sharp minimum codegree condition to ensure a perfect matching in an n-vertex k-graph for large n and all $k \geq 3$: the extremal examples are as follows.

Construction 1. *Given $n \geq k \geq 3$, let $\mathcal{H}_{n,k}$ be the collection of k-graphs H such that there is a partition of $V(H) = X \cup Y$ such that $n/k - |X|$ is odd and all edges of H intersect X in an odd number of vertices.*

Any matching M in $H \in \mathcal{H}_{n,k}$ cannot be perfect; indeed, if it were we would have $\sum_{e \in M}(|e \cap X| - 1) = |X| - n/k$, which is impossible, as the left hand side is even but the right hand side is odd. There is a large literature on minimum degree conditions that force a perfect matching, see [1, 3, 5, 8, 18–21, 24, 25, 27, 28, 31–33] and the surveys [26, 35], yet there are still many open problems, such as determining the minimum 1-degree condition that forces a perfect matching.

1.2 Algorithms Let $\mathbf{DPM}_k(n, m)$ be the decision problem of determining whether an n-vertex k-graph H with $\delta_{k-1}(H) \geq m$ contains a perfect matching. When can $\mathbf{DPM}_k(n, m)$ be decided in polynonmial time? The result of [28] mentioned above shows that the decision problem is trivial for $m \geq n/2 - k + 3$ (then there is a perfect matching iff $k \mid n$). Szymańska [30] proved that if $\delta < 1/k$ then $\mathbf{DPM}_k(n, \delta n)$ is polynomial-time reducible to $\mathbf{DPM}_k(n, 0)$, and so NP-complete. Karpiński, Ruciński and Szymańska [14] showed that there exists $\varepsilon > 0$ such that $\mathbf{DPM}_k(n, (1/2 - \varepsilon)n)$ is in P

[*]JH is partially supported by the Simons Collaboration Grant for Mathematicians 630884. PK is supported in part by ERC Consolidator Grant 647678.

[†]Department of Mathematics, University of Rhode Island, Kingston, RI, 02881, USA. Email: jie_han@uri.edu

[‡]Mathematical Institute, University of Oxford, Woodstock Road, Oxford, OX2 6GG, UK. Email: keevash@maths.ox.ac.uk

[1]In 2012, the Nobel Memorial Prize in Economics was awarded to Shapley and Roth "for the theory of stable allocations and the practice of market design."

and posed the question of determining for the complexity of $\mathbf{DPM}_k(n, \delta n)$ for $\delta \in [1/k, 1/2)$.

This was resolved for $\delta > 1/k$ by Keevash, Knox and Mycroft [15, 16], who showed that the problem is in P and also gave a polynomial-time algorithm that either finds a perfect matching or a certificate that none exists. Their proof was very long and technical, and left open the 'threshold case' $\delta = 1/k$, which poses additional challenges (discussed in Section 2). Han [10] completely resolved the question of [14] by showing that $\mathbf{DPM}_k(n, \delta n)$ is in P for $\delta \in [1/k, 1/2)$. His proof was much simpler than that in [16] (it relied on some theory from [16] and also developed a lattice-based absorbing method which has found many other applications) but his result only concerned the decision problem, and left open the constructive problem, i.e. finding a perfect matching or a certificate that none exists.

Main result We give a short proof of the following even stronger result.

THEOREM 1.1. *Let $k \geq 3$ and H be an n-vertex k-graph with $\delta_{k-1}(H) \geq n/k - c$ for some $c = o(n)$. Then there is an algorithm running in time $O(n^{k \cdot \max\{4^k, c\}})$ that finds a perfect matching in H or a certificate that none exists. In particular, if c is constant then $\boldsymbol{DPM}_k(n, n/k - c)$ is in P.*

2 Overview of the paper

In the next section we describe the algorithm referred to in Theorem 1.1 and reduce its proof of correctness to two theorems (proved in the two subsequent sections) that respectively handle the 'non-extremal' and 'extremal' cases for H. To explain this distinction, we consider the following construction that appears naturally around the codegree threshold n/k.

CONSTRUCTION 2. (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 on V whose edges are all k-sets that intersect S.*

We note that the minimum codegree of H is $|S|$ and any matching in H has at most $|S| < n/k$ edges, so cannot be perfect. On the other hand, Han [9] showed that any n-vertex k-graph with $\delta_{k-1}(H) \geq n/k-1$ contains a matching of size $n/k - 1$, thus determining the tight codegree condition for a matching that is just one edge short of being perfect. This rather surprising phenomenon indicates that the key issue for whether there is a perfect matching near the codegree threshold n/k is whether H is close to a space barrier, or equivalently, whether H has an independent set of size about $n - n/k$: we say H is *ε-extremal* if it has an independent set of size $(1 - \varepsilon)\frac{k-1}{k}n$.

Our strategy of separating the non-extremal and extremal cases follows that of Han [10], and indeed several aspects of his proof carry over to our setting (despite our weaker assumption $\delta_{k-1}(H) \geq n/k - c$, which is significantly more challenging to work with, as it does not rule out space barriers). However, the crucial difficulty that prevented Han from finding a perfect matching (as opposed to just testing for its existence) is the algorithmic intractability of finding (or even approximating) the largest independent set (this is NP-hard even in graphs, let alone hypergraphs).

The starting point for the new approach in this paper is to observe that this issue can be avoided via an algorithm of Han [9] that either finds an almost perfect matching or a large independent set. In the latter case, we can algorithmically extend to a large *maximal* independent set, which may not be of maximum size, but nevertheless gives us enough power to analyse the extremal case with some additional arguments. A final contribution of this paper (see the concluding remarks) is an algorithmic reduction of the perfect matching problem (see Problem 1) of independent interest.

3 The algorithm

In this section we state our algorithm and prove our main theorem assuming two theorems (concerning the non-extremal and extremal cases) whose proofs will be given in the following two sections.

The first ingredient of our algorithm is Procedure ListPartitions from [16, Section 2], which can test[2] in time $O(n^{k+1})$ whether $H \in \mathcal{H}_{n,k}$ (defined in Construction 1). If $H \in \mathcal{H}_{n,k}$ then the procedure finds a partition certifying that H does not have a perfect matching, so we reduce to the case $H \notin \mathcal{H}_{n,k}$.

Next we require the following definitions from [16].

DEFINITION 1. *Let $H = (V, E)$ be a k-graph and $\mathcal{P} = \{V_1, V_2, \ldots, V_d\}$ be a partition of V. The index vector of a set $S \subseteq V$ is $\mathbf{i}_{\mathcal{P}}(S) = (|S \cap V_1|, \ldots, |S \cap V_d|) \in \mathbb{Z}^d$. Given $\mu > 0$, let $I^{\mu}_{\mathcal{P}}(H)$ denote the set of all k-vectors $\mathbf{i} \in \mathbb{Z}^d$ such that at least $\mu|V|^k$ edges $e \in H$ have $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$, and let $L^{\mu}_{\mathcal{P}}(H)$ denote the lattice in $\mathbb{Z}^d$ generated by $I^{\mu}_{\mathcal{P}}(H)$.*

The intuition behind $L^{\mu}_{\mathcal{P}}(H)$ is that it captures the robust 'divisibility' constraints in H (generalising the parity condition in Construction 1) and that H should not have any divisibility obstruction to a perfect matching if it is possible to delete a small matching so that the index vector of the uncovered set is in $L^{\mu}_{\mathcal{P}}(H)$.

We will also use the reachability methods introduced by Lo and Markström [22, 23].

[2] See Lemma [16, Section 2] and note that the proof is valid assuming $\delta_{k-1}(H) \geq \omega n$ for any fixed ω and $n > n_0(k, \omega)$.

Definition 2. *Let H be an n-vertex k-graph. We say that two vertices u and v 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 say that $U \subseteq V(H)$ is* (β, i)-closed *in H if any two vertices $u, v \in U$ are (β, i)-reachable in H.*

The following statement is a simplified form of [10, Lemma 2.5]. Throughout this paper, $x \ll y$ means that for any $y > 0$ there exists $x_0 > 0$ such that for any $x < x_0$ the subsequent statement holds.

Lemma 3.1. *Let $1/n \ll \beta, \mu \ll \gamma \ll 1/k$, where $k \geq 3$ is an integer. Then for each k-graph H on n vertices with $\delta_{k-1}(H) \geq n/k - \gamma n$, we can find in time $O(n^{2^{k-1}k+1})$ a partition $\mathcal{P} = \{V_1, \ldots, V_d\}$ such that each V_i is $(\beta, 2^{k-1})$-closed in H and has $|V_i| \geq n/k - 2\gamma n$.*

The following theorem handles the non-extremal case: given the partition $\mathcal{P}$ from Lemma 3.1, it provides an algorithm that finds a perfect matching or outputs an independent set witnessing that H is in the extremal case.

