HAL
open science

On Subgraphs of Bounded Degeneracy in Hypergraphs

Kunal Dutta, Arijit Ghosh

To cite this version:

Kunal Dutta, Arijit Ghosh. On Subgraphs of Bounded Degeneracy in Hypergraphs. 2017. hal-01669886

HAL Id: hal-01669886

https://hal.science/hal-01669886v1

Preprint submitted on 21 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ON SUBGRAPHS OF BOUNDED DEGENERACY IN HYPERGRAPHS

KUNAL DUTTA AND ARIJIT GHOSH

Abstract. A k-uniform hypergraph is d-degenerate if every induced subgraph has a vertex of degree at most d. Given a k-uniform hypergraph $H = (V(H), E(H))$, we show there exists an induced subgraph of size at least

$$\sum_{v \in V(H)} \min\left\{1,\, c_k \left(\frac{d+1}{d_H(v)+1}\right)^{1/(k-1)}\right\},$$

where $c_k = 2^{-\left(1+\frac{1}{k-1}\right)}\left(1-\frac{1}{k}\right)$ and $d_H(v)$ denotes the degree of vertex v in the hypergraph H. This connects, extends, and generalizes results of Alon-Kahn-Seymour (Graphs and Combinatorics, 1987), on d-degenerate sets of graphs, Dutta-Mubayi-Subramanian (SIAM Journal on Discrete Mathematics, 2012) on d-degenerate sets of linear hypergraphs, and Srinivasan-Shachnai (SIAM Journal on Discrete Mathematics, 2004) on independent sets in hypergraphs to d-degenerate subgraphs of hypergraphs. Our technique also gives optimal lower bounds for a more generalized definition of degeneracy introduced by Zaker (Discrete Applied Mathematics, 2013). We further give a simple non-probabilistic proof of the Dutta-Mubayi-Subramanian bound for linear k-uniform hypergraphs, which extends the Alon, Kahn and Seymour (Graphs and Combinatorics, 1987) proof technique to hypergraphs. Finally we provide several applications in discrete geometry, extending results of Payne-Wood (SIAM Journal on Discrete Mathematics, 2013) and Cardinal-Tóth-Wood (Journal of Geometry, 2016). We also address some natural algorithmic questions.

The proof of our main theorem combines the *random permutation* technique of Bopanna-Caro-Wei and Beame and Luby, together with a new *local density* argument which may be of independent interest.

1. Introduction

For $k \geq 2$, a k-uniform hypergraph H is a pair $H = (V, E)$ where $E(H) \subseteq \binom{V(H)}{k}$. We will call $V(H)$ and $E(H)$ the vertex set and edge set of H

DataShape Group, INRIA Sophia Antipolis – Méditerranée, France
Advanced Computing and Microelectronics Unit, Indian Statistical Institute, Kolkata, India
E-mail addresses: `Kunal.Dutta@inria.fr, agosh@mpi-inf.mpg.de`.
Date: November 13, 2017.
Key words and phrases. degenerate graphs, independent sets, hypergraphs, and random permutations.

respectively. When there is no chance of confusion, we will use V and E to denote $V(H)$ and $E(H)$. For a vertex $v \in V(H)$, degree $d_H(v)$ of $V(H)$ will denote $|\{e : e \in E(H), v \in e\}|$. For readability, $k-1$ will be denoted by t.

For a subset $I \subseteq V(H)$, the induced k-uniform hypergraph $H(I)$ of I denotes the hypergraph $(I, E(H) \cap \binom{I}{k})$. A hypergraph is *linear* if every pair of vertices are contained in at most a single hyperedge, i.e. any pair of hyperedges intersect in at most one vertex. A hypergraph $H = (V, E)$ is *d-degenerate* if the induced hypergraph of all subsets of V has a vertex of degree at most d, i.e., for all $I \subseteq V$, there exists $v \in I$ such that $d_{H(I)}(v) \leq d$. For a k-uniform hypergraph $H = (V, E)$, we will denote by $\alpha_{k,d}(H)$ the size of a maximum-sized subset of V whose induced hypergraph is d-degenerate, i.e.,

$$\alpha_{k,d}(H) = \max\{|I| : I \subseteq V,\ H(I) \text{ is } d\text{-degenerate}\}.$$

Observe that $\alpha(H) := \alpha_{k,0}(H)$ is the *independence number* of the hypergraph H.

1.1. **Previous Results.** Turán [Tur41] gave a lower bound on the independence number of graphs: $\alpha(G) \geq \frac{n}{d+1}$ where d is the average degree of vertices in G. Caro [Car79] and Wei [Wei81] independently showed that for graphs

$$\alpha(G) \geq \sum_{v \in V(G)} \frac{1}{d_G(v)+1},$$

see [AS08]. This degree-sequence based bound matches the original average-degree based lower bound of Turán in the case when all degrees are equal, and improves it for general degree sequences.

For hypergraphs, Spencer [Spe72] gave a bound on the independence number, based on the average degree d: $\alpha(H) \geq c_k \left(\frac{n}{d^{1/t}}\right)$, where c_k is independent of n and d. Caro and Tuza [CT91] generalized the Caro-Wei result to the case of hypergraphs:

Theorem 1.1. *For all k-uniform hypergraph H, we have*

$$\alpha(H) \geq \sum_{v \in V} \frac{1}{\binom{d_H(v)+1/t}{d_H(v)}}.^{1}$$

The above theorem directly implies the following corollary, which gives Spencer's bound:

Corollary 1.2. *For all $k \geq 2$, There exists $d_k > 0$ such that all k-uniform hypergraphs H satisfy*

$$\alpha(H) \geq d_k \sum_{v \in V} \frac{1}{(1+d_H(v))^{1/t}}.$$

[1] Where for $l \in \mathbb{N} \cup \{0\}$ and $r \in \mathbb{R}$, $\binom{r}{l} = \frac{r(r-1)\ldots(r-l+1)}{l!}$.

Further, Thiele [Thi99] obtained a lower bound on the independence number of arbitrary (non-uniform) hypergraphs, in terms of the *degree rank*, a generalization of the degree sequence.

