Extremal and Probabilistic Graph Theory

Instructor: Jie Ma, Scribed by Tianchi Yang and Xinyang Ye

Apr 29th 2020, Wednesday

1 Lec 18. Random Algebraic Constructions

Theorem 1.1 (Bukh, 2015)**.** *For any s, there exists C relevant to s such that* $\mathrm{ex}(n, K_{s,C+1}) = \Theta(n^{2-\frac{1}{s}})$.

In this lecture, we use algebraic construction to prove the theorem. Let q be a prime power, and F_q be the field of order q. Let $s \geq 4$ be fixed and $q \gg s$. Let $d = s^2 - s + 2$, and $n = q^s$.

Definition 1.2. Let $\vec{X} = \{x_1, x_2, ..., x_s\} \in F_q^s$ and $\vec{Y} = \{y_1, y_2, ..., y_s\} \in F_q^s$. Let $\mathcal{P}$ be all polynomials $f(\vec{X}, \vec{Y})$ of degree at most d in each of $\vec{X}$ and $\vec{Y}$, that is,

$$f(\vec{X}, \vec{Y}) = \sum_{(\vec{a},\vec{b})} \alpha_{\vec{a},\vec{b}} \cdot x_1^{a_1} x_2^{a_2} \cdots x_s^{a_s} \cdot y_1^{b_1} y_2^{b_2} \cdots y_s^{b_s},$$

over all possible choices that $\sum_{i\in[s]} a_i \leq d$ and $\sum_{j\in[s]} b_j \leq d$, where $\alpha_{\vec{a},\vec{b}} \in F_q$.

Definition 1.3. For any $f(\vec{X}, \vec{Y}) \in \mathcal{P}$, we can define a bipartite graph G_f on partition (L, R) as follows:

$$L = R = F_q^s, \text{ and } \vec{X} \in L \sim \vec{Y} \in R \text{ if and only if } f(\vec{X}, \vec{Y}) = 0.$$

Then by the linearity of expectation, $E[e(G)] = n^2/q$. The key idea is to choose a polynomial $f \in \mathcal{P}$ randomly at uniform and use it to define a bipartite graph G_f.

Lemma 1.4. *For any $\vec{u}, \vec{v} \in F_q^s$, $Pr[f(\vec{u}, \vec{v}) = 0] = 1/q$.*

Proof. Note that if c is a uniformly random constant in F_q, then $f(\vec{u}, \vec{v})$ and $f(\vec{u}, \vec{v}) + c$ are identically distributed. Since all constant elements of $f \in \mathcal{P}$ are distributed uniformly at random in F_q, then $Pr[f(\vec{u}, \vec{v}) = 0] = Pr[f(\vec{u}, \vec{v}) = 1] = \cdots$. So $Pr[f(\vec{u}, \vec{v}) = 0] = 1/q$. ∎

Fact 1.5 (Sampling conditional probability)**.** *Let A be an event in a probability space: $P(A) = \sum_{events\ B} P[A|B] \cdot P[B]$. If $P[A|B] = a$ for any event B, then $P(A) = a$.*

Lemma 1.6. *Suppose $r, s \leq \min\{\sqrt{q}, d\}$. Let $U \subseteq F_q^s$ and $V \subseteq F_q^s$ be sets with $|U| = s$ and $|V| = r$. Then*

$$Pr[f(\vec{u}, \vec{v}) = 0 \text{ for all } \vec{u} \in U, \text{ and } \vec{v} \in V] = 1/q^{sr}.$$

Proof. Call a set of points in F_q^s simple if the first coordinate of the points are distinct.
(1). First, we give the proof when both U and V are simple. In this case, we decompose $f = g + h$, where h contains the sr monomials $x_1^i y_1^j$ for $i = 0, 1, ..., s-1$ and $j = 0, 1, ..., r-1$, and g is the linear combination of other monomials.

To prove that $Pr[f(\vec{u}, \vec{v}) = 0$ for all $\vec{u} \in U$, and $\vec{v} \in V] = 1/q^{sr}$, it suffices to prove that the system of sr equations $h(\vec{u}, \vec{v}) = -g(\vec{u}, \vec{v})$ for all $\vec{u} \in U, \vec{v} \in V$ has a unique solution when all

$-g(\vec{u},\vec{v})$ are given. Note that $h(\vec{X},\vec{Y}) = \sum_{i<s,j<r} \alpha_{ij} x_1^i y_1^j$ has sr terms and the system consists of sr equations with sr unknown variables $\alpha_{ij}, 0 \le i \le s-1$ and $0 \le j \le r-1$. This is a consequence of the Lagrange interpolation theorem twice:

- The first application gives for all fixed $\vec{u} \in U$, the single-variable polynomials $h_{\vec{u}}(\vec{Y})$ of degree $r-1$ such that $h_{\vec{u}}(\vec{v}) = -g(\vec{u},\vec{v})$ for all $\vec{v} \in V$.
- The second application gives a polynomial $h(\vec{X},\vec{Y}) = \sum_{0\le j\le r-1} a_j(x_1) y_1^j$ such that each of the coefficients of $h(\vec{u},\vec{Y})$ is equal to the respective coefficient of $h_{\vec{u}}(\vec{Y})$ for all $\vec{u} \in U$.