Theorem 3.1. *Suppose $1/n \ll \beta, \mu \ll \gamma \ll 1/k$. Let H be a k-graph on n vertices such that $\delta_{k-1}(H) \geq n/k - \gamma n$. Suppose $\mathcal{P}$ is a partition of $V(H)$ found by Lemma 3.1 such that there is a matching M_1 with $|M_1| \leq k$ and $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M_1)) \in L^{\mu}_{\mathcal{P}}(H)$. Then there is an algorithm that finds in time $O(n^{2^{k+1}k^2})$ a perfect matching in H or an independent set in H of size $(1 - 5k\gamma)\frac{k-1}{k}n$.*

The extremal case is handled by the following theorem (recall that H is ε-extremal if $V(H)$ contains an independent subset of size at least $(1-\varepsilon)\frac{k-1}{k}n$).

Theorem 3.2. *Assume $1/n \ll \varepsilon \ll 1/k$ and $c \leq \varepsilon n/k$. Let H be a k-graph on n vertices such that $\delta_{k-1}(H) \geq n/k - c$. Suppose $H \notin \mathcal{H}_{n,k}$ and H is ε-extremal. Then there is an algorithm that finds in time $O(n^{\max\{ck, 4(k-1)\}})$ a perfect matching in H or a certificate that none exists.*

Now we are ready to state our main algorithm.

Procedure PerfectMatching

Data : an n-vertex k-graph H such that $\delta_{k-1}(H) \geq n/k - c$.
Output : a perfect matching in H or a certificate that none exists.

```
if Procedure ListPartitions finds that H ∈ H_{n,k} then
    Output the certifying partition and halt.
Choose constants 1/n ≪ β, μ ≪ γ ≪ 1/k ;
Apply Lemma 3.1 to find a partition P of V(H) and L^μ_P(H) ;
if there is a matching M_1 of size at most k with i_P(V(H) ∖ V(M_1)) ∈ L^μ_P(H) then
    if Theorem 3.1 finds a perfect matching M then
        Output M and halt.
    else
        Apply Theorem 3.2 with ε = 5kγ and halt with appropriate output.
else
    Output P as a certificate of "no perfect matching" and halt.
```

We conclude this section by showing correctness of the above algorithm, thus proving Theorem 1.1, assuming Theorems 3.1 and 3.2, which will be proved in the following two sections. Suppose L is an edge-lattice in $\mathbb{Z}^{|\mathcal{P}|}$, where $\mathcal{P}$ is a partition of a set V, then the *coset group* of $(\mathcal{P}, L)$ is $G = G(\mathcal{P}, L) = L^{|\mathcal{P}|}_{\max}/L$, where $L^d_{\max} = \{x \in \mathbb{Z}^d : k \text{ divides } \sum_{i \in [d]} x_i\}$.

Proof. [Proof of Theorem 1.1] We show correctness of Procedure PerfectMatching. As discussed above, Procedure ListPartitions from [16, Section 2], tests in time $O(n^{k+1})$ whether $H \in \mathcal{H}_{n,k}$ and if so outputs a partition certifying that H does not have a perfect matching. Thus we can assume $H \notin \mathcal{H}_{n,k}$.

Next we apply Lemma 3.1 which finds $\mathcal{P} = (V_1, \ldots, V_d)$ in time $O(n^{2^{k-1}k+1})$; we note that $d \leq k$. Then we test each set of at most k edges in H (of which there are $O(n^{k^2})$) to see if it is a matching M_1 satisfying $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M_1)) \in L^{\mu}_{\mathcal{P}}(H)$. If we find any such M_1 then we use it to apply Theorem 3.1, which finds a perfect matching or an independent set that can be used to apply Theorem 3.2, which in turn finds a perfect matching or a certificate that none exists. We only need to apply Theorems 3.1 and 3.2 for one such M_1 if it exists, so the running time is

$$O(\max\{n^{2^{k+1}k^2}, n^{\max\{ck, 4(k-1)\}}\}) = O(n^{k \cdot \max\{4^k, c\}}).$$

To complete the proof, it remains to show that if there is a perfect maching M then some such M_1 exists. To see this we argue similarly to [16]. First we note the following property of $I^{\mu}_{\mathcal{P}}(H)$ that follows easily from the minimum codegree condition of H: for any $\mathbf{v} \in \mathbb{Z}^d$ with non-negative coordinates summing to $k-1$ there is some $i \in [d]$ such that $\mathbf{v} + \mathbf{u}_i \in I^{\mu}_{\mathcal{P}}(H)$. Indeed, consider all $(k-1)$-sets with index vector $\mathbf{v}$, by the minimum codegree condition H contains at least $(n/k - 2\gamma n - k)^{k-1} \cdot \delta_{k-1}(H)/k! \geq (1/k - 3\gamma)^k n^k / k!$ edges whose index vector equals $\mathbf{v} + \mathbf{u}_i$ for $i \in [d]$; then by averaging, there exists $i \in [d]$ such that the number of edges with index vector $\mathbf{v} + \mathbf{u}_i$ is at least μn^k. Then the proof of [16, Lemma 6.4] shows that the coset group G of $L^{\mu}_{\mathcal{P}}(H)$ in the lattice $\{\mathbf{v} \in \mathbb{Z}^d : k \mid \sum v_i\}$ has size $|G| \le d \le k$. Now we repeatedly apply the pigeonhole principle to reduce M to M_1 as in [16, Proposition 6.10]. We start with $M_1 = M$ and note that as M is perfect we have $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M)) = \mathbf{0} \in L^{\mu}_{\mathcal{P}}(H)$. While $|M_1| > k$ we consider any edges $e_1, \dots, e_k$ in M_1 and the partial sums $\sum_{j=1}^{i} \mathbf{i}_{\mathcal{P}}(e_j)$ for $0 \le i \le k$. By the pigeonhole principle, some two of these sums lie in the same coset of $L^{\mu}_{\mathcal{P}}(H)$, namely, there exist $0 \le i_1 < i_2 \le k$ such that

$$\sum_{j=i_1+1}^{i_2} \mathbf{i}_{\mathcal{P}}(e_j) = \sum_{j=1}^{i_2} \mathbf{i}_{\mathcal{P}}(e_j) - \sum_{j=1}^{i_1} \mathbf{i}_{\mathcal{P}}(e_j) \in L^{\mu}_{\mathcal{P}}(H).$$

So we can delete $e_{i_1+1}, \dots, e_{i_2}$ from M_1 while preserving $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M_1)) \in L^{\mu}_{\mathcal{P}}(H)$. We terminate with $|M_1| \le k$, as required.

4 The non-extremal case

In this section we prove Theorem 3.1, which finds a perfect matching or a large independent set, thus establishing correctness of Procedure PerfectMatching in the non-extremal case. We adapt (and simplify) the approach of Han [10] via lattice-based absorption and also incorporate a derandomisation argument of Garbe and Mycroft [6] so that we can find a perfect matching (not just test for its existence).

4.1 Almost perfect matching or large independent set The key idea of proof via absorption is that it simplifies the problem of finding a perfect matching to that of finding an almost perfect matching. Accordingly, we start by showing how to find an almost perfect matching or a large independent set. The following lemma is essentially [9, Lemma 1.6]; the proof is algorithmic, although this is not made explicit, so for the convenience of the reader we do this here.

Lemma 4.1. *Suppose that $1/n \ll \gamma \ll 1/k$ and $k \mid n$. Let H be a k-graph on n vertices with $\delta_{k-1}(H) \ge n/k - \gamma n$. Then in time $O(n^{k+1})$ we can find either a matching that leaves at most k^2/γ vertices uncovered, or an independent set of size $(1 - 2k\gamma)\frac{k-1}{k}n$.*

Proof. Consider any matching $M = \{e_1, e_2, \dots, e_m\}$ in H. Let V' be the set of vertices covered by M and let U be the set of vertices which are not covered by M. Assume that $|U| > k^2/\gamma$ and U is an independent set (otherwise we can trivially enlarge M). We will show that we can find either a matching of size $m + 1$, or an independent set of size $(1 - 2k\gamma)\frac{k-1}{k}n$.

We arbitrarily partition all but at most $k - 2$ vertices of U into disjoint $(k-1)$-sets $A_1, \dots, A_t$ where $t = \lfloor \frac{|U|}{k-1} \rfloor > \frac{k}{\gamma}$. Let D be the set of vertices $v \in V'$ such that $\{v\} \cup A_i \in E(H)$ for at least k choices of A_i. First we consider the case that there is some $i^* \in [m]$ with $|e_{i^*} \cap D| \ge 2$. We fix distinct x, y in $e_{i^*} \cap D$ and apply the definition of D to pick distinct A_i, A_j such that $\{x\} \cup A_i$ and $\{y\} \cup A_j$ are edges. Then we can enlarge M by replacing e_{i^*} by $\{x\} \cup A_i$ and $\{y\} \cup A_j$. Thus we may assume each $|e_i \cap D| \le 1$.