On the algorithmic side, Srinivasan and Shachnai [SS04], used the random permutation method of Beame and Luby [BL90] and also Bopanna-Caro-Wei (see e.g. [AS08], [aDMS12]), together with the FKG correlation inequality, to obtain a randomized parallel algorithm for independent sets, which matched the asymptotic bound of Caro and Tuza [CT91]. Dutta, Mubayi and Subramanian [aDMS12] also used the Bopanna-Caro-Wei method; alongwith some new combinatorial identities, they obtained degree-sequence based lower bounds on the independence numbers of K_r-free graphs and linear k-uniform hypergraphs, which generalized the earlier average-degree based bounds of Ajtai, Komlós, and Szemerédi [AKS80], Shearer [She83, She95] and Duke, Lefmann and Rödl [DLR95], in terms of degree sequences. [2]

Average degree vs. degree-sequence. Intuitively, a bound using the degree sequence should be expected to be better than a bound using just the average degree, since it has more information about the graph. For the above bounds on the independence numbers, this essentially follows from the convexity of the function $x^{-1/t}$. Dutta, Mubayi and Subramanian [aDMS12] gave constructions of hypergraphs which show that the bounds based on the degree-sequence can be stronger than those based on the average degree by a polylogarithmic factor in the number of vertices.

Large d-degenerate subgraphs. Unlike independent sets in graphs, d-degenerate subgraphs have been less well-investigated. However, it includes as special cases zero-degenerate subgraphs i.e. independent sets, as well as 1-degenerate subgraphs, i.e. maximum induced forests, whose complements are the well-known hitting set and feedback vertex set problems respectively. The best known result on this question is that of Alon, Kahn and Seymour [AKS87], who proved the following lower bound for $\alpha_{2,d}(G)$: [3]

Theorem 1.3 ([AKS87])**.** *For all graphs $G = (V, E)$ we have*

$$\alpha_{2,d}(G) \geq \sum_{v \in V} \min\left\{1, \frac{d+1}{d_G(v)+1}\right\}.$$

This bound is sharp for every G which is a disjoint union of cliques. Moreover, they gave a polynomial time algorithm that finds in G an induced d-degenerate subgraph of at least this size.

[2]Their proof also yields an elementary proof of the main bound of Srinivasan and Shachnai [SS04] without using correlation inequalities, though not explicitly stated.

[3]Alon, Kahn and Seymour [AKS87] actually defined a d-degenerate graph as one where every subgraph has a vertex of degree *less than* d.

On the algorithmic side, Pilipczuk and Pilipczuk [PP12] addressed the question of finding a maximum d-degenerate subgraph of a graph, giving the first algorithm with running time $o(2^n)$. Zaker [Zak13] studied a more general version of degeneracy and gave upper and lower bounds for finding the largest subgraph of a 2-uniform graph having a given generalized degeneracy.

The proof of Dutta-Mubayi-Subramanian [aDMS12] implies the following lower bound on $\alpha_{k,d}$ for linear hypergraphs (though not explicitly stated in their paper):

Theorem 1.4 ([aDMS12])**.** *Let $G = (V, E)$ be a linear k-uniform hypergraph, and for all $v \in V$, $d_G(v)$ denote the degree of v in G. Then*

$$\alpha_{k,d}(G) \geq w(G) := \sum_{v \in V} w_G(v), \tag{1}$$

where

$$w(v) = \begin{cases} 1 & \text{if } d_G(v) \leq d \\ \frac{1}{1+((k-1)(d+1))^{-1}} \frac{\binom{d_G(v)}{d+1}}{\binom{d_G(v)+\frac{1}{k-1}}{d_G(v)-d-1}} & \text{if } d_G(v) > d. \end{cases} \tag{2}$$

A possible extension and a counterexample. When $d = 0$, i.e. the subgraph is an independent set, the above expression reduces to the Caro-Tuza bound in Theorem 1.1. In the proof of the above theorem, the term corresponding to a given vertex v in the above expression is actually the fraction of orderings of the vertices, in which there are at most d hyperedges involving v and some vertices occuring prior to v, in the ordering. For linear k-uniform hypergraphs, the above expression is exact, but this does not hold for general k-uniform hypergraphs. However, in the case when $d = 0$, the above expression still gives an upper bound on the fraction of such orderings. One may wonder if a similar property might hold for general k-uniform graphs, for general d.

That is, does it hold that for any given vertex $v \in V$, the fraction of orderings of vertices where at most d hyperedges can be formed using only the vertex v and vertices occuring prior to v, is at most $\frac{1}{1+((k-1)(d+1))^{-1}} \frac{\binom{d_G(v)}{d+1}}{\binom{d_G(v)+\frac{1}{k-1}}{d_G(v)-d-1}}$?

This however can be seen to be false, from the following example:

Counterexample Consider the 3-uniform hypergraph $H = (V, E)$, given by $V = \{1, 2, 3, 4, 5\}$ and $E = \{\{1, 2, 3\}, \{1, 3, 4\}, \{1, 4, 5\}, \{1, 5, 2\}\}$. Taking $d = 1$, we get that the expression obtained from the above expression for the vertex 1, is

$$\frac{(3-1) \cdot 2}{1 + (3-1) \cdot 2} \frac{\binom{4}{2}}{\binom{4+\frac{1}{3-1}}{4}} = \frac{216}{315} = \frac{24}{35}.$$

On the other hand, the fraction of orderings of vertices, where the vertices occuring prior to 1 together with the vertex 1, constitute at most one hyperedge, is $\frac{72}{120} = \frac{3}{5} < \frac{24}{35}$. Therefore, the expression in Theorem 1.4 clearly

overshoots the actual fraction of orderings.

Thus, the bound in Theorem 1.4 does not hold for general d.

1.2. **Our Results.** We first give a completely different and an extremely simple proof of Theorem 1.4 using a weight function. Our proof follows along the lines of the proof of Theorem 1.3 due to Alon, Kahn and Seymour [AKS87].

