Lecture 15. Kleitman-Winston graph container lemma

Hong Liu

15th June 2021

In this lecture, we will go over the graph container lemma of Kleitman and Winston, and present an application of it on counting intersecting families.

1 Graph container lemma

The following graph containers theorem provides a way to bound the number of independent sets in a graph. Roughly speaking, it states that given a graph with certain local denseness condition (inequality (2) below), we can find a relatively small collection of *containers* (C_i), each having not too large size, such that every independent set of G is contained in one of these containers.

Theorem 1.1. *Let G be a graph on N vertices, let $R, q \in \mathbb{N}$, and take real $0 < \beta < 1$. Then, provided*

$$e^{-\beta q} N \leq R, \tag{1}$$

and, for every subset $S \subset V(G)$ of at least R vertices, we have

$$e_G(S) \geq \beta \binom{|S|}{2}, \tag{2}$$

there is a collection of sets $C_i \subset V(G)$, $1 \leq i \leq \binom{N}{q}$, such that $|C_i| \leq R + q$ for every i and, for every independent set $I \subset V(G)$, there is some i satisfying $I \subset C_i$. In particular, the number of independent sets of size m, for every $m \geq q$, is at most

$$\binom{N}{q}\binom{R}{m-q}.$$

The idea is that when the graph is locally dense, we can group independent sets up efficiently using small fingerprints (the set S in the proof below, which is of size at most q). We need a definition for the proof. Given a graph G and a vertex subset $A \subseteq V(G)$, the *max-degree ordering* of A is the ordering obtained by iteratively picking out highest degree vertices, that is $(v_1, \ldots, v_{|A|})$ such that for each $1 \leq i \leq |A|$, v_i is a maximum-degree vertex in the subgraph of G induced on $A \setminus \{v_1, \ldots, v_{i-1}\}$.

Proof. Let I be an independent set in G. Run the following algorithm.

Input: G and I. Set initially $A = V(G)$ (active set) and $S = \varnothing$ (selected ones). For the i-th iteration, $i = 1, \ldots, q$, do the following:

Step 1. *Select the next I-vertex.* Take the max-degree ordering of A, say $(v_1, \ldots, v_{|A|})$. Let v_{t_i} be the first (i.e. minimum index) vertex lying in I. Move v_{t_i} from A to S.

Step 2. *Update A.* Delete $\{v_1, \ldots, v_{t_i-1}\} \cup (N_G(v_{t_i}) \cap A)$ from A.

Output: $S = (v_{t_1}, \ldots, v_{t_q})$ and A.

Claim 1.2. *The final A is small: $|A| \le R$.*

Proof of claim. This follows from local denseness of G. For each vertex we select in Step 1, A gets shrunk by a constant factor in Step 2.

For each $i \in [q]$, let A_i be the active set prior to the i-th iteration (so $A_1 = V(G)$ and $A_{q+1} = A$). Suppose to the contrary that $|A| > R$. Let $A' = A_i \setminus \{v_1, \dots, v_{t_i - 1}\}$. Then $|A'| \ge |A| > R$ and by (2), $e_G(A') \ge \beta\binom{|A'|}{2}$. Thus, the chosen vertex v_{t_i} has degree at least

$$2e_G(A')/|A'| = \beta(|A'| - 1)$$

in $G[A']$. Consequently, at least

$$t_i + \beta(|A'| - 1) = t_i + \beta((|A_i| - t_i + 1) - 1) = t_i + \beta(|A_i| - t_i) \ge \beta|A_i|$$

are deleted from A_i at the i-th iteration. Hence, the active set shrinks by a factor of $(1 - \beta)$ after each iteration, and so by (1) the final A has size at most

$$(1-\beta)^q \cdot N \le e^{-\beta q} N \le R,$$

a contradiction. ■

Here are some key observations of the algorithm:

- $S \subseteq I$ and $I \setminus S \subseteq A$;
- the final active set $A = A(S)$ depends only on S, not on the independent set I! Indeed, given G, we always know the max-degree ordering, then knowing what S is, we can run the above algorithm to recover A.

