Model theory and combinatorial geometry, II

Artem Chernikov

UCLA

Model Theory conference

Będlewo, Poland, July 4, 2017

Joint work with David Galvin and Sergei Starchenko, and with Sergei Starchenko.

Zarankiewicz's problem for general graphs

- Let $G = (U, V, I)$ with $I \subseteq U \times V$ be a bipartite graph (U, V infinite).
- For $A \subseteq U, B \subseteq V$, $I(A, B)$ denotes the bipartite graph induced on $A \times B$.
- For $k \in \mathbb{N}$, let $K_{k,k}$ be the complete bipartite graph with each part of size k.

Fact

[Kővári, Sós, Turán, '54] For each $k \in \mathbb{N}$ there is some $c \in \mathbb{R}$ such that: for any bipartite graph G and $A \subseteq U, B \subseteq V$ with $|A| = |B| = n$, if $I(A, B)$ is $K_{k,k}$-free, then $|I(A, B)| \leq cn^{2-\frac{1}{k}}$.

- For simplicity, we will only discuss the *balanced* case $|A| = |B| = n$, most of the results have unbalanced versions with $|A|, |B|$ of different sizes as well.
- [Bohman, Keevash, '10] $\forall k \geq 5$, there exists a bipartite $K_{k,k}$-free graph with $\geq cn^{2-\frac{2}{k+1}}$ edges.

Vapnik-Chervonenkis dimension and density

- Let U be an infinite set, and $\mathcal{F}$ a family of subsets of U.
- For $A \subseteq U$, let $\mathcal{F} \cap A := \{S \cap A : S \in \mathcal{F}\}$.
- Let $\pi_{\mathcal{F}}(n) := \max\{|\mathcal{F} \cap A| : A \subseteq U, |A| = n\}$.
- The *VC-density* of $\mathcal{F}$ is $\mathrm{vc}(\mathcal{F}) := \inf\{r \in \mathbb{R} : \pi_{\mathcal{F}}(n) = O(n^r)\}$, or ∞ if no such r exists.
- Given a relation $I \subseteq U \times V$, we have the family $\mathcal{F}_I := \{I_b : b \in V\}$ of subsets of U, where $I_b := \{a \in U : (a, b) \in I\}$. Let $\mathrm{vc}(I) := \mathrm{vc}(\mathcal{F}_I)$.

Example

1. [Sauer-Shelah lemma] Let $\mathcal{M}$ be NIP. Then for any definable I, $\mathrm{vc}(I) < \infty$.
2. [Aschenbrenner, Dolich, Haskell, Macpherson, Starchenko] Let $\mathcal{M}$ be strongly minimal or *o*-minimal, then for any definable $I(x, y) \subseteq M^{|x|} \times M^d$, $\mathrm{vc}(I) \leq d$.

Better bound for graphs with bounded VC-density

Fact

[Fox, Pach, Sheffer, Suk, Zahl '15] For every d, k there is some constant $c = c(k, d) \in \mathbb{R}$ satisfying the following.
Let $G = (U, V, I)$ be a bipartite graph with $\mathrm{vc}(I) \leq d$. Then for any $A \subseteq U, B \subseteq V$ with $|A| = |B| = n$, if $I(A, B)$ is $K_{k,k}$-free, then $|I(A, B)| \leq cn^{2-\frac{1}{d}}$.

- Conversely, independence of the bounding exponent from k implies that I is NIP.
- In particular, if $d = 2$, the exponent is $\frac{3}{2}$.

Points-lines incidence, char p

- In $\mathcal{K} \models \mathrm{ACF}_p$, we have a matching lower bound:

Example

1. Let $U = V = K^2$.
2. Let $\mathbb{F}_q \subseteq K$ be a finite field, q a power of p.
3. Let $A = (\mathbb{F}_q)^2$ be the set of all points on the plane over $\mathbb{F}_q$.
4. Let B be the set of all lines (i.e. subsets of $\mathbb{F}_q^2$ given by $y = ax + b$, $(a, b) \in (\mathbb{F}_q)^2$).
5. Let $I \subseteq K^2 \times K^2$ be the (definable) incidence relation.
6. Then $\mathrm{vc}(I) = 2$ and I is $K_{2,2}$-free (only one line passes through a given pair of points)
7. We have $|A| = |B| = q^2$ and $|I(A, B)| = q|B| = q^3$.
8. Let $n := q^2$, then $|A| = |B| = n$ and $|I(A, B)| \geq n^{\frac{3}{2}}$.