Next we show that $|D| \ge (\frac{1}{k} - 2\gamma)n$. As $\delta_{k-1}(H) \ge n/k - \gamma n$ and U is independent,

$$t\left(\frac{1}{k} - \gamma\right)n \le \sum_{i=1}^{t} \deg(A_i) \le |D|t + n \cdot k < t(|D| + \gamma n),$$

so we have

$$|D| > \left(\frac{1}{k} - 2\gamma\right)n.$$

Let $V_D := \bigcup\{e_i : e_i \cap D \neq \varnothing\}$. Note that $|V_D \setminus D| = (k-1)|D| \ge (k-1)(\frac{1}{k} - 2\gamma)n$. If $V_D \setminus D$ is independent then we can output it to conclude the proof. Thus we can assume there is some edge $e_0 \in H[V_D \setminus D]$. Suppose that e_0 intersects $e_{i_1}, \dots, e_{i_l}$ in M for some $l \le k$ and write $\{v_{i_j}\} = e_{i_j} \cap D$ for all $j \in [l]$. By definition of D, we can greedily pick $A_{i_1}, \dots, A_{i_l}$ such that $\{v_{i_j}\} \cup A_{i_j} \in E(H)$ for all $j \in [l]$. Then we can enlarge M by replacing the edges $e_{i_1}, \dots, e_{i_l}$ by e_0 and $\{v_{i_j}\} \cup A_{i_j}$ for $j \in [l]$.

For the running time, note that there are at most n/k iterations in which we enlarge the matching. In each iteration we find D by calculating the degree of each vertex v in time $O(n^{k-1})$. Then we check if $V_D \setminus D$ is an independent set in time $O(n^k)$. The overall running time is $O(n^{k+1})$.

4.2 An absorbing lemma Now we implement the absorption part of the algorithm, Lemma 4.2, by derandomising a lemma from [10]. For the statement, we recall Definitions 1 and 2, and say that a set T *absorbs* or *is absorbing for* a set S if both $H[T]$ and $H[T \cup S]$ contain perfect matchings.

Lemma 4.2. (Absorbing Lemma) *Let $1/n \ll 1/c' \ll \beta, \mu \ll 1/k, 1/t, 1/C'$. Suppose that $\mathcal{P} = \{V_1, \dots, V_d\}$ is a partition of $V(H)$ such that each V_i is (β, t)-closed in H. Then in time $O(n^{4tk^2})$ we can find a family $\mathcal{F}_{abs}$ of at most $c' \log n$ disjoint tk^2-sets such that $H[V(F)]$ contains a perfect matching for all $F \in \mathcal{F}_{abs}$ and every k-vertex set S with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$ has at least C' absorbing tk^2-sets in $\mathcal{F}_{abs}$.*

The proof uses the following lemma which is [10, Claim 3.6].

LEMMA 4.3. *Suppose V_i is (β,t)-closed in H for all $i \in [d]$. Then any k-set S with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$ has at least $\frac{\mu\beta^k}{2^{k+1}} n^{tk^2}$ absorbing tk^2-sets.*

For the derandomisation we use the following lemma of Garbe and Mycroft [6, Proposition 4.7, Procedure SelectSet] (see also Karpiński, Ruciński and Szymańska [13]).

LEMMA 4.4. *Fix constants $\beta > \tau > 0$ and integers m, M, N and $r \le N$ such that r and N are sufficiently large, and that $M \le (1/8)\exp(\tau^2 r/(3\beta))$. Let U and W be disjoint sets of sizes $|U| = M$ and $|W| = N$. Let G be a graph with vertex set $U \cup W$ such that $G[U]$ is empty, $G[W]$ has precisely m edges, and $\deg_G(u) \ge \beta N$ for every $u \in U$. Then in time $O(N^4 + MN^3)$ we can find an independent set $R \subseteq W$ in G such that $(1-\nu)r \le |R| \le r$ and $|N_G(u) \cap R| \ge (\beta - \tau - \nu)r$ for all $u \in U$, where $\nu = 2mr/N^2$.*

We are now ready to prove the absorbing lemma.

Proof. [Proof of Lemma 4.2] We will apply Lemma 4.4 to the graph G with parts $U = \{S : \mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)\}$ and $W = \{T \subseteq V(H) : |T| = tk^2\}$, where T, T' in W are adjacent iff $T \cap T' \ne \varnothing$, and $S \in U$ and $T \in W$ are adjacent iff T absorbs S. In the notation of Lemma 4.4 we have $N = \binom{n}{tk^2}$, $M \le \binom{n}{k}$ and $m = |E(G[W])| \le tk^2\binom{n}{tk^2-1}\binom{n}{tk^2} = t^2k^4N^2/(n - tk^2 + 1)$. We let $\beta' := \mu\beta^k/2^{k+1}$, $\tau := \beta'/3$ and $r = c'\log n$. Then by Lemma 4.3, $\deg_G(u) \ge \beta' N$ for every $u \in U$, and

$$\exp\left(\frac{\tau^2 r}{3\beta'}\right) = \exp\left(\frac{\beta' c' \log n}{27}\right) \ge 8\binom{n}{k} \ge 8M,$$

as n is large enough. Thus, by Lemma 4.4, in time $O(N^4 + MN^3) = O(n^{4tk^2})$, we can find a set $R \subseteq W$ which is independent in G with $(1-\tau)r \le |R| \le r$ and $|N_G(u) \cap R| \ge (\beta' - \tau - \nu)r$ for all $u \in U$, where

$$\nu = \frac{2mr}{N^2} \le \frac{2t^2k^4N^2c'\log n}{(n - tk^2 + 1)N^2} = \frac{2t^2k^4c'\log n}{n - tk^2 + 1} < \beta'/3.$$

Note that R consists of disjoint tk^2-sets of H by definition of $G[W]$. We now remove tk^2-sets in R that do not have a perfect matching, and denote the resulting family of tk^2-sets by $\mathcal{F}_{abs}$. Thus $|\mathcal{F}_{abs}| \le c'\log n$, each member of $\mathcal{F}_{abs}$ has a perfect matching, and every k-vertex set S with $\mathbf{i}_{\mathcal{P}}(S) \in I^{\mu}_{\mathcal{P}}(H)$ has at least $\beta' r/3 \ge \beta' c' \log n/3 \ge C'$ absorbing tk^2-sets in $\mathcal{F}_{abs}$, as n is large enough.

4.3 Proof of Theorem 3.1 We start the proof by fixing a parameter hierarchy $1/n \ll 1/c' \ll \beta, \mu \ll \gamma, 1/C \ll 1/k$, such that C has the following property: for any $I \subseteq \{\mathbf{v} \in \mathbb{Z}^d : \sum_i |v_i| \le k\}$, where $d \le k$, and $\mathbf{u}$ in the integer span of I with $\sum_i |u_i| \le k^2(1+\gamma^{-1})$, we can write

$$\mathbf{u} = \sum_{\mathbf{v} \in I} a_{\mathbf{v}}(\mathbf{u})\mathbf{v}. \tag{4.1}$$

where $a_{\mathbf{v}}(\mathbf{u}) \in \mathbb{Z}$ and $|a_{\mathbf{v}}(\mathbf{u})| \le C$ for each $\mathbf{v} \in I$.

Let H be a k-graph on n vertices such that $\delta_{k-1}(H) \ge n/k - \gamma n$. Let $\mathcal{P} = \{V_1, \dots, V_d\}$ be the partition found by Lemma 3.1; we note that $d \le k$ by Lemma 3.1. Write

$$C' := (C+k+2)k + k/\gamma, \quad I = I^{\mu}_{\mathcal{P}}(H), \quad \text{and } L = L^{\mu}_{\mathcal{P}}(H).$$

Let M_1 be a matching of size at most k such that $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M_1)) \in L$. We first apply Lemma 4.2 to H with $t = 2^{k-1}$, obtaining a family $\mathcal{F}_{abs}$ of $2^{k-1}k^2$-sets with $|\mathcal{F}_{abs}| \le c'\log n$ such that every S with $\mathbf{i}_{\mathcal{P}}(S) \in I$ has at least C' absorbing $2^{k-1}k^2$-sets in $\mathcal{F}_{abs}$. We obtain $\mathcal{F}_0$ from $\mathcal{F}_{abs}$ by removing any sets that intersect $V(M_1)$, of which there are at most $|V(M_1)| \le k^2$; then every S with $\mathbf{i}_{\mathcal{P}}(S) \in I$ has at least $C' - k^2$ absorbing sets in $\mathcal{F}_0$. We let M_0 be a perfect matching on $V(\mathcal{F}_0)$ consisting of perfect matchings on each member of $\mathcal{F}_0$.

