SOME VC-DENSITY COMPUTATIONS IN SHELAH-SPENCER GRAPHS

ANTON BOBKOV

1. Preliminaries

VC density was introduced in [**?**] by Aschenbrenner, Dolich, Haskell, MacPherson, and Starchenko as a natural notion of dimension for NIP theories. In a NIP theory we can define a VC function

$$\mathrm{vc} : \mathbb{N} \longrightarrow \mathbb{N}$$

Where $vc(n)$ measures complexity of definable sets in an n-dimensional space. Simplest possible behavior is $\mathrm{vc}(n) = n$ for all n. Theories with that property are known to be dp-minimal, i.e. having the smallest possible dp-rank. In general, it is not known whether there can be a dp-minimal theory which doesn't satisfy $\mathrm{vc}(n) = n$.

In this paper, we investigate vc-density of definable sets in Shelah-Spencer structures. We follow notations in [**?**]

2. Definitions

Definition 2.1. A formula $\phi(x, y)$ is called *basic* if

- $\phi(x, y)$ is a minimal chain extension $\{M_i\}_{i\in[0..k]}$ with $M_0 = \{x, y\}$
- $\phi(x, y)$ determines edges and non-edges on $\{x, y\}$.
- there are no edges between x and y.
- all elements of y that are connected to $M_k - \{x, y\}$. (see note 1.2)

Note 2.2. We add the final condition to simplify our analysis. Similar techniques can be used to acquire bounds on formulas not subject to that condition.

Definition 2.3. For a basic formula $\phi(x, y)$ let

- $\dim (M_i/M_{i-1}) = -\epsilon_i(\phi)$.
- if x or y aren't positive, then $\epsilon_L(\phi) = \epsilon_U(\phi) = \infty$. Otherwise
- $\epsilon_L(\phi) = \sum_{[1..k]} \epsilon_i(\phi)$.
- $\epsilon_U(\phi) = \min_{[1..k]} \epsilon_i(\phi)$.
- Let $Y(\phi) = \dim(y)$ considering y as a graph. In particular if $\{y\}$ are disconnected then $Y = |y|$.

Some of our results only apply to a special family of parameter sets

Definition 2.4.

$$\mathscr{A} = \{A \subset \mathcal{U}^y \mid \text{finite, disconnected, strongly embedded}\}$$

3. Lower bound

We work with formulas that are boolean combinations of basic formulas written in disjunctive-conjunctive form. Define dimensions for those inductively.

Definition 3.1 (Negation)**.** If ϕ is a basic formula, then

$$\epsilon_L(\neg\phi) = \epsilon_L(\phi)$$

Definition 3.2 (Conjunction)**.** Take a collection of formulas $\phi_i(x, y)$ where each ϕ_i is positive or negative basic formula. If both positive and negative formulas are present then $\epsilon_L(\phi) = \infty$. If different formulas define graphs for x differently then $\epsilon_L(\phi) = \infty$. Otherwise

$$\epsilon_L(\bigwedge \phi_i) = \sum \epsilon_L(\phi_i)$$

Definition 3.3 (Disjunction)**.** Take a collection of formulas ψ_i where each instance is a conjunction of positive and negative instances of basic formulas.

$$\epsilon_L(\bigvee \psi_i) = \min \epsilon_L(\psi_i)$$

We use the ϵ_L for a lower bound of *vc*-density of a formula.

Here we show that for a formula ϕ we have

$$\mathrm{vc}\,\phi \geq \left\lfloor \frac{Y(\phi)}{\epsilon_L(\phi)} \right\rfloor$$

First we discuss some cases with ϵ_L infinite.

If x or y are not positive they don't have realizations.

Also each basic formula has to describe graph of x. If there are two formulas in the conjunct that disagree on that, then there are no realizations of the conjunction. We ignore those components as well.

Given a disjunction of formulas, choose one with the smallest ϵ_L. Assume that our formula is conjunction of positive instances of basic formulas and we will consider the negative case later.

$$\psi = \bigwedge \phi_j \in \Phi$$

$$\epsilon_L(\psi) = \sum \phi_j \in \Phi$$

Let n be the integer such that $n\epsilon_L(\psi) < Y$ and $(n + 1)\epsilon_L(\psi) > Y$.

Take an abstract realization of y, and label it by b.

Pick parameter set

$$A = \bigcup_{i<N} b_i$$

a disjoint union where each b_i is an ordered tuple of size $|x|$ connected according to ψ. We also require A to be strong.

Fix n arbitrary elements out of b_i, label them a_i.

Fix an individual formula $\phi \in \Phi$, with minimal sequence M_i.