Using this twice, we show the solution is unique.
(2). Now we consider the general U and V. It suffices to find invertible linear transformation T and $S : F_q^s \to F_q^s$ such that TU and SV are simple. Indeed, $\mathcal{P}$ is invariant under the actions of these transformations on the first s variables $\vec{X}$ and then on the latter s variables $\vec{Y}$. Hence, if we array for TU and SV to be the simple, we reduce to (1). To find such $T : F_q^s \to F_q^s$, it suffices to find a linear map $T_1 : F_q^s \to F_q$, that injective on U. We then find an invertible map $T : F_q^s \to F_q^s$, where first coordinate is T_1. To find such a T_1, we pick T_1 uniformly at random among all linear maps $F_q^s \to F_q$. Then for all points $(\vec{u_1},\vec{u_2}) \in U$, $Pr[T_1(\vec{u_1}) = T_1(\vec{u_2})] = 1/q$. So by union bound,

$$Pr[\vec{u_1},\vec{u_2} \in U \ with \ T_1(\vec{u_1}) = T_1(\vec{u_2})] = \frac{1}{q}\binom{|U|}{2} < 1,$$

impling the existence of the desired $T_1 : F_q^s \to F_q$. And the construction for S is similar. ∎

Fix $U \subseteq F_q^s$ with $|U| = s$. We want to count the common neighbours of the vertices in U. We will use the **moments method**. Let $I(\vec{v}) = 1$ if $\vec{v}$ is adjacent to any $\vec{u} \in U$, and otherwise $I(\vec{v}) = 0$. Let $X = |N(U)|$. Then $X = \sum_{\vec{v}} I(\vec{v})$, and

$$E[X^d] = E[(\sum_{\vec{v}\in F_q^s} I(\vec{v}))^d] = \sum_{\vec{v_1},\cdots,\vec{v_d}\in F_q^s} E[I(\vec{v_1})I(\vec{v_2})\cdots I(\vec{v_d})] = \sum_{1\le r\le d} \binom{q^s}{r} q^{-rs} M_r \le \sum_{r\le d} M_r \triangleq M,$$

where M_r is defined to the number of surjective mappings from $[d]$ to $[r]$. By Markov's inequality,

$$Pr(X \ge \lambda) = Pr(X^d \ge \lambda^d) \le \frac{E[X^d]}{\lambda^d} \le \frac{M}{\lambda^d}$$

Lemma 1.7. *For all s, d, there exists a constant C such that if $f_1(\vec{Y}), f_2(\vec{Y}), ..., f_s(\vec{Y})$ are polynomials over $Y \in F_q^s$ of degree at most d, then*

$$\{\vec{y} \in F_q^s : f_1(\vec{y}) = f_2(\vec{y}) = ... = f_s(\vec{y}) = 0\}$$

has size either at most C or at least $q - C\sqrt{q} \ge q/2$.

Remark 1.8. This lemma can be reduced for an important result in algebraic geometry, known as the Lang-Weil Bound (1954). It says that roughly, the number of points in an r-dimensional algebraic variety in F_q^s is around q^r (assuming some irreducibility conditions)

Let X be the number of common neighbours of vectors $\vec{u_1}, \vec{u_2}, ..., \vec{u_s} \in U$, then

$$X = |\{\vec{v} \in F_q^s : \vec{v} \sim \vec{u_i}, i \in [s]\}| = |\{\vec{v} \in F_q^s : f(\vec{u_i}, \vec{v}) = 0, i \in [s]\}|$$
$$= |\{\vec{y} \in F_q^s : f_{\vec{u_1}}(\vec{y}) = f_{\vec{u_2}}(\vec{y}) = ... = f_{\vec{u_s}}(\vec{y}) = 0\}|.$$

By lemma 1.7, if $X > C$, then $X > q/2$ implies

$$Pr(X > C) = Pr(X \geq \frac{q}{2}) \leq \frac{E[X^d]}{(q/2)^d} \leq \frac{M}{(q/2)^d}.$$

So the number of s-subsets in L or in R with more than C common neighbours is at most $2\binom{n}{s}\frac{M}{(q/2)^d} = O(q^{s-2})$ in expectation. Take such a G and remove a vertex from every such s-subset to creat a new graph G'. We see that G' is $K_{s,C+1}$-free, $v(G') \leq 2n$, and

$$e(G') \geq e(G) - |\#s\text{-subsets}| \cdot n \geq \frac{n^2}{q} - O(q^{s-2})n = (1 - o(1))n^{2-\frac{1}{s}}.$$