Next, we greedily find a matching M_2 in $V(H) \setminus V(M_0 \cup M_1)$ which contains C edges e with $\mathbf{i}_{\mathcal{P}}(e) = \mathbf{i}$ for every $\mathbf{i} \in I$. This is possible because H contains at least μn^k edges for each k-vector $\mathbf{i} \in I$, whereas $|V(M_0 \cup M_1 \cup M_2)| \le 2^{k-1}k^2c'\log n + k^2 + kC\binom{k+d-1}{k} < \mu n$, as n is large enough.

Let $H' := H[V(H) \setminus V(M_0 \cup M_1 \cup M_2)]$. Note that $|V(H')| \ge n - \mu n$ as n is large, and

$$\delta_{k-1}(H') \ge \delta_{k-1}(H) - \mu n \ge n/k - 2\gamma n \ge (1/k - 2\gamma)|V(H')|.$$

We apply Lemma 4.1 to H' with parameter 2γ in place of γ. If Lemma 4.1 finds an independent set of size $(1 - 2k(2\gamma))\frac{k-1}{k}|V(H')|$, then we can output this independent set and halt, as

$$\begin{aligned}(1-4k\gamma)\frac{k-1}{k}|V(H')| &\ge (1-4k\gamma)\frac{k-1}{k}(n - \mu n)\\ &\ge (1-5k\gamma)\frac{k-1}{k}n.\end{aligned}$$

Thus we can assume that Lemma 4.1 finds a matching M_3 covering all but a set S_0 of at most $k^2/(2\gamma) \le k^2/\gamma$ vertices of $V(H')$.

Recall that $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M_1)) \in L$. By the definition of M_2, we have $\mathbf{i}_{\mathcal{P}}(V(H) \setminus V(M_1 \cup M_2)) \in L$, which implies $\mathbf{i}_{\mathcal{P}}(S_0) + \sum_{e \in M_0 \cup M_3} \mathbf{i}_{\mathcal{P}}(e) \in L$. As in the previous section, by the pigeonhole principle we can find in time $O(n)$ edges $e_1, \dots, e_{d'} \in M_0 \cup M_3$ for some $d' \le r - 1$ such that $\mathbf{i}_{\mathcal{P}}(S_0) + \sum_{i \in [d']} \mathbf{i}_{\mathcal{P}}(e_i) \in L$.

We delete $e_1, \dots, e_{d'}$ from our matching, thus leaving an unmatched set $D := \bigcup_{i \in [d']} e_i \cup S_0$. Then $\mathbf{i}_\mathcal{P}(D)$ is in the integer span of I and $|D| \le k^2(1+\gamma^{-1})$, so by definition of C, we have $\mathbf{i}_\mathcal{P}(D) = \sum_{\mathbf{v} \in I} a_\mathbf{v} \mathbf{v}$ as in (4.1) with each $|a_\mathbf{v}| \le C$. We write each as $a_\mathbf{v} = b_\mathbf{v} - c_\mathbf{v}$ such that one of $b_\mathbf{v}, c_\mathbf{v}$ is a nonnegative integer and the other is zero. Thus we have

$$\mathbf{i}_\mathcal{P}(D) + \sum_{\mathbf{v} \in I} c_\mathbf{v} \mathbf{v} = \sum_{\mathbf{v} \in I} b_\mathbf{v} \mathbf{v}. \tag{4.2}$$

Let $\mathcal{F} \subseteq M_2$ contain $c_\mathbf{v}$ sets of index $\mathbf{v}$ for each $\mathbf{v} \in I$. We delete the family $\mathcal{F}$ of edges from M_2, thus leaving $V(\mathcal{F}) \cup D$ as the unmatched set.

The combinatorial meaning of (4.2) is that $V(\mathcal{F}) \cup D$ can be expressed as the disjoint union over $\mathbf{v} \in I$ of $b_\mathbf{v}$ sets of index $\mathbf{v}$. Thus we regard $V(\mathcal{F}) \cup D$ as the union of $|\mathcal{F}| + |D|/k \le kC + k + k/\gamma$ sets S with $\mathbf{i}_\mathcal{P}(S) \in I$. Each such S has at least $C' - k^2 \ge d' + |\mathcal{F}| + |D|/k$ absorbing sets in $\mathcal{F}_0$, of which at most d' may be unavailable due to deleting $e_1, \dots, e_{d'}$, so we can greedily absorb all such S, thus obtaining a perfect matching of H.

For the running time, note that we can find $\mathcal{F}_{abs}$ in time $O(n^{2^{k+1}k^2})$, as $4tk^2 = 2^{k+1}k^2$. Then we find M_2 and M_3 both in time $O(n^{k+1})$, and $e_1, \dots, e_{d'}$ in time $O(n)$. Finally, it takes constant time to find $b_\mathbf{v}$ and $c_\mathbf{v}$ for all $\mathbf{v} \in I$, pick $\mathcal{F}$ and partition $V(\mathcal{F}) \cup D$; the absorption can also be done in constant time. The overall running time is $O(n^{2^{k+1}k^2})$.

5 The extremal case

In this section we prove Theorem 3.2, thus completing the proof of Theorem 1.1. We start by recalling the assumptions of the theorem, which will be assumed throughout the section. Let $k \ge 3$, suppose $\varepsilon > 0$ is sufficiently small, $n > n_0 = n_0(k, \varepsilon)$ is sufficiently large, and let $0 \le c \le \varepsilon n/k$ be an integer. Suppose H is a k-graph on n vertices with $\delta_{k-1}(H) \ge n/k - c$. Assume $H \notin \mathcal{H}_{n,k}$ and that H is ε-extremal, i.e. there is an independent subset $S \subseteq V(H)$ with $|S| \ge (1-\varepsilon)\frac{k-1}{k}n$. We will define an algorithm that finds either a perfect matching in H or a certificate that H has no perfect matching.

5.1 Proof strategy and notation We start by introducing some notation and terminology, also used throughout the section, which we will motivate with reference to the strategy of the proof. We define a partition (A, B, C) of $V(H)$ as follows. Let C be any *maximal* independent subset of $V(H)$ that contains S. Note that one can construct C from S greedily in time $O(n^k)$. For any $x \in V \setminus C$, we write $\deg(x, C)$ for the number of edges e of H containing x with $e \setminus \{x\} \subseteq C$. We let

$$A = \{x \in V \setminus C : \deg(x, C) \ge (1-\alpha)\binom{|C|}{k-1},$$

where $a = \varepsilon^{1/3}$. The final part of the partition is $B = V(H) \setminus (A \cup C)$. We will see in Lemma 5.6 that B is small.

As C is independent, we have $|e \cap C| \le k-1$ for any edge e, so we can assume $|C| \le (k-1)n/k$, as otherwise it cannot be covered by a matching. Thus $|C| \approx (k-1)n/k$ and $|A| \approx n/k$, so the bulk of any perfect matching can be thought of as a matching in an auxiliary graph between $A' \subseteq A$ and a partition of $C' \subseteq C$ into $(k-1)$-sets. The main task of the proof will be to efficiently check whether we can delete a small matching to reduce to such an auxiliary graph where a perfect matching can easily be found by derandomising a theorem of Pikhurko (see Theorem 5.1).

The *slack* of a matching M in H is defined as

$$s_M := |(A \cup B) \setminus V(M)| - (n/k - |M|).$$

As noted above, we can assume

$$s := s_\varnothing = |A| + |B| - n/k = (k-1)n/k - |C| \ge 0.$$

In general, to use M as part of a perfect matching we require $s_M \ge 0$. Indeed, writing $X = (A \cup B) \setminus V(M)$ and $Y = C \setminus V(M)$, if $s_M < 0$ then $|Y| = n - k|M| - |X| = (k-1)|X| - ks_M > (k-1)|X|$, so as Y is independent there cannot be a perfect matching.

We call a matching M a *cleaner* if M covers B and s_M is non-negative and even. Note that we impose the final condition to avoid a parity obstruction: we may only have edges that have odd intersection with X, in which case a perfect matching of $H[X \cup Y]$ requires $|X|$ and $n/k - |M|$ to have the same parity, namely, s_M is even. We summarise the above remarks as follows.

OBSERVATION 1. *Suppose M is contained in a perfect matching and $X = (A \cup B) \setminus V(M)$. Then:*

(i) $s_M \ge 0$, and (ii) if all $|e \cap X|$ with $e \in E(H)$ are odd then s_M is even.