Although the counterexample in the previous subsection shows that the bound in Theorem 1.4 does not hold for general d and general k-uniform hypergraphs, we show that the asymptotic expression (in terms of d and $d_G(v)$) still holds, thus extending Theorem 1.4 asymptotically to the case of general hypergraphs:

Theorem 1.5. *Let $G = (V, E)$ be a k-uniform hypergraph, and for all $v \in V$, $d_G(v)$ denote the degree of v in G. Then*

$$\alpha_{k,d}(G) \geq \sum_{v \in V} \min\left\{1,\, c_k \left(\frac{d+1}{d(v)+1}\right)^{1/t}\right\}, \tag{3}$$

where $t = k-1$ and $c_k = 2^{-\left(1+\frac{1}{(k-1)}\right)}\left(1-\frac{1}{k}\right)$. Further, there exists a randomized algorithm that can extract a d-degenerate set of above size in expectation.

Our proof uses the *random permutation method* [AS08] of Bopanna-Caro-Wei, together with a new local density argument, avoiding advanced correlation inequalities. As a consequence, we obtain a simpler proof as well as a generalization of the result of Srinivasan and Shachnai [SS04].

The rest of the paper is organised as follows: Section 2 has the simpler proof of Theorem 1.4, Section 3 has the proof of Theorem 1.5 and in Section 4 we prove lower bounds for a more generalized definition of degeneracy. In Section 4, we will show that our proof technique for Theorem 1.5 gives optimal lower bounds for a more generalized definition of degeneracy introduced recently by Zaker [Zak13]. In Section 5 we give several applications of our results in discrete and combinatorial geometry. Finally in the conclusion there are some remarks and open questions.

2. Linear Hypergraphs

In this section we will give an alternative proof of the Theorem 1.4. The proof will follow exactly along the lines of the proof by Alon, Kahn and Seymour [AKS87] of Theorem 1.3.

First observe that

$$\frac{\binom{d_G(v)}{r}}{\binom{d_G(v)+1/t}{d_G(v)-r}} = \frac{1}{\left(1+\frac{1}{t(r+1)}\right)\cdots\left(1+\frac{1}{td_G(v)}\right)}$$

This implies that $w(v)$ is decreasing in $d_G(v)$ for all values of $d_G(v) \geq r$. Also, observe that

$$\frac{\binom{d_G(v)-1}{r}}{\binom{d_G(v)-1+1/t}{d_G(v)-1-r}} = \left(1 + \frac{1}{t\, d_G(v)}\right) \frac{\binom{d_G(v)}{r}}{\binom{d_G(v)+1/t}{d_G(v)-r}} \tag{4}$$

The alternative proof will be by induction on the number n of vertices of the k-uniform hypergraph G. The base case of $n = 1$ follows trivially. Assuming the result holds for $n - 1$, we will now show that the result also holds for n. Given a vertex $v \in V$, let H_v denote the hypergraph $G(V')$ where $V' = V \setminus \{v\}$.

Case 1. If we have a vertex $v \in V(G)$ with $d_G(v) \leq d$, then consider $H = H_v$. Observe that $\alpha_{k,d}(G) = \alpha_{k,d}(H) + 1$. Since $\forall u \in V'$, we have from Equation (4), $w_G(u) \leq w_H(u)$. This implies

$$w(H) = \sum_{u \in V'} w_H(u) \geq \sum_{u \in V'} w_G(u) = w(G) - 1.$$

The last inequality follows from the fact that $w_G(v) = 1$ since $d_G(v) \leq d$. Using the induction hypothesis $\alpha_{k,d}(G) \geq w(G)$ and the fact that $\alpha_{k,d}(G) = \alpha_{k,d}(H) + 1$, we get

$$\alpha_{k,d}(G) = \alpha_{k,d}(H) + 1 \geq w(H) + 1 \geq w(G).$$

Case 2. Now we will consider the case where $d_G(v) > d$, $\forall v \in V(G)$. Let $\Delta = \max_{u \in V(G)} d_G(u)$, and let $v \in V(G)$ be a vertex with $d_G(v) = \Delta$. Let $u_1, \ldots, u_l$, where $l = t\Delta$, be the neighbors of v in G. Note that $l = t\Delta$ follows from the fact that G is a linear hypergraph. We will now show that $w(H) \geq w(G)$, where $H = H_v$.

$$\begin{aligned} w(H) &= \sum_{u \in V'} w_H(u) \\ &= w(G) - w_G(v) - \sum_{i=1}^{l} w_G(u_i) + \sum_{i=1}^{l} w_H(u_i) \\ &= w(G) - w_G(v) + \sum_{i=1}^{l} \frac{w_G(u_i)}{t\, d_G(u_i)} \\ &\geq w(G) \end{aligned}$$

The second last inequality follows from the facts that $d_H(u_i) = d_G(u_i) - 1$ (as G is a linear hypergraph) and Equation (4). The last inequality follows from the facts that $d_G(u) \leq \Delta$ for all $u \in V$ and $w_G(u_i) \geq w_G(v)$ (direct consequence of Equation (4)). From induction hypothesis we have

$$\alpha_{k,d}(H) \geq w(H) \geq w(G).$$

This completes the proof of Theorem 1.4 since $\alpha_{k,d}(G) \geq \alpha_{k,d}(H)$.

3. General k-uniform Hypergraphs

In this section we shall prove a lower bound on $\alpha_{k,d}(H)$ for general k-uniform hypergraph H in terms of its degree sequence. We will give a very simple randomized algorithm to obtain an d-degenerate subgraph of a k-uniform hypergraph, whose analysis in expectation will yield the desired bound in Theorem 1.5.

3.1. Details of the Algorithm.

Before we can give the details of the algorithm, we will need some definitions.

Definition 3.1. Let σ be an ordering of the vertices of H.

- Fix a vertex $v \in V(H)$. Call a hyperedge $e \in E(H)$ with $v \in e$ a *backward edge* with respect to σ, if $\forall u \in e \setminus \{v\}$, $\sigma(u) < \sigma(v)$.
- We will denote by $b_\sigma(v)$ the number of backward edges of the vertex v with respect to the ordering σ.