Points-lines incidence, char 0

- On the other hand, over the reals a better bound holds (optimal up to a constant, by Erdős):

Fact

[Szémeredi-Trotter '83] Let $I \subseteq \mathbb{R}^2 \times \mathbb{R}^2$ be the incidence relation between points and lines on the affine plane over $\mathbb{R}$. Then $|I(A, B)| = O\left(n^{\frac{4}{3}}\right)$.

- Note: $\frac{4}{3} < \frac{3}{2}$.
- In fact, even in ACF_0 the bound is better:

Fact

[Tóth '03] Let $I \subseteq \mathbb{C}^2 \times \mathbb{C}^2$ be the incidence relation between points and lines on the affine plane over $\mathbb{C}$. Then $|I(A, B)| = O\left(n^{\frac{4}{3}}\right)$.

- Reason? ACF_0 is a reduct of a distal theory, while ACF_p is not.
- More precisely, because cutting lemma holds in ACF_0.

o-minimal "Szémeredi-Trotter"

- Generalizing a result of [Fox, Pach, Sheffer, Suk, Zahl '15] in the semialgebraic case, we have e.g.:

Theorem

Let $\mathcal{M}$ be an o-minimal expansion of a field and $I(x, y) \subseteq M^2 \times M^2$ definable. Then for any $k \in \omega$ there is some c satisfying the following.
For any $A, B \subseteq M^2$, if $I(A, B)$ is $K_{k,k}$-free, then $|I(A, B)| \leq cn^{\frac{4}{3}}$.

- Independently, [Basu, Raz]: same conclusion, under a stronger assumption that the whole graph $I\left(M^2, M^2\right)$ is $K_{k,k}$-free. Their proof uses the crossing number inequality, which appears specific to *o*-minimality.

Ingredients of the proof

Theorem

(Distal cutting lemma) Assume $I(x, y)$ admits a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(S)| = O\left(|S|^d\right)$. Then there is a constant c s.t. for any finite $S \subseteq M^{|y|}$ of size n and any real $1 < r < n$, there is a covering $X_1, \ldots, X_t$ of $M^{|x|}$ with $t \leq cr^d$ and each X_i crossed by at most $\frac{n}{r}$ of the sets $\{I(x, b) : b \in S\}$.

Theorem

(Optimal distal cell decomposition) If $\mathcal{M}$ is an o-minimal expansion of a field and $I(x, y)$ with $|x| = 2$ definable. Then $I(x, y)$ admits a distal cell decomposition $\mathcal{T}$ with $|\mathcal{T}(S)| = O\left(|S|^2\right)$ for all finite sets S.

- Combining, every $I \subseteq M^2 \times M^2$ has an r-cutting of quadratic size, and $\mathrm{vc}(I^*) = 2$ by o-minimality.
- Starting with the general incidence bound given by the VC-density in o-minimal structures, recursively can improve it using cutting lemma for a certain careful choice of r.

Generalizing Elekes-Szabo

- Even in a situation without precise bounds, can get something.

Theorem

Let $\mathcal{M}$ be strongly minimal, interpretable in a distal structure, and $I \subseteq M^2 \times M^2$ is $K_{k,2}$-free. Then there is some $\varepsilon > 0$ such that if I is $K_{k,2}$-free, then $|I(A, B)| \leq n^{\frac{3}{2}-\varepsilon}$.

- Or just "finite combinatorial dimension" as in Elekes-Szabo.
- This can be combined with the group configuration theorem [Tao, Hrushovski, Raz-Scharir-Solymosi] to generalize Elekes-Ronyai theorem to strongly minimal theories iterpretable in distal theories.

The exponents

- Given a bipartite graph $G = (U, V, I)$, let $f(n) := \max\{|I(A, B)| : A \subseteq U, B \subseteq V, |A| = |B| = n\}$.

Definition

The *upper density* of G is $\bar{d}(I) := \inf\{c \in \mathbb{R} : f(n) = O(n^c)\}$.