Next we identify a subset D of vertices in B with relatively high degree within C; we will see in Lemma 5.8 below that it is easy to greedily extend any small matching to one covering D. We list the vertices of B as $v_1, \dots, v_{|B|}$ so that $\deg(v_i, C)$ is a non-increasing sequence. We find (in linear time) the largest $d > 0$ such that $\deg(v_d, C) > (d+c)(k-1)\binom{|C|}{k-2}$; if no such d exists then let $d = 0$. Let $D := \{v_1, \dots, v_d\}$ if $d > 0$, or $D := \varnothing$ otherwise.

We conclude this subsection with some compact notation for describing the type of a set or edge with respect to the partition (A,B,C) of $V(H)$. We say that S is an $A^iB^jC^l$ set if $|S\cap A|=i$, $|S\cap B|=j$ and $|S\cap C|=l$; we also say S has the form $A^iB^jC^l$. If any index is 0, we omit it, e.g. we write B^jC^l instead of $A^0B^jC^l$. If S is an edge of H we call it an $A^iB^jC^l$ edge. We also call it an $(i+j)$-edge. We write $N_H(v,A^iB^jC^l)$ for the $(k-1)$-sets in $N_H(v)$ of form $A^iB^jC^l$, namely,

$$\begin{aligned}&N_H(v,A^iB^jC^l)\\ :=&\{S\in N_H(v):|S\cap A|=i,|S\cap B|=j,|S\cap C|=l\}\end{aligned}$$

and also $\deg(v,A^iB^jC^l)=|N_H(v,A^iB^jC^l)|$.

5.2 The algorithm and proof modulo lemmas

Now we define our algorithm (it refers to some lemmas stated below). The notation is as in the previous subsection; we also let

$$t=n/k-|A|=|B|-s \quad\text{and}\quad d=|D|.$$

Procedure PerfectMatchingEXT

Data : An ε-extremal n-vertex k-graph $H\notin\mathcal{H}_{n,k}$ with $\delta_{k-1}(H)\geq n/k-c$, an independent set S with $|S|\geq(1-\varepsilon)\frac{k-1}{k}n$, and sets A, B, C, D as defined above.

Output : A perfect matching in H or a certificate that none exists.

if $n<n_0$ **then**
  Examine every set of n/k edges in H, and halt with appropriate output.
if $|C|>(k-1)n/k$ **then**
  Output "no PM" and C, and halt.
if *$s=|A|+|B|-n/k$ is odd and H contains no j-edge with $j\leq s+1$ even* **then**
  Output "no PM" and H, and halt.
if *there is no matching M_0 of size $\max\{t-d,0\}$ in $H[(B\setminus D)\cup C]$* **then**
  Output "no PM" and $(B\setminus D)\cup C$, and halt.
else
  Output "PM" ;
Use one of Lemmas 5.3, 5.4 or 5.5 to find a cleaner M, then Lemma 5.2 to find a perfect matching.

Next we state five lemmas needed for the proof of Theorem 3.2 (the proofs are deferred to later in the section). The first lemma is required so that the search for M_0 is feasible, the second shows that is sufficient to find a cleaner, and the others show how to find a cleaner in various cases.

Lemma 5.1. *We have $t-d\leq c$, so M_0 can be found in time $O(n^{ck})$.*

Lemma 5.2. *If H has a cleaner M then $H\setminus V(M)$ has a perfect matching, which can be found in time $O(n^{4(k-1)})$.*

Lemma 5.3. *Suppose H contains no j-edge for all even $0\leq j\leq k$ and $H\notin\mathcal{H}_{n,k}$. If Procedure PerfectMatchingEXT outputs "PM" then we can find a cleaner in time $O(n^k)$.*

Lemma 5.4. *Suppose $s=0$ or that H has an i-edge for some even $i\in[2,k]$ but does not have any ABC^{k-2} edge. If Procedure PerfectMatchingEXT outputs "PM" then we can find a cleaner in time $O(n^k)$.*

Lemma 5.5. *Suppose $s>0$ and H has some ABC^{k-2} edge e_0. If Procedure PerfectMatchingEXT outputs "PM" then we can find a cleaner in time $O(n^k)$.*

We conclude this subsection by assuming these lemmas and proving Theorem 3.2. First we make the following observation which will henceforth be used without comment.

Observation 2. *If a matching M has n_i i-edges then $s_M=s-\sum_{i\geq1}n_i(i-1)$.*

Proof. [Proof of Theorem 3.2] We need to show that Procedure PerfectMatchingEXT runs in time $O(n^{\max\{ck,4(k-1)\}})$, and either finds a perfect matching in H or a certificate that none exists.

For the running time, we find C as an arbitrary maximal independent set containing S in time $O(n^k)$, and then find A and B by determining their degrees to C in time $O(n^k)$. Next we sort the degrees of vertices from B in time $O(n\log n)$ and find d and D in linear time. The first two tests runs in constant time and the third in time $O(n^k)$. By Lemma 5.1, the search for M_0 takes time $O(n^{ck})$. If the output is "PM" then the conclusion of the algorithm finds a perfect matching in time $O(n^{4(k-1)})$ (assuming the lemmas).

It remains to show that there is no perfect matching if the algorithm outputs "no PM". The correctness of the first three tests follows from Observation 1. Indeed, for the third test, let $M=\{e\}$ for any j-edge e with $j\geq s+3$ even. Then $s_M=s-(j-1)\leq-2$, so no matching containing e is perfect. On the other hand, as s is odd, any matching consisting of odd edges only cannot be perfect either. For the final test, note that in any matching $A\cup D$ can be incident to at most $|A\cup D|=n/k-t+d$ edges, so any perfect matching must contain a matching of size $\max\{t-d,0\}$ completely within $(B\setminus D)\cup C$.

5.3 Using a cleaner In this subsection we prove Lemma 5.2, which shows how to find a perfect matching assuming that there is a cleaner. We start by estimating the sizes of the parts A, B, C; as discussed above, B is small, $|A| \approx n/k$ and $|C| \approx n - n/k$.

LEMMA 5.6. *$|A| \geq n/k - \alpha^2 n$, $|B| \leq \alpha^2 n$ and $(1 - \varepsilon)\frac{(k-1)n}{k} \leq |C| \leq \frac{(k-1)n}{k}$.*

Proof. The upper bound on C follows from Observation 1 and the lower bound from our assumptions for Theorem 3.2, which give $|C| \geq |S| \geq (1-\varepsilon)(k-1)n/k$. As $|A| + |B| + |C| = n$, we have

$$(5.3)\quad 0 \leq s = |A| + |B| - n/k = n - n/k - |C| \leq \varepsilon(k-1)n/k.$$

By the definitions of A and B, we have

$$\begin{aligned}\left(\frac{n}{k} - c\right)\binom{|C|}{k-1} &\leq \sum_{x \in A \cup B} \deg(x, C) \\ &\leq (1-\alpha)\binom{|C|}{k-1}|B| + \binom{|C|}{k-1}|A|.\end{aligned}$$

We deduce $n/k - c \leq |A| + |B| - \alpha|B|$, so $\alpha|B| \leq |A| + |B| - n/k + c \leq \varepsilon n$ by (5.3) and $c \leq \varepsilon n/k$. Thus $|B| \leq \alpha^2 n$ and by (5.3) again, $|A| \geq n/k - |B| \geq n/k - \alpha^2 n$.

We also require the following algorithmic version of a special case of a result of Pikhurko [25], concerning perfect matchings in a k-graph H that is k-partite, i.e. $V(H)$ has a partition $(V_1, \dots, V_k)$ so that every edge intersects all k parts. For $S \subseteq [k] = \{1, \dots, k\}$ we write $\delta_S(H)$ for the minimum degree $\deg_H(\{v_i : i \in S\})$ of any set consisting of one vertex in each of the parts $(V_i : i \in S)$.

THEOREM 5.1. [25, Theorem 3] *Suppose $1/n \ll \gamma \ll 1/k$. Let H be a k-partite k-graph with parts $V_1, \dots, V_k$ each of size n. Suppose $\delta_{\{1\}}(H) \geq (1-\gamma)n^{k-1}$ and $\delta_{[k]\setminus\{1\}}(H) \geq (1-\gamma)n$. Then there is an algorithm that finds a perfect matching in H in time $O(n^{4(k-1)})$.*

As Pikhurko's proof is probabilistic, we give an alternative derandomised proof via Lemma 4.4 (see also [11] for a similar proof).