Algorithm 1 RandPermute

Input: $H := (V, E)$ and d;
// H is a k-uniform hypergraph
Random odering: Let σ be a random ordering of the vertex set V;
Initialization: $I \leftarrow \emptyset$;
// I will be the degenerate subset we output;
for $v \in V$ **do**
  Compute: $b_\sigma(v)$;
  if $b_\sigma(v) \leq d$ **then**
    $I \leftarrow I \cup \{v\}$;
  end if
end for
Output: I;

3.2. Analysis of the Algorithm.

Theorem 1.5 directly follows from the following result.

Claim 3.2.

$$\mathrm{E}\left[|I|\right] \geq \sum_{v \in V} \min\left\{1,\, c_k \left(\frac{d+1}{d_H(v)+1}\right)^{1/t}\right\},$$

where $c_k = c_k = 2^{-\left(1+\frac{1}{k-1}\right)}\left(1-\frac{1}{k}\right) = 2^{-(1+o_k(1))}$.

Proof. For all vertices $v \in V$, we will denote by $N(v)$ the neighbors of v in H.

Given an arbitrary vertex $v \in V$, and a random ordering of the vertices σ, we need to bound $\Pr\left[v \in I\right]$ from below. Since the event of v being selected depends on the relative ordering of the vertices in $N(v)$, therefore, the probability v being selected in I in a random ordering is the number

of orderings for which v is selected, divided by $(|N(v)|+1)!$. Let σ be an ordering of the vertices of V, such that v is selected in I in the ordering σ. Given a vertex $v \in V$, consider now $L_v := (V(L_v), E(L_v))$, the $(k-1)$-uniform *link hypergraph* on the neighbourhood of v, defined as follows:

$$V(L_v) := N(v), \text{ and}$$
$$E(L_v) := \{S \subset V(L_v) : S \cup \{v\} \in E\},$$

i.e., the vertices are the neighbours of v, and the edges are those edges of the original hypergraph H which contained v, but with v removed. Clearly $|E(L_v)| = d_H(v)$. Let $F \subset V(L_v)$ be

$$F := \{u \in N(v) : \sigma(u) < \sigma(v)\},$$

i.e., the vertices in the neighbourhood of v which occur before v in the ordering σ. We want $L_v(F)$ to have at most d hyperedges. The vertices occurring before v can be ordered arbitrarily amongst themselves, and similarly for the vertices occuring after v. So we get that the probability that v is selected in I is given by:

$$\begin{aligned}\Pr[v \in I] &= \sum_{J \subset V(L_v):\ |E(L_v(J))| \le d} \frac{(|J|)!\,(|V(L_v)|-|J|)!}{(|V(L_v)|+1)!} \\ &= \frac{1}{|V(L_v)|+1} \sum_{J \subset V(L_v):\ |E(L_v(J))| \le d} \frac{1}{\binom{|V(L_v)|}{|J|}}\end{aligned}$$

For $k = 2$, the link hypergraph is a 1-graph i.e. a set of vertices, each vertex being a 1-edge. Hence the summation in the RHS evaluates to $d+1$ (counting 1 for each case when there are exactly $0, 1, \ldots, d$ vertices before v, in the random ordering). Therefore

$$\begin{aligned}\mathrm{E}[|I|] &= \sum_{v \in V} \Pr[v \in I] \\ &= \sum_{v \in V} \min\left\{1, \frac{d+1}{d(v)+1}\right\},\end{aligned}$$

and we get the theorem of Alon, Kahn and Seymour (Theorem 1.3).

For general k-uniform hypergraphs, observe that if $d_H(v) \le d$, then $\Pr[v \in I] = 1$. However, if $d_H(v) > d$, then we need to look at the link hypergraph which can be an arbitrary $k-1$-uniform hypergraph. In this case, we shall prove the following general lemma, (which may be of independent interest).

Lemma 3.3. *For each k-uniform hypergraph $H = (V, E)$, such that $|V| = n$, $|E| = m$, we have*

$$\sum_{J \subset V(H):\, |E(J)| \le a} \frac{1}{\binom{n}{|J|}} \ge c'_k n \Big(\frac{a+1}{m}\Big)^{1/k}.$$

where $c'_k = 2^{-(1+1/k)}$.

Indeed, we get that the probability that v is selected in I is given by:

$$\begin{aligned}
\Pr[v \in I] &= \frac{1}{|V(L_v)|+1} \sum_{J \subset V(L_v):\ |E(L_v(J))| \le d} \frac{1}{\binom{|V(L_v)|}{|J|}} \\
&\ge \frac{c'_{k-1}\,|V(L_v)|}{|V(L_v)|+1} \times \frac{(d+1)^{1/(k-1)}}{|E(L_v)|^{1/(k-1)}} \qquad \text{(from Lemma 3.3)} \\
&\ge \frac{c'_{k-1}\,|V(L_v)|}{|V(L_v)|+1} \times \frac{(d+1)^{1/(k-1)}}{d_H(v)^{1/(k-1)}} \qquad \text{(as } |E(L_v)| = d_H(v)\text{)} \\
&\ge c_k \left(\frac{d+1}{d_H(v)+1}\right)^{1/(k-1)}, \qquad (5)
\end{aligned}$$

where

$$c_k = 2^{-\left(1+\frac{1}{k-1}\right)}\left(1-\frac{1}{k}\right) = 2^{-(1+o_k(1))}.$$

Note that Inequality (5) follows from the fact that since $d_H(v) > d \ge 0$, we must have at least $k-1$ vertices in the hypergraph L_v, i.e., $|V(L_v)| \ge k-1$. □

It only remains to prove Lemma 3.3, which we will prove using a *local density* argument.