Abstractly adjoin $M_i/\{a_i, b\} = M/\{x, y\}$ for each i. Let $\bar{M}_\phi = \bigcup M_i$ (disjointly).

We can join those for all $\phi \in \Phi$. Let $\bar{M} = \bigcup M_\phi$ (disjointly).

Claim 3.4. $(A \cap \bar{M}) \leq \bar{M}$.

Proof. It's total dimension is $Y - n\epsilon_L(\psi) > 0$ and all subextensions are positive as well. $\square$

Thus a copy of $\bar{M}$ can be embedded over A into our ambient model. Our choice of b_i was arbitrary, so we get $\binom{N}{n}$ choices out of $N|x|$ many elements. Thus we have $O(|A|^n)$ many traces.

Lemma 3.5. *There are arbitrarily large sets with properties of A.*

This shows

$$\operatorname{vc}\phi \geq n = \left\lfloor \frac{Y}{\epsilon_L} \right\rfloor$$

Now consider the case when the chose conjunction consists of negative basic formulas

$$\psi = \bigwedge \neg\phi_i$$

Let

$$\bar{\psi} = \bigwedge \phi_i$$

Do the construction above for $\bar{\psi}$ and suppose its trace is $X \subset A$ for some b. Then over b the same construction gives trace $(A - X)$ for ψ. Thus we get as many traces.

Claim 3.6. *We can find a minimal extension $M/\{x, y\}$ with arbitrarily small dimension.*

This shows that vc function is infinite in Shelah-Spencer random graphs.

$$\operatorname{vc}(n) = \infty$$

4. Upper bound

We compute an upper bound on a collection of formulas $\{\phi_i\}$. This would automatically bound any boolean combinations of such formulas.

We define a parameter controlling an upper bound on the vc-density of the formula. It is defined inductively.

Definition 4.1. Let ϕ be a basic formula with M_i a minimal chain, with ϵ_i corresponding dimensions, and M it's total size.

$$U_\phi = \frac{M}{\min \epsilon_i}$$

Definition 4.2 (Negation)**.** Let ϕ be basic

$$U_{\neg\phi} = U_\phi$$

Definition 4.3 (Conjunction and Disjunction)**.** Let ϕ_{ij} be basic or a negation of a basic formula.

$$\psi = \bigvee \bigwedge \phi_{ij}$$

$$U_\psi = \max U_{\phi_{ij}}$$

Work with a parameter set $A^{|x|}$, with $A \subset M$.
Pick a trace of $\phi(x, y)$ on A by a parameter b. Record $A \cap b$.
Let

$$A_b = \left\{ a \in A^{|x|} \mid \phi(a, b) \right\}$$

Enumerate $A_b = \{a_i\}_{i \leq I}$. This is a trace of ϕ on A.

Definition 4.4. Suppose ϕ is in positive form. A *witness* of $\phi(a_i, b)$ is the union of realizations of all the positive existential formulas.

Let $\bar{A} = A \cup b$ Let $\bar{M}$ be a union of all witnesses for $\phi(a_i, b), i \in I$. We consider several quantities.

Definition 4.5.

$$\begin{aligned} d &= \dim \bar{M}/\bar{A} \\ s &= |\bar{M} - \bar{A}| \\ b &= \partial(\bar{M}, \bar{A}) \end{aligned}$$

Suppose M is arbitrary and $(M_i, a_i b)$ is some chain-minimal extension. Then we claim

Claim 4.6.

$$\begin{aligned} &\left|\partial(M \cup M_i, \bar{A}) - \partial(M, \bar{A})\right| \leq |M_i| \\ &\dim(M \cup M_i \cup \bar{A}/M\bar{A}) > -\epsilon_U \end{aligned}$$

Thus as we consider $\bar{M}$ as an increasing union of witnesses to chain-minimal extensions, we see the extension with the largest ratio can contribute most to the boundary. Thus is our upper bound for the boundary.

Let $M_i/\{a_i, b\}$ be a witness of $\phi(a_i, b)$ for each $i \leq I$. Let $\bar{M} = \bigcup M_i$. Consider $\bar{M}/A$.

Definition 4.7. Define the boundary of C over $A \cup b$

$\partial(C, A \cup b) = \{a \in (A \cup b) \mid \text{there is an edge between } a \text{ and element of } D - (A \cup b)\}$

Let $\bar{M}_i = \bigcup_{j<i} M_j$.