Proof. We will apply Lemma 4.4 to the auxiliary graph G with parts $U = V_1$ and $W = \prod_{i=2}^k V_i$, whose edges consist of all $\{u, S\}$ with $u \in U$, $S \in W$ and $\{u\} \cup S \in E(H)$, and all $\{S, S'\}$ with $S, S' \in W$ and $S \cap S' \neq \emptyset$. According to the notation of Lemma 4.4 we have $M = n$, $N = n^{k-1}$ and $m = |G[W]| \leq (k-1)n^{2k-3}$. For any $u \in U$ we have $\deg_G(u, W) \geq \delta_{\{1\}}(H) \geq (1-\gamma)n^{k-1}$, so we can take $\beta = 1 - \gamma$. To satisfy the conditions of Lemma 4.4 we let $\tau := 1/4$ and $r := 4\gamma n$. Then for large n we have

$$\exp\left(\frac{\tau^2 r}{3(1-\gamma)}\right) \geq \exp\left(\frac{\gamma n}{12}\right) \geq 8n = 8M.$$

By Lemma 4.4, in time $O(N^4 + MN^3) = O(n^{4(k-1)})$, we can find an independent set R in $G[W]$ such that $(1-\nu)r \leq |R| \leq r$ and $|N_G(u) \cap R| \geq (\beta - \tau - \nu)r$ for all $u \in U$, where

$$\nu = \frac{2mr}{N^2} \leq \frac{8(k-1)\gamma n^{2k-n}}{n^{2k-2}} = 8(k-1)\gamma.$$

Thus R is a collection of disjoint $(k-1)$-sets of the form $v_2 \dots v_k$ with each $v_i \in V_i$ with $(1-\nu)4\gamma n \leq |R| \leq 4\gamma n$ and $|N_G(u) \cap R| \geq (1 - 8k\gamma - 1/4)r > r/2 = 2\gamma n$ for all $u \in V_1$, for small γ.

Next we partition $(\bigcup_{2 \leq i \leq k} V_i) \setminus V(R)$ arbitrarily into $n - |R|$ disjoint $(k-1)$-sets $S_1, S_2, \dots, S_{n-|R|}$ each having one vertex in each of $V_2, \dots, V_k$. We consider the bipartite subgraph G' of G induced by V_1 and $\{S_1, \dots, S_{n-|R|}\}$.

We claim that the set V_1' of vertices $v \in V_1$ with $\deg_{G'}(v) < (n - |R|)/2$ has size at most $2\gamma n$. To see this, first note that $|E(G')| \geq (1-\gamma)n(n-|R|)$ as $\delta_{[k]\setminus\{1\}}(H) \geq (1-\gamma)n$. Thus G' has at most $\gamma n(n - |R|)$ non-edges, whereas every vertex in V_1' is in at least $(n-|R|)/2$ non-edges of G', so the claim holds.

We now proceed as follows.

1. Greedily match the vertices of V_1' with $(k-1)$-sets in R. This is possible because each vertex v forms an edge with more than $2\gamma n$ sets in R, and $|V_1'| \leq 2\gamma n$.

2. Greedily match the remaining sets of R with vertices in $V_1 \setminus V_1'$. This is possible as $|R| \leq 4\gamma n$ and each set in R forms an edge with at least $(1-\gamma)n$ vertices in V_1.

3. Let V_1'' be the set of uncovered vertices in V_1 and G'' be the bipartite subgraph of G induced by V_1'' and $\{S_1, \dots, S_{n-|R|}\}$. We note that both parts of G'' have size $n - |R|$ and G'' has minimum degree at least $(n-|R|)/2$. Thus G'' has a perfect matching by Hall's theorem, which can be found in time $O(n^3)$ by the Hungarian algorithm.

Clearly the union of the matchings above gives a perfect matching in H.

Proof. [Proof of Lemma 5.2] Suppose H has a cleaner, i.e. a matching M that covers B such that $s_M = |(A \cup B) \setminus V(M)| - (n/k - |M|)$ is non-negative and even. We will show that $H \setminus V(M)$ has a perfect matching, which can be found in time $O(n^{4(k-1)})$.

We start by finding a matching M_* of size $2\alpha^2 n$ with $V(M_*) \cap V(M) = \emptyset$ consisting only of A^2C^{k-2} edges or only of A^3C^{k-3} edges. To do so, we apply a greedy algorithm to choose edges one by one, which we call 'good' if they avoid $V(M)$ and all previous choices, until we obtain $2\alpha^2 n$ edges of one of the required forms. At each step, we consider any good set S consisting of 2 vertices in A and $k-3$ vertices in C. As $\deg_H(S) \geq n/k - c > |B| + k|M| + 2k\alpha^2 n$ we can complete S to a good edge of the form A^2C^{k-2} or A^3C^{k-3}. Thus the algorithm to find M_* can be completed, and clearly takes time $O(n^2)$.

We obtain M' from M by adding some edges of M_*: we add s_M edges if they have the form A^2C^{k-2} or $s_M/2$ edges if they have the form A^3C^{k-3}. In both cases we obtain $s_{M'} = 0$, i.e. $|C'| = (k-1)|A'|$ where $A' = A \setminus V(M')$ and $C' = C \setminus V(M')$. We partition C' arbitrarily into $k-1$ parts $C^1, C^2, \dots, C^{k-1}$ each of size $m := |A'|$. We note by (5.3) that $m \geq |A| - k|M'| \geq |A| - k(|B| + s_M) > n/k - 2k\alpha^2 n$.

To find a perfect matching, it suffices to show that Theorem 5.1 applies to the k-partite sub-k-graph H' of H with parts $A', C^1, C^2, \dots, C^{k-1}$.

To bound $\delta_{[k]\setminus\{1\}}(H')$ we consider any set S formed by $k-1$ vertices $v_i \in C^i$ for $i \in [k-1]$. As C is independent, the number of non-neighbours of S in $A \cup B$ is at most

$$|A| + |B| - n/k + c \leq \varepsilon \frac{(k-1)n}{k} + c \leq k\varepsilon m,$$

where we use (5.3) in the first inequality and the last inequality follows from $m = |A'| \geq n/k - 2k\alpha^2 n > \frac{k-1}{k^2} n$. We deduce $\delta_{[k]\setminus\{1\}}(H') \geq m - k\varepsilon m = (1 - k\varepsilon)m$.

To bound $\delta_{\{1\}}(H')$, we consider any $v \in A' \subseteq A$, and note that by definition of A, the number of $(k-1)$-sets $S \subseteq C$ with $S \notin N_H(v)$ is at most

$$\alpha \frac{|C|^{k-1}}{(k-1)!} \leq \alpha \frac{\left(\frac{k-1}{k} n\right)^{k-1}}{(k-1)!} \leq \alpha \frac{(km)^{k-1}}{(k-1)!} = \alpha c_k m^{k-1},$$

where $c_k = \frac{k^{k-1}}{(k-1)!}$. This implies $\delta_{\{1\}}(H') \geq (1 - \alpha c_k) m^{k-1}$.

By Theorem 5.1 with $\gamma = \alpha c_k$, we find a perfect matching in H' in time $O(n^{4(k-1)})$.

5.4 Greedy extension In this subsection we gather various results that will be used in the following subsection for greedy extension of matchings during the construction of a cleaner.

Lemma 5.7. *Let $B' \subseteq B$ and $X \subseteq A \cup C$ with $|X| \leq k|B|$. Then in time $O(n^2)$ we can find a matching M_2 with $V(M_2) \cap X = \varnothing$ that covers all vertices of B' by edges of the form ABC^{k-2} or BC^{k-1}.*

Proof. For each $v \in B'$ in turn we pick $k-2$ arbitrary vertices from $C \setminus X$, and an uncovered vertex in $V \setminus (B \cup X)$ to complete an edge, which we add to M_2. Since $|B| + |X| \leq 2k|B| < n/k - c \leq \delta_{k-1}(H)$, such an edge always exists.

Next we recall that D consists of d vertices each having degree in C larger than $(d+c)(k-1)\binom{|C|}{k-2}$. To see the following lemma, we consider a greedy algorithm for constructing M, and note that at each step at most $(k-1)(d-1+c+1)$ vertices in C are already chosen, so the number of unavailable $(k-1)$-sets in C is at most $(d+c)(k-1)\binom{|C|}{k-2}$.

Lemma 5.8. *Given any set $C_* \subseteq C$ of size at most $(c+1)(k-1)$, we can find a matching of M of size d such that $D \subseteq V(M) \subseteq D \cup (C \setminus C_*)$ in time $O(n^k)$.*

Now we give a degree bound that will be used to greedily cover $B \setminus D$.

