Triangle-intersecting families of graphs

Yuval Filmus

January 31, 2014

1 Simonovits–Sós conjecture

In 1938, Erdős, Ko and Rado proved the basic result known as the *Erdős–Ko–Rado theorem*: (curiously, the paper [2] was published only in 1961)

Theorem (Erdős–Ko–Rado). Suppose $k \leq n/2$ and $\mathcal{F} \subseteq \binom{[n]}{k}$ is an intersecting family (any two sets intersect). Then $|\mathcal{F}| \leq \binom{n-1}{k-1}$. If $k < n/2$, then this bound is achieved only for dictators (families of the form $\{S \in \binom{[n]}{k} : i \in S\}$).

Their paper opened up an entire field in extremal combinatorics. One of the questions, asked by Simonovits and Sós [5] in 1976, concerned *triangle-intersecting families.* A collection $\mathcal{F} \subseteq 2^{K_n}$ of graphs on n vertices is *triangle-intersecting* if the intersection of any two graphs contains some triangle. It will be convenient to measure such families using the measure $\mu(\mathcal{F}) = |\mathcal{F}|/2^{\binom{n}{2}}$. One way of constructing such a family is a *triangle-junta*: take a fixed triangle and all graphs containing it. Such a family contains 1/8 of the graphs. Simonovits and Sós conjectured that this is the best that can be achieved, and furthermore triangle-juntas are the unique maximizers. Unfortunately, all they could prove was an upper bound of 1/2, which follows from the fact that a graph and its complement cannot both be in the family.

Chung, Graham, Frankl and Shearer [1] were able to prove an upper bound of 1/4, using Shearer's lemma. The lemma states that if you project the family $\mathcal{F}$ into m subsets $X_1, \ldots, X_m$ such that each element is covered exactly k times, then

$$\mu(\mathcal{F}) \leq \sqrt[k]{\mu(\mathcal{F}_1) \cdots \mu(\mathcal{F}_m)},$$

where $\mathcal{F}_i$ is the projection to X_i, and the measure μ is normalized to be a probability measure on each of the sets. The idea is to take as the sets X_i all complements of complete bipartite graphs. For each bipartite graph G, if we project $\mathcal{F}$ to $\overline{G}$ then we get an intersecting family, since every triangle contains an edge outside of G. Therefore $\mu(\mathcal{F}_i) \leq 1/2$, since $\mathcal{F}_i$ cannot contain both a graph and its complement. On the other hand, each edge appears in half the families, so $k = m/2$. Therefore $\mu(\mathcal{F}) \leq ((1/2)^m)^{2/m} = 1/4$.

The proof only used the fact that a triangle is not bipartite. It therefore applies for a larger class of families, *non-bipartite-intersecting* or *odd-cycle-intersecting.* We can also improve on the proof in another respect. Instead of considering *intersecting* families, we can consider *agreeing* families. These are families in which the condition for each pair A, B of sets is applied not to the intersection $A \cap B$ but to the *agreement* $A \triangledown B = \overline{A \triangle B}$, which is the set of positions on which both sets "agree". For any bipartite G, if we project an odd-cycle-agreeing family to $\overline{G}$ then we get an agreeing family, and such families have measure at most 1/2, for the same reason as above. So the bound 1/4 applies even for odd-cycle-agreeing families.

In the rest of the talk, we prove the Simonovits–Sós conjecture for odd-cycle-agreeing families.

2 Hoffman's bound

The basic idea is to use a spectral bound due to Hoffman [3]. The bound, which is a special case of the Lovász bound (better known as the θ function), was devised to bound the size of independent sets in graphs. In our case, the graph is the non-agreement graph of our problem: the vertices are the graphs on n vertices, and the edges connect any two graphs which are not odd-cycle-agreeing. An independent set in this graph is the same as an odd-cycle-agreeing family.