Proof of Lemma 3.3. For all $1 \le s \le n$, we define

$$\rho_s := \mathrm{E}_{|S|=s}\left[|E(H(S))|\right] = \frac{\sum_{S \subseteq V,\, |S|=s} E(H(S))}{\binom{n}{s}}.$$

Note that the expectation is taken over all subsets of V of size s, and

$$E(H(S)) = \{e \in E(H) \,:\, e \subseteq S\}.$$

Counting the number of pairs (e, S), where $e \in E(H)$, and $S \subset V : |S| = s, e \in S$, in two ways, we get the average local density of sets of size s is

$$\rho_s = \frac{m\binom{n-k}{s-k}}{\binom{n}{s}} = \frac{m(s)_k}{(n)_k} \le \frac{ms^k}{n^k}.$$

(Here $\binom{a}{b} := 0$ if $b < 0$). This is as follows: let

$$z := \#\left\{(e,S) : e \in E, S \in \binom{V}{s}, e \subset S\right\}.$$

Then each of the $\binom{n}{s}$ sets of size s contributes, on average, ρ_s-many entries to z. On the other hand, each edge $e \in E(H)$ belongs to $\binom{n-k}{s-k}$-many sets of size s. Equating the two summations gives the claimed average local density.

Now, we use the above observation to prove the lemma. Partition the summands on the LHS into n parts, depending on the size of the set J

(i.e. the number of neighbouring vertices which precede v in the random ordering):

$$\sum_{i=1}^{n}\left(\sum_{J\in\binom{V}{i}:|E(H(J))|\leq a}\frac{1}{\binom{n}{i}}\right) \quad .$$

When $i<k$, it is easy to see that the inner summation is 1. The main idea of the proof is the following: first, observe that for any $i\in[n]$, the inner summation is just the probability that a randomly picked set of exactly i vertices has fewer than $a+1$ edges. Next, the expected number of edges in a randomly chosen i-vertex subset, is upper bounded by ρ_i, and for small enough i, is much smaller than $(a+1)/2$. So the probability that a random i-set contains more than twice the expected number, is at most half. So for all such i, the contribution to the outer sum is at least $1/2$. The number of such terms in the outer sum then gives the claimed lower bound.

Formally, let X_i be a random variable giving the number of edges contained in a randomly chosen set on i vertices. By Markov's inequality:

$$\Pr\left[X_i\geq 2.\mathrm{E}\left[X_i\right]\right] \quad\leq\quad \frac{1}{2}.$$

We have that $\mathrm{E}\left[X_i\right]=\rho_i$. Therefore, the LHS becomes:

$$\sum_{i=1}^{n}\Pr\left[X_i<a+1\right] \quad=\quad 1-\sum_{i=1}^{n}\Pr\left[X_i\geq a+1\right]$$

With foresight, we split the above sum into two parts, when $i\leq t^*:=\frac{n(a+1)^{1/k}}{(2m)^{1/k}}$, and when $i>t^*$. Observe that when $i\leq t^*$, we have that $\frac{mi^k}{n^k}\leq\frac{a+1}{2}$. We get

$$\begin{aligned}
\sum_{i=1}^{n}\Pr\left[X_i<a+1\right] &= \sum_{i=1}^{t^*}\Pr\left[X_i<a+1\right]+\sum_{i>t^*}\Pr\left[X_i<a+1\right]\\
&\geq \sum_{i=1}^{t^*}\Big(1-\frac{\mathrm{E}\left[X_i\right]}{a+1}\Big)\\
&= \sum_{i=1}^{t^*}\Big(1-\frac{m(i)_k}{(a+1)(n)_k}\Big)\;\geq\;\sum_{i=1}^{t^*}\big(1-\frac{mi^k}{(a+1)n^k}\big)\\
&\geq \sum_{i=1}^{t^*}\big(1-\frac{1}{2}\big)\;\geq\;\frac{n(a+1)^{1/k}}{2(2m)^{1/k}}.
\end{aligned}$$

where in the second step we used Markov's inequality on the first summation, and in the penultimate step we used the observation on t^* noted above. This completes the proof of Lemma 3.3.

□

Remark 3.4. For simplicity of exposition we did not try to optimize c_k in the proof of Theorem 1.5. We observe that

$$c_k = 2^{-\left(1+\frac{1}{k-1}\right)}\left(1-\frac{1}{k}\right) \geq \frac{1}{8},$$

for all $k \geq 2$, and $c_k \to \frac{1}{2}$ as $k \to \infty$.

4. Generalized degeneracy

In a recent paper Zaker [Zak13] generalized the definition of definition of degeneracy to arbitrary function over the vertex set of graphs taking values from non-negative integers. This is a strict generalization of degeneracy to graphs, and in this section we will now show that techniques from this paper can be directly used to get asymptotically bounds for this definition as well.

Let H be a k-uniform hypergraph with $|V(H)| = n$, and let $\kappa : V(H) \to \mathbb{N} \cup \{0\}$ be a non-negative integer assignment to the vertices of H. We will say H is κ-degenerate if the vertices of H can be ordered as $v_1, \ldots, v_n$ such that the degree of vertex v_i in the hypergraph $H_i = H(\{1, \ldots, i\})$ is at most $\kappa(v_i)$.

Let H be k-uniform hypergraph, and let $\kappa : V(H) \to \mathbb{N} \cup \{0\}$, then

$$\alpha_{k,\kappa}(H) = \max\left\{|I| \,:\, I \subseteq V(H),\, H(I) \text{ is } \eta\text{-degenerate where } \eta = \kappa_{|_I}.\right\}$$

Directly applying the random permutation technique from Section 3 will give us the following result. We have included the proof of completeness.

Theorem 4.1. *For all graphs $G = (V, E)$ and $\kappa : V(H) \to \mathbb{N} \cup \{0\}$, we have exists an induced subgraph H of G with*

$$\alpha_{k,\kappa}(G) = \sum_{v \in V} \min\left\{1, \frac{\kappa(v)+1}{d_G(v)+1}\right\}.$$

Moreover there exists a randomized polynomial time algorithm that finds in G an induced subgraph H which is η-degenerate, where $\eta = \kappa_{|_{V(H)}}$, and the expected size of $|V(H)|$ is $\sum_{v \in V} \min\left\{1, \frac{\kappa(v)+1}{d_G(v)+1}\right\}$.