Lemma 5.9. *If $D \subsetneq B$ then $\deg(v, AC^{k-2}) > k|B||A|\binom{|C|}{k-3}$ for every vertex $v \in B \setminus D$.*

Proof. Suppose for contradiction that there is $v \in B \setminus D$ with $\deg(v, AC^{k-2}) \leq k|B||A|\binom{|C|}{k-3}$. By considering the degree of v together with each $(k-2)$-set in C we have

$$\begin{aligned} &\deg(v, AC^{k-2}) + \deg(v, BC^{k-2}) + (k-1)\deg(v, C^{k-1}) \\ &\geq \binom{|C|}{k-2}\left(\frac{n}{k} - c\right). \end{aligned}$$

As $|B| \leq \alpha n^2$ by Lemma 5.6 and α is small we deduce

$$\begin{aligned} &(k-1)\deg(v, C^{k-1}) \\ &\geq \binom{|C|}{k-2}\left(\frac{n}{k} - c\right) - k|B||A|\binom{|C|}{k-3} - |B|\binom{|C|}{k-2} \\ &\geq k^2(|B| + c)\binom{|C|}{k-2}. \end{aligned}$$

However, this contradicts the definition of D, so the lemma holds.

We conclude this subsection by establishing the lower bound $d = |D| \geq t - c$.

Proof. [Proof of Lemma 5.1] Note that for any vertex $b \in B \setminus D$, if $d > 0$ then $\deg(b, C) \leq (d+c)(k-1)\binom{|C|}{k-2}$; if $d = 0$ then $\deg(b, C) \leq (1+c)(k-1)\binom{|C|}{k-2}$. So it always holds that $\deg(b, C) \leq (d+1+c)(k-1)\binom{|C|}{k-2}$. Recall that (i) $\delta_{k-1}(H[B \cup C]) \geq t - c$ and independence of C, and (ii) $|B| \leq \alpha^2 n$ and $|C| \leq (k-1)n/k$, we have

$$\begin{aligned} &\binom{|C|}{k-1}(t-c) \leq \sum_{b \in B} \deg(b, C) \\ &\leq d\binom{|C|}{k-1} + |B|(d+1+c)(k-1)\binom{|C|}{k-2} \\ &\leq (d+\alpha)\binom{|C|}{k-1}. \end{aligned}$$

This implies the lemma as $t, c, d \in \mathbb{N}$.

5.5 Finding a cleaner We conclude this section by proving the three lemmas that find a cleaner according to the various cases of the proof of Theorem 3.2.

Proof. [Proof of Lemma 5.3] Assume that H contains no even edges and $H \notin \mathcal{H}_{n,k}$. Then the slack $s = |A|+|B|-n/k$ must be even. We apply Lemma 5.7 to find a matching M covering B, which must consist of BC^{k-1} edges, as there is no 2-edge. This preserves the slack, i.e. $s' = s = 0$, so M is a cleaner, as required.

Proof. [Proof of Lemma 5.4] We start with the case $s = 0$. Procedure PerfectMatchingEXT finds a matching M_0 of size $t - d = |B| - |D|$ in $(B \smallsetminus D) \cup C$, as otherwise it would output "no PM". By Lemma 5.8, we can enlarge M_0 to a matching M of size $|B|$ covering B and $(k-1)|B|$ vertices in C. Thus we preserve the slack, i.e. $s' = s = 0$, so M is a cleaner, as required.

It remains to consider $s > 0$. By assumption, H has an i-edge for some even $i \in [2,k]$ and does not have any ABC^{k-2} edge. Thus whenever we apply Lemma 5.7, we always obtain edges of the form BC^{k-1}. If s is even then we can simply apply Lemma 5.7 to construct a matching M covering B with edges of the form BC^{k-1}. Indeed, this preserves the slack, i.e. $s' = s$, so M is a cleaner, as required.

Finally, we can assume that s is odd. By assumption, there is an i-edge e_0 with i even. We fix any e_0 that minimises i. Note that $s \geq i - 1$, otherwise Procedure PerfectMatchingEXT would output "no PM". We apply Lemma 5.7 to construct a matching M consisting of e_0 and $|B \smallsetminus e_0|$ edges covering $B \smallsetminus e_0$ with edges of the form BC^{k-1}. Thus we obtain slack $s' = |A \smallsetminus V(M)| - (n/k - |M|) = s - i + 1$, which is even and non-negative, so again M is a cleaner, as required.

Proof. [Proof of Lemma 5.5] Suppose $s > 0$ and H has some ABC^{k-2} edge e_0, i.e. $|e_0 \cap A| = |e_0 \cap B| = 1$, and Procedure PerfectMatchingEXT outputs "PM". We write $e_0 \cap B = \{x\}$. Our proof will use two different strategies for finding a cleaner that in combination cover all possible cases. The main strategy (which applies to all but one case) finds in time $O(n^k)$ a matching $M = M_1 \cup M_2$ covering B, such that, writing $s_2 := s_{M_2} = |(A \cup B) \smallsetminus V(M_2)| - (n/k - |M_2|)$,

1. every edge e of M has $|e \cap B| = 1$ and $|e \cap A| \leq 1$,
2. $|M_1| \geq t + 1 = n/k - |A| + 1 = |B| - s + 1$,
3. the number of edges of M_1 that intersect A is $(s_2 \bmod 2) \in \{0,1\}$.

Note that if we find such a matching M then $|M_2| = |M| - |M_1| \leq |B| - (t+1) = s - 1$ and $s_2 = s + |M_2| - |(A \cup B) \cap V(M_2)| \geq s - |M_2| \geq 1$, so $s_M = s_2 - (s_2 \bmod 2)$ is even and non-negative, i.e. M is a cleaner, as required. We will apply the main strategy to 3 cases below.

Case 1. Suppose $D = B$. By Lemma 5.8, we can find a matching M' covering B by BC^{k-1} edges such that $V(M')$ is disjoint from $e_0 \smallsetminus B$. Then $|M'| = |B| = t + s \geq t + 1$, as $s > 0$. Let e_x be the edge of M' containing x (the vertex in $e_0 \cap B$). We let $M_2 = \varnothing$, so $s_2 = s = |A| + |B| - n/k$, and let $M_1 = M'$ if s_2 is even or otherwise let $M_1 = (M' \smallsetminus \{e_x\}) \cup \{e_0\}$. Then M satisfies (1-3).

Case 2. Suppose $t \leq |D| < |B|$. Fix any $v \in B \smallsetminus D$. By Lemma 5.9, we can find an ABC^{k-2} edge e' containing v. Moreover, as C is a maximal independent set, we can find a BC^{k-1} edge e_v containing v. Next, by Lemma 5.8 we can find a matching M' covering D by BC^{k-1} edges with $V(M') \cap (e_v \cup e') = \varnothing$. Then by Lemma 5.7 we find a matching M_2 covering $B \smallsetminus (D \cup \{v\})$ with $V(M_2) \cap (e_v \cup e' \cup V(M')) = \varnothing$. We let $M_1 = M' \cup \{e_v\}$ if s_2 is even or $M_1 = M' \cup \{e'\}$ otherwise. Then M satisfies (1-3).

Case 3. Suppose $|D| = d < t$ and there is a BC^{k-1} edge e_* disjoint from $D \cup V(M_0)$, where M_0 is the matching of size $t - d$ in $(B \smallsetminus D) \cup C$ found by Procedure PerfectMatchingEXT. We write $e_* \cap B = \{v\}$. By Lemma 5.8, we can find BC^{k-1} edges covering D that extend $M_0 \cup \{e_*\}$ to a matching M' of size $t+1$. Since $v \notin D$, by Lemma 5.9 we have $\deg(v, AC^{k-2}) > k|B||A|\binom{|C|}{k-3} \geq k|M'||A|\binom{|C|}{k-3}$, so we can find an ABC^{k-2} edge e_v containing v such that $e_v \cap V(M') = \{v\}$. By Lemma 5.7 we can find M_2 covering $B \smallsetminus V(M')$ such that $V(M_2) \cap (V(M') \cup e_v) = \varnothing$. We let $M_1 = M'$ if s_2 is even or $M_1 = (M' \smallsetminus \{e_*\}) \cup \{e_v\}$ otherwise. Then M satisfies (1-3).

Finally, we describe the second strategy for finding a cleaner, which will complete the proof when none of the above 3 cases apply, i.e. when $|D| = d < t$ and there is no BC^{k-1} edge e_* disjoint from $D \cup V(M_0)$. By Lemma 5.8, we can find BC^{k-1} edges covering D that extend M_0 to a matching M' of size t.

Then we claim that we can find ABC^{k-2} edges covering $B \smallsetminus V(M')$ that extend M' to a matching M. To see this, we apply a greedy algorithm, where in each step, to cover some vertex x of $B \smallsetminus V(M')$ we consider any BC^{k-2} set S containing x disjoint from all previous edges. By our assumptions for this case, $N(S)$ is disjoint from C, so as $|N(S)| \geq n/k - c > k|B|$ we can complete S to an ABC^{k-2} edge as required, so we can construct M as claimed.

Writing $t' = |V(M_0) \cap B|$, we note that M_0 reduces the slack by $t' - |M_0| = t' - t + d$, and $M' \smallsetminus M_0$ keeps it unchanged. Finally, $M \smallsetminus M'$ reduces the slack by

$|M \setminus M'| = |B| - d - t'$. Therefore $s_M = s - (t' - t + d) - (|B| - t' - d) = 0$, so M is a cleaner.