Lemma (Hoffman's bound). Let A be a symmetric matrix indexed by the graphs on n vertices such that (i) $A_{GH} = 0$ whenever G, H are odd-cycle-agreeing, (ii) $A\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ is the constant vector. For every odd-cycle-agreeing family $\mathcal{F}$, $\mu(\mathcal{F}) \leq \frac{-\lambda_{\min}}{1-\lambda_{\min}}$, where $\lambda_{\min}$ is the minimal eigenvalue of A.

Proof. Let f be the characteristic vector of $\mathcal{F}$. We consider f under the inner product $\langle g, h\rangle = \mathbb{E} g(x)h(x)$. Under this inner product, $\langle f, f\rangle = \langle f, \mathbf{1}\rangle = \mu(\mathcal{F})$. Decompose $f = \mu(\mathcal{F})\mathbf{1} + g$. By construction $g \perp \mathbf{1}$, and so $\|g\|^2 = \|f\|^2 - \mu(\mathcal{F})^2\langle \mathbf{1}, \mathbf{1}\rangle = \mu(\mathcal{F}) - \mu(\mathcal{F})^2$. The conditions on A imply that

$$0 = \langle f, Af\rangle = \mu(\mathcal{F})^2\langle \mathbf{1}, \mathbf{1}\rangle + \langle g, Ag\rangle \geq \mu(\mathcal{F})^2 - \lambda_{\min}\|g\|^2.$$

Substituting the value of $\|g\|^2$, we obtain $\mu(\mathcal{F})^2 \leq \lambda_{\min}(\mu(\mathcal{F}) - \mu(\mathcal{F})^2)$, and so $\mu(\mathcal{F}) \leq \lambda_{\min}(1 - \mu(\mathcal{F}))$. The lemma easily follows. □

Hoffman's bound isn't always tight, but in our case it is. How do we come up with the matrix A? The first idea is to use some *symmetry*. If $\mathcal{F}$ is an odd-cycle-agreeing family then so is $\mathcal{F} \oplus G$ given by $(\mathcal{F} \oplus G)(H) = \mathcal{F}(G \oplus H)$. We can do the same operation on the matrix A, by defining $A^{\oplus G}(H, K) = A(H \oplus G, K \oplus G)$. Since $(H \oplus G)\triangledown(K \oplus G) = \overline{(H \oplus G) \oplus (K \oplus G)} = \overline{H \oplus K} = H \triangledown K$, we see that $A^{\oplus G}$ satisfies condition (i) in Hoffman's bound. It is easy to see that condition (ii) is also satisfied, and furthermore $\lambda_{\min}(A^{\oplus G}) = \lambda_{\min}(A)$. We can therefore consider $A' = \mathbb{E}_G A^{\oplus G}$. Clearly $\lambda_{\min}(A') \geq \lambda_{\min}(A)$, and so since $-\lambda_{\min}/(1 - \lambda_{\min} = 1 - 1/(1 - \lambda_{\min})$ is decreasing in $\lambda_{\min}$, replacing A with A' can only result in a better bound. The matrix A', in turn, is symmetric, that is $A'^{\oplus G} = A'$. A straightforward calculation shows that the Fourier characters $\chi_G(H) = (-1)^{|G \cap H|}$ are all eigenvectors of A', and so consistute its eigenvectors (since they form a basis). Summarizing, without loss of generality we can conclude that A has the Fourier characters as eigenvectors.

We can say even more. The space of 2×2 matrices whose eigenvectors are the Fourier characters is spanned by two matrices: the identity matrix $I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and the swapping matrix $X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. As a linear operator, the first matrix leaves its input unaffected, and the second flips it. Taking tensor products, we obtain a basis $B_G(H) = G \oplus H$. When G is the complement of a bipartite graph, B_G satisfies the properties in Hoffman's bound. Condition (ii) is easy to check. To verify condition (i), suppose that H, K are odd-cycle-intersecting. Then $1'_H B_G 1_K = 1'_H 1_{K \oplus G} = [H = K \oplus G] = [H \triangledown K = \overline{G}] = 0$, since $\overline{G}$ is bipartite. Therefore all matrices of the form $\sum'_G \alpha_G B_{\overline{G}}$ satisfy the conditions of Hoffman's bound, where G goes over all bipartite graphs, and the α_G sum to 1. A straightforward inductive argument proves the converse: these are all the matrices satisfying the conditions.