Proof. Let σ be a random ordering of the vertex set V. Without loss of generality set $\sigma := v_1 < v_2 < \cdots < v_n$. The vertex set of V is processed in the increasing order with respect to σ starting with v_1. We will include a vertex v_i in the set I if the number of neighbors of v appearing before v in the random ordering σ, i.e., if

$$|N(v_i) \cap \{v_1, \ldots, v_{i-1}\}| \leq \kappa(v_i)$$

then $v_i \in I$. The output of this procedure $H = G(I)$. Observe that H is η-degenerate where $\eta = \kappa_{|_I}$. So the only thing left to prove is the lower bound

on $\mathrm{E}_\sigma[|I|]$. Observe that if $d_G(v_i) \le \kappa(v_i)$ then $\Pr[v_i \in I] = 1$, otherwise

$$\Pr[v_i \in I] = \sum_{i=0}^{\kappa(v_i)} \frac{\binom{d_G(v_i)}{j} j!\,(d_G(v_i)-j)!}{(d_G(v_i)+1)!} = \frac{\kappa(v_i)+1}{d_G(v_i)+1}.$$

Therefore

$$\mathrm{E}_\sigma[|I|] = \sum_{i=1}^{n} \Pr[v_i \in I] = \sum_{i=1}^{n} \min\left\{1,\ \frac{\kappa(v_i)+1}{d_G(v_i)+1}\right\}$$

This completes the proof. □

Using the inductive argument in Section 2 we will get the following generalization of Theorem 1.4

Theorem 4.2. *Let $G = (V, E)$ be a linear k-uniform hypergraph, and $\kappa : V \to \mathbb{N} \cup \{0\}$. Then*

$$\alpha_{k,\kappa}(G) \ge w(G,\kappa) := \sum_{v \in V} w_G(v,\kappa), \tag{6}$$

where

$$w_G(v,\kappa) = \begin{cases} 1 & \text{if } d_G(v) \le \kappa(v) \\ \frac{1}{1+(t(\kappa(v)+1))^{-1}} \frac{\binom{d_G(v)}{\kappa(v)+1}}{\binom{d_G(v)+1/t}{d_G(v)-\kappa(v)-1}} & \text{if } d_G(v) > \kappa(v). \end{cases} \tag{7}$$

For general k-uniform hypergraphs, using random permutation and the proof technique from Section 3, we can prove the following generalization of Theorem 1.5.

Theorem 4.3. *Let $G = (V, E)$ be a k-uniform hypergraph, and $\kappa : V \to \mathbb{N} \cup \{0\}$. Then*

$$\alpha_{k,\kappa}(G) \ge \sum_{v \in V} \min\left\{1,\ c_k \left(\frac{\kappa(v)+1}{d_G(v)+1}\right)^{1/t}\right\}, \tag{8}$$

where $t = k-1$ and $c_k = 2^{-\left(1+\frac{1}{(k-1)}\right)}\left(1-\frac{1}{k}\right)$. There exists a randomized algorithm that can extract a d-degenerate set of above size in expectation.

5. Applications in Discrete Geometry

In this section, as an application of Theorem 1.5 we will prove several generalizations of a result of Payne and Wood [PW13] in incidence geometry, on the maximum size of a subset, of an n point set in the plane, such that no three points in the subset are collinear.

Lemma 5.1. *(1) Let P be a set of n points in the plane such that for any line l in the plane $|l \cap P| \le \ell$. For $d \le O(n \log \ell + \ell^2)$ there*

exists a subset $S \subseteq P$ with at most $d|S|$ collinear triples in S and

$$|S| = \Omega\left(\sqrt{\frac{d\,n^2}{n\log\ell + \ell^2}}\right).$$

And if $\ell \leq O(\sqrt{n})$, then

$$|S| = \Omega\left(\sqrt{\frac{d\,n}{\log\ell}}\right).$$

(2) *Let P be a set of n points in the plane such that for any line l in the plane $|l \cap P| \leq \ell$. Let $k \geq 4$ be a constant and $d \leq O(\ell^{k-3}n + \ell^{k-1})$. Then there exists a subset $S \subseteq P$ of size*

$$\Omega\left(n\left(\frac{d}{\ell^{k-3}n + \ell^{k-1}}\right)^{1/(k-1)}\right)$$

such that S has at most $d|S|$ collinear k-tuples in S. And if $\ell \leq O(\sqrt{n})$, then

$$|S| = \Omega\left(\left(\frac{n^{k-2}d}{\ell^{k-3}}\right)^{1/(k-1)}\right).$$

The following result of Payne and Wood [PW13], proved using Szemerédi-Trotter theorem [ST83] on incidence geometry, will be used to prove Lemma 5.1.

Lemma 5.2 ([PW13])**.** (1) *Let P be a set of n points in the plane such that for any line l in the plane $|l \cap P| \leq \ell$. Then the number of collinear 3-tuples in P is at most $O(n^2\log\ell + n\ell^2)$.*

(2) *Let P be a set of n points in the plane such that for any line l in the plane $|l \cap P| \leq \ell$. Then, for $k \geq 4$, the number of collinear k-tuples in P is at most $O(\ell^{k-3}n^2 + \ell^{k-1}n)$.*

Proof of Lemma 5.1. Let H be a k-uniform hypergraph with $V(H) = P$, and $\{p_1, \ldots, p_k\} \in E(H)$ if there exists a line a line l in the plane with $\{p_1, \ldots, p_k\} \in l$. Lemma 5.2 bounds the size of $E(H)$. The result now follows directly from Theorem 1.5. □

In a followup work Cardinal, Tóth and Wood [CTW16], using Elekes and Tóth's [ET05, Theorem 2.3] generalisation to hyperplane point incidences in $\mathbb{R}^d$ of Szemeredi-Trotter's theorem [ST83] on incidence geometry, proved the following theorem.