Thus, $C := A \cup S \supseteq I$ and by the above claim $|C| \le R + q$. As A is completely determined, the number of such containers C is at most the number of choices of S, which is at most $\binom{N}{q}$. □

2 Counting intersecting families

We combine spectral methods with the graph contain theorem to count intersecting families/hypergraphs. The following result is due to Balogh, Das, Delcourt, Liu and Sharifzadeh.

Theorem 2.1. *For $k \ge 3$ and $n \ge 2k+1$, let $I(n,k)$ denote the number of intersecting k-uniform hypergraphs on $[n]$. Then*

$$I(n,k) = 2^{(1+o(1))\binom{n-1}{k-1}}.$$

For this, we shall formulate intersecting hypergraphs as independent sets in Kneser graphs.

2.1 Supersaturation for Kneser graph

The *Kneser graph* $KG(n,k)$ is a graph with vertex set $\binom{[n]}{k}$ and an edge between vertices $F_1, F_2 \in \binom{[n]}{k}$ if and only if $F_1 \cap F_2 = \varnothing$. This graph has $N = \binom{n}{k}$ vertices and is D-regular, where $D = \binom{n-k}{k}$. Moreover, subsets of vertices of $KG(n,k)$ correspond to k-uniform hypergraphs on $[n]$, and independent sets correspond directly to intersecting hypergraphs. Our problem thus reduces to counting the number of independent sets in $KG(n,k)$.

The supersaturation condition of (2) in Theorem 1.1 requires large vertex subsets to induce subgraphs of positive density. We use spectral methods to show that the Kneser graph satisfies this property. The following version of expander-mixing lemma, due to Alon and Chung, relates the eigenvalues of a graph to its distribution of edges.

Theorem 2.2 (Alon–Chung)**.** *Let G be a D-regular graph on N vertices, and let λ be its minimum eigenvalue. Then for all $S \subseteq V(G)$,*

$$e(G[S]) \geq \frac{D}{2N}|S|^2 + \frac{\lambda}{2N}|S|\,(N - |S|)\,.$$

To employ this result, we require the spectrum of the Kneser graph, which was determined by Lovász in 1979. In particular, the minimum eigenvalue of the Kneser graph $KG(n,k)$ is $\lambda = -\binom{n-k-1}{k-1} = -\frac{k}{n-k}D$. Combined with Theorem 2.2, this gives the following supersaturation bound.

Proposition 2.3. *Given $\varepsilon > 0$, any set S of at least $(1+\varepsilon)\binom{n-1}{k-1}$ vertices in the Kneser graph $KG(n,k)$ induces at least $\left(1 - \frac{1}{1+\varepsilon}\right)\frac{Dn}{N(n-k)}\binom{|S|}{2}$ edges.*

Proof. Given a vertex set S with $|S| \geq (1+\varepsilon)\binom{n-1}{k-1} = (1+\varepsilon)\frac{kN}{n}$, we apply Theorem 2.2 and the fact that $\lambda = -\frac{k}{n-k}D$ to find

$$\begin{aligned} e(G[S]) &\geq \frac{D}{2N}|S|^2 + \frac{\lambda}{2N}|S|\,(N - |S|) \geq \left(\frac{D-\lambda}{N} + \frac{\lambda}{|S|}\right)\binom{|S|}{2} \\ &\geq \left(1 - \frac{1}{1+\varepsilon}\right)\frac{Dn}{N(n-k)}\binom{|S|}{2}. \end{aligned}$$

□

We remark that the bound on size of S is tight, as by Erdős-Ko-Rado, Kneser graph has an independent set of size $\binom{n-1}{k-1}$

We can now use Theorem 1.1 to find a small set of containers for independent sets in the Kneser graph.

Proposition 2.4. *For $\varepsilon > 0$ and $2 \leq k \leq \frac{n-1}{2}$, let $R = (1+\varepsilon)\binom{n-1}{k-1}$ and*

$$q = \frac{1+\varepsilon}{\varepsilon} \cdot \frac{(n-k)\binom{n}{k}}{n\binom{n-k}{k}} \ln \frac{n}{(1+\varepsilon)k}.$$

Then there exist k-uniform hypergraphs $\mathcal{F}_i$ on $[n]$, $1 \leq i \leq \binom{\binom{n}{k}}{q}$, each of size at most $R + q$, such that every intersecting k-uniform hypergraph $\mathcal{F}$ on $[n]$ is a subhypergraph of $\mathcal{F}_i$ for some i.