It remains to choose the coefficients α_G. To that end, we should understand what the eigenvalues of $B_{\overline{G}}$ look like. The eigenvalues of I are both 1, while X has eigenvalues $1, -1$ for its eigenvectors $\chi_\emptyset, \chi_{\{}1\}$ (here 1 is a dummy element). The matrix $B_{\overline{G}}$ can be thought of as a tensor product of copies of I and X, where a copy of X is used for each edge in $\overline{G}$. Therefore the eigenvalue corresponding to χ_H is $\lambda_H = (-1)^{|H \cap \overline{G}|} = (-1)^{|H|}(-1)^{|H \cap G|}$. The general matrix therefore has eigenvalues

$$\lambda_H = (-1)^{|H|} \sum_G{}' \alpha_G (-1)^{|H \cap G|}.$$

Call a vector of eigenvalues (or *spectrum*) *admissible* if it can be written in this form, ignoring the condition that the α_G sum to 1. Straightforward induction shows that for each bipartite G and each function $f\colon G \to \mathbb{R}$, the following spectrum is admissible: $(-1)^{|H|} f(H \cap G)$. Conversely, every feasbile spectrum is a linear combination of functions of this form.

At this point, a flash of inspiration is needed. We consider the following process: take a random complete bipartite graph G, and let $q_K(H)$ be the probability that $H \cap G$ is isomorphic to G, and $q_k(H)$ be the probability that $|H \cap G| = k$. By taking the weighted average over all bipartite G, we see that $(-1)^{|H|} q_K(H)$ and $(-1)^{|H|} q_k(H)$ are both admissible spectra. We will be looking for an admissible spectrum of the following form:

$$(-1)^{|H|} \sum_k c_k q_k(H).$$

The intuition here is that for large $|H|$, this spectrum is close to 0, while for small $|H|$, we might have enough degrees of freedom to control its minimum. to that end, we consider the following table:

H	$q_0(H)$	$q_1(H)$	$q_2(H)$	$q_3(H)$	$q_4(H)$
$\emptyset$	1	0	0	0	0
$-$	1/2	1/2	0	0	0
$\wedge$	1/4	1/2	1/4	0	0
$\triangle$	1/4	0	3/4	0	0
F_4	1/16	1/4	3/8	1/4	1/16
K_4^-	1/8	0	1/4	1/2	1/8

Here F_4 is any forest having 4 edges, and K_4^- is the diamond graph. Note that all rows sum to one, so there is no need to have more columns. The spectrum we're looking for must satisfy $\lambda_\emptyset = 1$, and so $c_0 = 1$. We can also deduce other constraints. In order to get a bound of 1/8, the spectrum must satisfy $\lambda_{\min} = -1/7$, and this gives us several inequality constraints. Furthermore, if we plug in a triangle-junta into Hoffman's bound then all the inequalities must be tight. That means that for every non-zero Fourier coefficient in this family, the corresponding eigenvalue must be $\lambda_{\min}$. This gives us more constraints. In this way, we can deduce $c_1 = -5/7$, $c_2 = -1/7$ and $4c_3 + c_4 = 3/7$. This gives us one degree of freedom. We arbitrarily choose $c_4 = 0$ to obtain the simplest possible expression,

$$(-1)^{|H|} \left(\frac{1}{7} q_0(H) - \frac{5}{7} q_1(H) - \frac{1}{7} q_2(H) + \frac{3}{28} q_3(H) \right).$$

A miracle happens and the minimal value of this expression, over all graphs, is $-1/7$. Intuitively, for large $|H|$ this expression is close to zero, while for small $|H|$ we engineered it to obtain the correct eigenvalues. This leaves open the case of medium $|H|$, which must be tediously checked. We conclude that an odd-cycle-agreeing family has measure at most 1/8, proving the Simonovits–Sós conjecture.