Lemma 5.3 (Lemma 4.5 [CTW16])**.** *Let P be a set of n-points in $\mathbb{R}^m$ such that for any hyperplane $\mathcal{H}$ in $\mathbb{R}^m$, we have $|P \cap \mathcal{H}| \leq \ell$ where $\ell = O(\sqrt{n})$. Then the number of cohyperplanar $(m+1)$-tuples in P is at most $O(n^m \log\ell)$.*

As in the proof of Lemma 5.1 using the above result together with Theorem 1.5 implies the following result.

Lemma 5.4. *Let P be a set of n points in $\mathbb{R}^m$ such that for any hyperplane $\mathcal{H}$ in $\mathbb{R}^m$ we have $|P \cap \mathcal{H}| \leq \ell$ where $\ell = O(\sqrt{n})$. Then there exists $S \subset P$ of size*

$$\Omega\left(\left(\frac{dn}{\log \ell}\right)^{1/m}\right)$$

such that S has at most $d|S|$ cohyperplanar $(m+1)$-tuples in S.

These geometric results can be easily extended to the polynomial settings. Let $\mathcal{F} \subset \mathbb{R}[X_1, \ldots, X_m]$ be a family of non-zero polynomial functions with real coefficients in m variables statisfying the following property: every $f \in \mathcal{F}$ can be written as a linear combination of the polynomials $f_0(X), \ldots, f_b(X)$ where $f_0(X) = 1$ and the polynomials $f_i(X)$ are linearly independent in $\mathbb{R}[X_1, \ldots, X_m]$. Now a point set P in $\mathbb{R}^m$, with $|P| > b$, is said to be in *general position* if no more than b points of P lies on a zero set of a polynomial in $\mathcal{F}$. In this situation we are interested in extracting a large subset S of P such that number of $(b+1)$-tuples in S which lie on a zero set of at least one polynomial in $\mathcal{F}$ is at most $d|S|$.

First observe that this definition captures many notions of general position with respect to algebraic and geometric objects of bounded complexity.

Remark 5.5.
(1) **Family of spheres in $\mathbb{R}^m$:** This set is generated by the polynomials $\left\{\sum_{i=1}^m X_i^2, X_1, \ldots, X_m, 1\right\}$. A set P in $\mathbb{R}^m$ is said to be in general position with respect to spheres if no more than $m+1$ points of P lie on any given sphere. This notion of general position is extremely important in computational geometry both from theoretical and practical sides. For example, see [BCKO08, EM90, ES95, ES97, BDG13].
(2) **Family of hyperplanes in $\mathbb{R}^m$:** This family is a linear combination of the polynomials $\{X_1, \ldots, X_m, 1\}$. A set P in $\mathbb{R}^m$ is said to be in general position with respect to spheres if no more than m points of P lie on any given hyperplane. For example, see [BCKO08, EM90, ES95, ES97, BDG13].
(3) **Polynomials in $\mathbb{R}[X_1, \ldots, X_m]$ with degree bounded by D:** This set is generated by the set $\mathcal{S} = \{X_1^{j_1} \ldots X_m^{j_m} \mid 0 \leq \sum_{i=1}^m j_i \leq D\}$. Observe that $|\mathcal{S}| = \binom{D+m}{m}$. A set P in $\mathbb{R}^m$ is said to be in general position with respect to this family if no more than $\binom{D+m}{m} - 1$ points of P lie on any given non-zero polynomial in $\mathbb{R}[X_1, \ldots, X_m]$ with degree bounded by D. This notion of general position appears in semialgebraic range searching [AMS13].

We can associate with the polynomial family $\mathcal{F}$ the following Veronese mapping:

$$\Phi_{\mathcal{F}} : \mathbb{R}^m \to \mathbb{R}^b, \text{ where } \Phi_{\mathcal{F}}(X) = (f_1(X), \ldots, f_b(X)).$$

It is easy to see that if there exists a set of points S in $\mathbb{R}^m$ and a non-zero polynomial $f(X) \in \mathcal{F}$ such that $S \subseteq Z(f(X))$, where $Z(f(X)) \subseteq \mathbb{R}^m$

denotes the zero set of the polynomial in $\mathbb{R}^m$, then there exists a hyperplane $\mathcal{H}$ in $\mathbb{R}^b$ with $\Phi_{\mathcal{F}}(S) \subset \mathcal{H}$.

The following result is a direct consequence of above construction and Lemma 5.4.

Lemma 5.6. *Let $\mathcal{F} \subseteq \mathbb{R}[X_1, \ldots, X_m]$ be a family of non-zero polynomials obtained by linear combination of non-zero polynomials $f_0(X)$, $f_1(X)$, $\ldots$, $f_b(X)$ where $X = (X_1, \ldots, X_m)$, $f_0(X) = 1$ and the polynomials $f_i(X)$ are linearly independent in $\mathbb{R}[X_1, \ldots, X_m]$. Let P be a n-point set in $\mathbb{R}^m$ such that for all non-zero polynomial $f(X) \in \mathcal{F}$ we have $|Z(f(X)) \cap P| \leq \ell$.*

(1) If $\ell = O(n^{1/3})$, then there exists $S \subset P$ of size

$$\Omega\left(\frac{dn}{\ell \log \ell}\right)$$

such that S has at most $d|S|$ $(b+1)$-tuples in S each of which is a subset of zero set of some polynomials in $\mathcal{F}$.

(2) If $\ell = O(\sqrt{n})$ and the Veronese map $\Phi_{\mathcal{F}} : \mathbb{R}^m \to \mathbb{R}^b$ restricted to the set P is injective, then there exists $S \subset P$ of size

$$\Omega\left(\frac{dn}{\log \ell}\right)$$

such that S has at most $d|S|$ $(b+1)$-tuples in S and each of these $(b+1)$-tuples are a subset of zero sets of some polynomials in $\mathcal{F}$.

6. Conclusion

Our randomized algorithm for finding d-degenerate subgraphs of k-uniform hypergraphs inherits the analysis of Srinivasan and Shachnai [SS04] for independent sets:

(i) The RandPermute algorithm runs in RNC, as long as d is polylogarithmic in the number of vertices and edges.
(ii) Our proof technique generalizes to non-uniform hypergraphs.
(iii) All our results generalize to the vertex-weighted scenario, where we want an induced d-degenerate subgraph of maximum weight.