Proof. We apply Theorem 1.1 to the Kneser graph $KG(n,k)$. By Proposition 2.3, condition (2) is satisfied by taking

$$\beta = \left(1 - \frac{1}{1+\varepsilon}\right)\frac{Dn}{N(n-k)},$$

where $D = \binom{n-k}{k}$ and $N = \binom{n}{k}$. In order to satisfy (1), we take

$$q = \frac{1}{\beta}\ln\frac{N}{R} = \frac{1}{\beta}\ln\frac{n}{(1+\varepsilon)k} = \frac{1+\varepsilon}{\varepsilon} \cdot \frac{(n-k)\binom{n}{k}}{n\binom{n-k}{k}} \ln \frac{n}{(1+\varepsilon)k}.$$

Applying Theorem 1.1, the result follows by taking $\mathcal{F}_i$ to be the hypergraph with edges $C_i \subset \binom{[n]}{k}$, since every intersecting hypergraph is an independent set of $KG(n,k)$. □

2.2 Proof of Theorem 2.1

Since there is an intersecting hypergraph of size $\binom{n-1}{k-1}$, and each of its subhypergraphs is also intersecting, we have a lower bound $\log I(n,k) \geq \binom{n-1}{k-1}$. We therefore need to show that $\log I(n,k) \leq (1+o(1))\binom{n-1}{k-1}$. Using Proposition 2.4, we will show that for any small $\varepsilon > 0$, $\log I(n,k) \leq (1+2\varepsilon)\binom{n-1}{k-1}$, provided $n \geq 2k+1$ is sufficiently large with respect to ε.

We know that every intersecting hypergraph is contained in one of $\binom{N}{q}$ containers, each of size at most $R+q$, where R and q are as in the statement of the proposition. Thus, the total number of intersecting hypergraphs is at most $\binom{N}{q}2^{R+q}$. Therefore, since $N = \binom{n}{k}$,

$$\log I(n,k) \leq R + q + q\log\frac{Ne}{q} = R + q\log\frac{2e\binom{n}{k}}{q}.$$

Because $R = (1+\varepsilon)\binom{n-1}{k-1}$, it is enough to show that $q\log\frac{2e\binom{n}{k}}{q} \leq \varepsilon\binom{n-1}{k-1}$. We have

$$q = \frac{(1+\varepsilon)(n-k)\ln\frac{n}{(1+\varepsilon)k}}{\varepsilon k\binom{n-k}{k}}\binom{n-1}{k-1} \leq \frac{2\frac{n}{k}\ln\frac{n}{k}}{\varepsilon\binom{n-k}{k}}\binom{n-1}{k-1},$$

and, provided $\varepsilon < \frac{1}{20}$,

$$\log\frac{2e\binom{n}{k}}{q} = \log\left(\frac{2\varepsilon en}{(1+\varepsilon)(n-k)\ln\frac{n}{(1+\varepsilon)k}}\cdot\binom{n-k}{k}\right) \leq \log\binom{n-k}{k}.$$

Hence it suffices to have $2\frac{n}{k}\ln\frac{n}{k} \leq \varepsilon^2\binom{n-k}{k}/\log\binom{n-k}{k}$. If $n = 2k+1$, the left-hand side is constant, while the right-hand side is $\Omega(n/\log n)$. If $n \geq 2k+2$, the left-hand side is $O(n\log n)$, while the right-hand side is $\Omega(n^2/\log n)$, and thus the inequality holds for large enough n.

Letting $\varepsilon \to 0$, we have $\log I(n,k) \leq (1+o(1))\binom{n-1}{k-1}$. This completes the proof.

Remark 2.5. The $n \geq 2k+1$ bound in Theorem 2.1 is best possible. When $n = 2k$, the k-sets in $[n]$ come in $\frac{1}{2}\binom{n}{k} = \binom{n-1}{k-1}$ complementary pairs, and a hypergraph is intersecting if and only if it does not contain both edges from a single pair. We thus have $I(n,k) = 3^{\binom{n-1}{k-1}}$ when $n = 2k$. For $n < 2k$, the complete hypergraph $\binom{[n]}{k}$ is itself intersecting, and thus $I(n,k) = 2^{\binom{n}{k}}$.