Simonovits and Sós conjectured that triangle-juntas are the unique maximal families. In order to prove this, we need to fudge a bit with our spectrum. The problem is that while $\lambda_{\min} = -1/7$, this is obtained on two many eigenvalues: on those corresponding to forests of $1, 2, 4$ edges, triangles and diamonds. Fortunately, we can fix that. Consider the expression

$$(-1)^{|H|} \left(\frac{1}{7} q_0(H) - \frac{5}{7} q_1(H) - \frac{1}{7} q_2(H) + \frac{3}{28} q_3(H) + \frac{2}{119} \sum_F{}' q_F(H) - \frac{2}{119} q_\square(H) \right),$$

where the sum ranges over all forests having exactly 4 edges. Some calculation shows that the minimal eigenvalue is now attained only on forests of 1 or 2 edges and triangles, and furthermore all other eigenvalues are at least $-135/952 > -1/7$. Suppose now that we have an odd-cycle-agreeing family of measure $1/8$. All the inequalities in Hoffman's bound must be tight, and so its Fourier expansion is supported on sets of at most 3 edges. Some simple arguments show that the family must depend on at most 3 edges, and so must be a triangle-semijunta (all graphs which intersect a fixed triangle in a specific way).

One advantage of the spectral approach is that it implies more than just an upper bound and a description of the optimal families: we can also get a stability result, showing that nearly-optimal families are close to optimal families. Consider an odd-cycle-agreeing family $\mathcal{F}$ of measure $1/8 - \epsilon$. Since there is a gap between the minimal eigenvalue and all other ones, an analysis of Hoffman's bound shows that a $1 - O(\epsilon)$ fraction of the Fourier expansion of the characteristic function f of the family lies on the first $3 + 1$ levels. A deep theorem of Kindler and Safra [4] then shows that $\mathcal{F}$ is $O(\epsilon)$-close to a family $\mathcal{G}$ depending on $O(1)$ coordinates. If the family $\mathcal{G}$ is not odd-cycle-aggreeing then consider two non-odd-cycle-agreeing graphs $G, H \in \mathcal{G}$; we can assume that G, H are supported on the $O(1)$ coordinates. For each graph K on the complement of these coordinates, $\mathcal{F}$ can contain at most one of $G \cup K$ and $H \cup \overline{K}$; therefore $\mathcal{F}$ is $\Omega(1)$-far from $\mathcal{G}$, and by assuming that ϵ is small enough, we can rule out this case. There are finitely many odd-cycle-agreeing families on the $O(1)$ coordinates which are not triangle-semijuntas, and so if ϵ is small enough, we can also rule out $\mathcal{G}$ being one of them. We conclude that for ϵ small enough, $\mathcal{F}$ is $O(\epsilon)$-close to a triangle-semijunta. We can drop the assumption on ϵ by adjusting the big O constant, obtaining the complete stability result.

References

[1] Fan R. K. Chung, Ronald L. Graham, Peter Frankl, and James B. Shearer. Some intersection theorems for ordered sets and graphs. *J. Comb. Theory, Ser. A*, 43(1):23–37, 1986.

[2] Paul Erdős, Chao Ko, and Richard Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford (2)*, 12:313–320, 1961.

[3] Alan J. Hoffman. On eigenvalues and colorings of graphs, I. In *Graph Theory and its Applications*, Proc. Advanced Sem., Math. Research Center, Univ. of Wisconsin, Madison, Wis., pages 79–91. Academic Press, New York, 1969.

[4] Guy Kindler and Shmuel Safra. Noise-resistant Boolean functions are juntas. *Manuscript.*

[5] Miklós Simonovits and Vera T. Sós. Graph intersection theorems. In *Proc. Colloq. Combinatorics and Graph Theory*, pages 389–391, Orsay, Paris, 1976.