- Note: $\bar{d}(I) \in \{0\} \cup [1, 2]$.
- [Blei, Körner] For any $\alpha \in [1, 2]$, there is some bipartite graph with $\bar{d}(I) = \alpha$ (probabilistic construction).
- What values can $\bar{d}(I)$ take when I is definable in a nice structure? E.g.,
- **Problem**: can $\bar{d}(I)$ be irrational for I definable in an NIP structure?
- [Bukh, Conlon] ($\approx$) If $\mathcal{K}$ is a pseudofinite field, then $\bar{d}(I)$ can be any rational $\alpha \in [1, 2]$.

Intermediate density

- As discussed above, if I is the point-line incidence relation on the affine plane over a field K, then:
- $\bar{d}(I) = \frac{4}{3}$ if $\mathrm{char}(K) = 0$,
- $\bar{d}(I) = \frac{3}{2}$ if $\mathrm{char}(K) = p$.
- Conversely, e.g.

Theorem

Assume that $\mathcal{M}$ is o-minimal and $I \subseteq M^2 \times M^k$ is a definable relation with $\bar{d}(I) \in (1, 2)$. Then $\mathcal{M}$ defines a field.

- Reason: strong bounds in the locally modular case + trichotomy in o-minimal structures.

Locally modular combinatorics, stable case

Definition

Call a structure $\mathcal{M}$ *combinatorially linear* if for every definable $I(x, y)$, $\bar{d}(I) \in \{0, 1, 2\}$.

- By the remark above, a combinatorially linear structure cannot define a field.
- Recall another familiar notion of *geometric* linearity:

Definition

1. A formula $I(x, y)$ is *weakly normal* if $\exists k \in \mathbb{N}$ s.t. the intersection of any k pairwise distinct sets of the form $I_b, b \in M^{|y|}$ is empty.
2. T is 1-based if every formula is a Boolean combination of weakly normal formulas.

- Note: this definition implies stability, and is equivalent to the definition in terms of forking.
- Stable 1-based theories satisfy a *linear* Zarankiewicz bound:

Locally modular combinatorics, stable case

Definition

Call a structure $\mathcal{M}$ *combinatorially linear* if for every definable $I(x, y)$, $\bar{d}(I) \in \{0, 1, 2\}$.

- By the remark above, a combinatorially linear structure cannot define a field.
- Stable 1-based theories satisfy a linear Zarankiewicz bound:

Theorem

Let $\mathcal{M}$ be stable, 1-based. Then for every definable $I(x, y) \subseteq M^{|x|} \times M^{|y|}$ and $k \in \mathbb{N}$, there is some $c \in \mathbb{R}$ satisfying: for any finite A, B, if $I(A, B)$ is $K_{k,k}$-free, then $|I(A, B)| \leq c(|A| + |B|)$.

- In particular, this implies that $\mathcal{M}$ is combinatorially linear.
- **Conjecture:** For any definable $I(x_1, \ldots, x_k)$, $\bar{d}(I) \in \mathbb{N}$.
- **Problem:** characterize combinatorial linearity among stable, or even strongly minimal, structures (Hrushovski's constructions?).

Locally modular combinatorics, *o*-minimal case

- **Conjecture:** Let $\mathcal{M}$ be *o*-minimal, locally modular. Then every definable $I(x, y)$ satisfies a linear Zarankiewicz bound.
- It seems difficult even for $I \subseteq \mathbb{R}^2 \times \mathbb{R}^4$ the incidence relation between points and rectangles on the plane.
- [Discussion with Sheffer] For 2-parametric families on the plane, $|I(A, B)| \leq c(n \log n)$.
- At least,

Theorem

Let $\mathcal{M}$ be o-minimal, locally modular. If $I \subseteq M^2 \times M^d$ is definable and $I\left(M^2, M^d\right)$ is $K_{k,k}$-free, then $|I(A, B)| \leq c(|A| + |B|)$.

Proof

- Reduce counting incidences to a family of curves (otherwise if I is of full dimension, it will contain an infinite box, so certainly $K_{k,k}$).
- Reduces to a normal family of curves by subdividing each curve in the family, and proving for each of these boundedly many families separately.
- Locally modular $\implies$ the family is 1-dimensional, can work with the dual family of subsets of a set of dimension one (easy to count directly).
- Not local enough: unit distance problem!