6 Concluding remarks

In this paper, we showed that $\mathbf{DPM}_k(n, m)$ is in P for $m \geq n/k - c$ for constant c. For simplicity of analysis we did not attempt to optimise the exponent in the running time. We remark that some improvements can be obtained by merging 'transferrals' and working with two partitions as in [10], and/or settling for just testing for a perfect matching rather than actually finding one.

As observed in [10], the argument in [30] actually shows that $\mathbf{DPM}_k(n, n/k - n^{\varepsilon})$ is NP-complete for any $\epsilon > 0$. Thus the complexity remains unknown for $n/k - n^{\varepsilon} \leq m < n/k - c$. Our algorithm is valid for any $c = o(n)$, but our proof only bounds the running time by $O(n^{ck})$. The bottleneck to improving this comes from the extremal case, where we showed that the existence of a perfect matching in H is equivalent to that of a matching of size $t := n/k - |A|$ in $H[B \cup C]$. Thus we have a reduction to the following potentially simpler problem, which we believe has independent interest.

PROBLEM 1. *Suppose $1/n \ll \varepsilon \ll 1/k$. Let X, Y be disjoint sets with $|X| = \varepsilon n$ and $|Y| = n$. Let $c = c(n)$ and $t = t(n)$ with $c \leq t \leq |X|$. Let H be a k-graph on $X \cup Y$ with $\delta_{k-1}(H) \geq t - c$ such that $H[Y]$ is independent. What is the complexity of deciding the existence of a matching of size t in H?*

References

[1] N. Alon, P. Frankl, H. Huang, V. Rödl, A. Ruciński, and B. Sudakov, *Large matchings in uniform hypergraphs and the conjecture of Erdős and Samuels*, J. Combin. Theory Ser. A **119** (2012), no. 6, 1200–1215. MR2915641 ↑1

[2] A. Asadpour, U. Feige, and A. Saberi, *Santa Claus meets hypergraph matchings*, Approximation, randomization and combinatorial optimization, 2008, pp. 10–20. MR2538773 ↑1

[3] A. Czygrinow and V. Kamat, *Tight co-degree condition for perfect matchings in 4-graphs*, Electron. J. Combin. **19** (2012), no. 2, Paper 20, 16. MR2928635 ↑1

[4] J. Edmonds, *Paths, trees, and flowers*, Canad. J. Math. **17** (1965), 449–467. MR0177907 (31 #2165) ↑1

[5] P. Frankl and A. Kupavskii, *The Erdős Matching Conjecture and concentration inequalities*, arXiv e-prints (2018Jun), arXiv:1806.08855, available at arXiv:1806.08855. ↑1

[6] F. Garbe and R. Mycroft, *Hamilton cycles in hypergraphs below the Dirac threshold*, J. Combin. Theory Ser. B **133** (2018), 153–210. MR3856709 ↑4, 5

[7] S. Glock, D. Kühn, A. Lo, and D. Osthus, *The existence of designs via iterative absorption*, arXiv e-prints (Nov. 2016), arXiv:1611.06827, available at arXiv:1611.06827. ↑1

[8] H. Hàn, Y. Person, and M. Schacht, *On perfect matchings in uniform hypergraphs with large minimum vertex degree*, SIAM J. Discrete Math **23** (2009), 732–748. ↑1

[9] J. Han, *Near perfect matchings in k-uniform hypergraphs*, Combin. Probab. Comput. **24** (2015), no. 5, 723–732. MR3371500 ↑2, 4

[10] ———, *Decision problem for perfect matchings in dense k-uniform hypergraphs*, Trans. Amer. Math. Soc. **369** (2017), no. 7, 5197–5218. ↑2, 3, 4, 11

[11] J. Han and Y. Zhao, *Minimum vertex degree threshold for C_4^3-tiling*, J. Graph Theory **79** (2015), no. 4, 300–317. ↑8

[12] R. M. Karp, *Reducibility among combinatorial problems*, Complexity of computer computations (Proc. Sympos., IBM Thomas J. Watson Res. Center, Yorktown Heights, N.Y., 1972), 1972, pp. 85–103. MR0378476 (51 #14644) ↑1

[13] M. Karpiński, A. Ruciński, and E. Szymańska, *Computational complexity of the Hamiltonian cycle problem in dense hypergraphs*, LATIN 2010: theoretical informatics, 2010, pp. 662–673. MR2673301 ↑5

[14] ———, *Computational complexity of the perfect matching problem in hypergraphs with subcritical density*, Internat. J. Found. Comput. Sci. **21** (2010), no. 6, 905–924. MR2740002 (2012c:68098) ↑1, 2

[15] P. Keevash, F. Knox, and R. Mycroft, *Polynomial-time perfect matchings in dense hypergraphs*, Proceedings of the 45th STOC (2013) (2013). ↑2

[16] ———, *Polynomial-time perfect matchings in dense hypergraphs*, Advances in Mathematics **269** (2015), no. 0, 265 –334. ↑2, 3, 4

[17] P. Keevash, *The existence of designs*, arXiv e-prints (Jan. 2014), arXiv:1401.3665, available at arXiv:1401.3665. ↑1

[18] I. Khan, *Perfect matchings in 3-uniform hypergraphs with large vertex degree*, SIAM J. Discrete Math. **27** (2013), no. 2, 1021–1039. MR3061464 ↑1

[19] ———, *Perfect matchings in 4-uniform hypergraphs*, J. Combin. Theory Ser. B **116** (2016), 333–366. ↑1

[20] D. Kühn and D. Osthus, *Matchings in hypergraphs of large minimum degree*, J. Graph Theory **51** (2006), no. 4, 269–280. MR2207573 (2006k:05177) ↑1

[21] D. Kühn, D. Osthus, and A. Treglown, *Matchings in 3-uniform hypergraphs*, J. Combin. Theory Ser. B **103** (2013), no. 2, 291–305. MR3018071 ↑1

[22] A. Lo and K. Markström, *Minimum codegree threshold for $(K_4^3 - e)$-factors*, J. Combin. Theory Ser. A **120** (2013), no. 3, 708–721. MR3007146 ↑2

[23] A. Lo and K. Markström, *F-factors in hypergraphs via absorption*, Graphs Combin. **31** (2015), no. 3, 679–712. ↑2

[24] K. Markström and A. Ruciński, *Perfect Matchings (and Hamilton Cycles) in Hypergraphs with Large Degrees*, European J. Comb. **32** (July 2011), no. 5, 677–687. ↑1

[25] O. Pikhurko, *Perfect matchings and K_4^3-tilings in hypergraphs of large codegree*, Graphs Combin. **24** (2008), no. 4, 391–404. MR2438870 (2009e:05214) ↑1, 8

[26] V. Rödl and A. Ruciński, *Dirac-type questions for hypergraphs — a survey (or more problems for endre to solve)*, An Irregular Mind **Bolyai Soc. Math. Studies 21** (2010), 561–590. ↑1

[27] V. Rödl, A. Ruciński, and E. Szemerédi, *Perfect matchings in uniform hypergraphs with large minimum degree*, European J. Combin. **27** (2006), no. 8, 1333–1349. MR2260124 (2007g:05153) ↑1

[28] ———, *Perfect matchings in large uniform hypergraphs with large minimum collective degree*, J. Combin. Theory Ser. A **116** (2009), no. 3, 613–636. MR2500161 (2010d:05124) ↑1

[29] H. Ryser, *Neuere probleme in der kombinatorik*, Vortrage über Kombinatorik, Oberwolfach (1967), 69–91. ↑1

[30] E. Szymańska, *The complexity of almost perfect matchings and other packing problems in uniform hypergraphs with high codegree*, European J. Combin. **34** (2013), no. 3, 632–646. MR2997414 ↑1, 11

[31] A. Treglown and Y. Zhao, *Exact minimum degree thresholds for perfect matchings in uniform hypergraphs*, J. Combin. Theory Ser. A **119** (2012), no. 7, 1500–1522. MR2925939 ↑1

[32] ———, *Exact minimum degree thresholds for perfect matchings in uniform hypergraphs II*, J. Combin. Theory Ser. A **120** (2013), no. 7, 1463–1482. MR3092677 ↑1

[33] ———, *A note on perfect matchings in uniform hypergraphs*, Electron. J. Combin. **23** (2016), P1.16. ↑1

[34] W. T. Tutte, *The factorization of linear graphs*, J. London Math. Soc. **22** (1947), 107–111. MR0023048 (9,297d) ↑1

[35] Y. Zhao, *Recent advances on dirac-type problems for hypergraphs*, Recent trends in combinatorics, 2016. ↑1