It is interesting to ask if the RandPermute algorithm can be used to solve the conjecture of Beame and Luby [BL90], which asks whether iterating the RandPermute algorithm always yields a maximal independent set.

Acknowledgements. Kunal Dutta and Arijit Ghosh are supported by European Research Council under Advanced Grant 339025 GUDHI (Algorithmic Foundations of Geometric Understanding in Higher Dimensions) and Ramanujan Fellowship (No. SB/S2/RJN-064/2015) respectively.

Part of this work was done when Kunal Dutta and Arijit Ghosh were Researchers at Max-Planck-Institute for Informatics, Germany, supported by the Indo-German Max Planck Center for Computer Science (IMPECS).

References

[aDMS12] Kunal Dutta amd Dhruv Mubayi and C. R. Subramanian. New Lower Bounds for the Independence Number of Sparse Graphs and Hypergraphs. *SIAM Journal on Discrete Mathematics*, 26(3):1134–1147, 2012.

[AKS80] Miklós Ajtai, János Komlós, and Endre Szemerédi. A Note on Ramsey Numbers. *Journal on Combinatorial Theory, Series A*, 29(3):354–360, 1980.

[AKS87] Noga Alon, Jeff Kahn, and Paul D. Seymour. Large Induced Degenerate Subgraphs. *Graphs and Combinatorics*, 3(1):203–211, 1987.

[AMS13] Pankaj K. Agarwal, Jiří Matoušek, and Micha Sharir. On Range Searching with Semialgebraic Sets. II. *SIAM Journal on Computing*, 42(6):2039–2062, 2013.

[AS08] Noga Alon and Joel H. Spencer. *The Probabilistic Method*. Wiley International, 2008.

[BCKO08] Mark de Berg, Otfried Cheong, Marc van Kreveld, and Mark Overmars. *Computational Geometry: Algorithms and Applications*. Springer-Verlag TELOS, Santa Clara, CA, USA, 3rd ed. edition, 2008.

[BDG13] Jean-Daniel Boissonnat, Ramsay Dyer, and Arijit Ghosh. The stability of Delaunay Triangulations. *International Journal of Computational Geometry & Applications*, 23(4-5):303–334, 2013.

[BL90] Paul Beame and Michael Luby. Parallel Search for Maximal Independence Given Minimal Dependence. In *Proceedings of the 1st Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA 1990, pages 212–218, 1990.

[Car79] Yair Caro. New results on the independence number. Technical report, Tel Aviv University, 1979.

[CT91] Yair Caro and Zsolt Tuza. Improved Lower Bounds on k-Independence. *Journal of Graph Theory*, 15(1):99–107, 1991.

[CTW16] Jean Cardinal, Csaba D. Tóth, and David R. Wood. General position subsets and independent hyperplanes in d-space. *Journal of Geometry*, pages 1–11, 2016.

[DLR95] Richard A. Duke, Hanno Lefmann, and Vojtech Rödl. On Uncrowded Hypergraphs. *Random Structures & Algorithms*, 6(2/3):209–212, 1995.

[EM90] Herbert Edelsbrunner and Ernst P. Mücke. Simulation of simplicity: A technique to cope with degenerate cases in geometric algorithms. *ACM Transactions on Graphics*, 9(1):66–104, 1990.

[ES95] Jeff Erickson and Raimund Seidel. Better Lower Bounds on Detecting Affine and Spherical Degeneracies. *Discrete & Computational Geometry*, 13:41–57, 1995.

[ES97] Jeff Erickson and Raimund Seidel. Erratum to Better Lower Bounds on Detecting Affine and Spherical Degeneracies. *Discrete & Computational Geometry*, 18(2):239–240, 1997.

[ET05] György Elekes and Csaba D. Tóth. Incidences of Not-Too-Degenerate Hyperplanes. In *Proceedings of the 21st Annual ACM Symposium on Computational Geometry*, SoCG 2005, pages 16–21, 2005.

[PP12] Marcin Pilipczuk and Michal Pilipczuk. Finding a Maximum Induced Degenerate Subgraph Faster Than 2^n. In *Proceedings of the 7th International Symposium on Parameterized and Exact Computation*, IPEC 2012, pages 3–12, 2012.

[PW13] Michael S. Payne and David R. Wood. On the General Position Subset Selection Problem. *SIAM Journal on Discrete Mathematics*, 27(4):1727–1733, 2013.

[She83] James B. Shearer. A note on the independence number of triangle-free graphs. *Discrete Mathematics*, 46(1):83–87, 1983.

[She95] James B. Shearer. On the Independence Number of Sparse Graphs. *Random Structures & Algorithms*, 7(3):269–272, 1995.

[Spe72] Joel H. Spencer. Turán's theorem for k-graphs. *Discrete Mathematics*, 2:183–186, 1972.

[SS04] Hadas Shachnai and Aravind Srinivasan. Finding Large Independent Sets in Graphs and Hypergraphs. *SIAM Journal on Discrete Mathematics*, 18(3):488–500, 2004.

[ST83] Endre Szemerédi and William T. Trotter. Extremal problems in discrete geometry. *Combinatorica*, 3(3):381–392, 1983.

[Thi99] Torsten Thiele. A Lower Bound on the Independence Number of Arbitrary Hypergraphs. *Journal of Graph Theory*, 30(3):213–221, 1999.

[Tur41] Paul Turán. On an extremal problem in graph theory [in Hungarian]. *Math Fiz. Lapok*, 48, 1941.

[Wei81] V. K. Wei. A lower bound on the stability number of a simple graph. Technical Memorandum TM 81-11217-9, Bell Laboratories, 1981.

[Zak13] Manouchehr Zaker. Generalized degeneracy, dynamic monopolies and maximum degenerate subgraphs. *Discrete Applied Mathematics*, 161(16–17):2716 – 2723, 2013.