Definition 4.8. $(j-1, j)$ is called a *jump* if some of the following conditions happen

- New vertices are added outside of A i.e.
$$\bar{M}_j - A \neq \bar{M}_{j-1} - A$$
- New vertices are added to the boundary, i.e.
$$\partial(\bar{M}_j, A) \neq \partial(\bar{M}_{j-1}, A)$$

Definition 4.9. We now let m_i count all jumps below i

$$m_i = |\{j < i \mid (j-1, j) \text{ is a jump}\}|$$

Lemma 4.10.

$$\begin{aligned} \dim(\bar{M}_i/(A \cup b)) &\leq -m_i \cdot \epsilon_U \\ |\partial(\bar{M}_i, (A \cup b))| &\leq m_i \cdot |M| \\ |\bar{M}_j - (A \cup b)| &\leq m_i \cdot |M| \end{aligned}$$

Proof. Proceed by induction. Second and third propositions are clear. For the first proposition base case is clear.

Induction step. Suppose $\bar{M}_j \cap (A \cup b) = \bar{M}_{j+1}$ and $\partial(\bar{M}_j, A) = \partial(\bar{M}_{j+1}, A)$. Then $m_i = m_{i+1}$ and the quantities don't change. Thus assume at least one of these equalities fails.

Apply Lemma 4.2 to $\bar{M}_j \cup (A \cup b)$ and $(M_{j+1}, a_{j+1}b)$. There are two options

- $\dim(\bar{M}_{j+1} \cup (A \cup b)/\bar{M}_i \cup (A \cup b)) \leq -\epsilon_U$. This implies the proposition.
- $M_{j+1} \subset \bar{M}_j \cup (A \cup b)$. Then by our assumption it has to be $\partial(\bar{M}_j, A) \neq \partial(\bar{M}_{j+1}, A)$. There are edges between $M_{j+1} \cap (\partial(\bar{M}_{j+1}, A) - \partial(\bar{M}_j, A))$ so they contribute some negative dimension $\leq \epsilon_U$.

□

Let $m = m_I$. Thus we have $\dim(\bar{M}/(A \cup b)) = \leq -m \cdot \epsilon_U$. Thus as A is strong we need $I \cdot \epsilon_U < Y$. Let $W = \frac{|M|Y}{\epsilon_U}$.

$$|\partial(\bar{M}, A)| \leq m \cdot |M| \leq W$$
$$|\bar{M} \cap A| \leq m \cdot |M| \leq W$$

Now, classify every trace by the isomorphism class of $\bar{M} - A \cup \partial(\bar{M}, A)$ and by $\partial(\bar{M}, A)$.

Lemma 4.11. *Suppose we have traces b_1, b_2 with the same components as above. Then $A_{b_1} = A_{b_2}$.*

Consider $\bar{M} - A \cup \partial(\bar{M}, A)$. Number of vertices is $\leq (2W)^2$. Thus number of isomorphism classes $\leq 2^{(2W)^2}$.

Consider $\partial(\bar{M}, A)$. Let $N = |A|$. Order matters, so the total number of choices for it is

$$N \cdot (N-1) \cdot \ldots \cdot (N - W + 1) = \frac{N!}{(N-W)!}$$

Thus the number of possible different traces is bounded by

$$2^{(2W)^2} \cdot \frac{N!}{(N-W)!} = O(N^W)$$

Since choice of A was arbitrary, this gives

$$\operatorname{vc} \phi \leq W = \frac{|M|Y}{\epsilon_U}$$

5. Technical Lemmas

Lemma 5.1. *Suppose we have a set B and a minimal pair (M, A) with $A \subset B$ and $\dim(M/A) = -\epsilon$. Then either $M \subseteq B$ or $\dim((M \cup B)/B) < -\epsilon$.*

Proof. By diamond construction

$$\dim((M \cup B)/B) \leq \dim(M/(M \cap B))$$

and

$$\dim(M/(M\cap B)) = \dim(M/A) - \dim(M/(M\cap B))$$
$$\dim(M/A) = -\epsilon$$
$$\dim(M/(M\cap B)) > 0$$

□

Lemma 5.2. *Suppose we have a set B and a minimal chain M_n with $M_0 \subset B$ and dimensions $-\epsilon_i$. Let ϵ be the minimal of ϵ_i. Then either $M_n \subseteq B$ or $\dim((M_n \cup B)/B) < -\epsilon$.*

Proof. Let $\bar{M}_i = M_i \cup B$

$$\dim(\bar{M}_n/B) = \dim(\bar{M}_n/\bar{M}_{n-1}) + \ldots + \dim(\bar{M}_2/\bar{M}_1) + \dim(\bar{M}_1/B)$$

Either $M_n \subseteq B$ or one of the summands above is nonzero. Apply previous lemma. □

6. Counterexamples

7. Upper bound on $\mathscr{A}$

Let n be the integer such that $n\epsilon_U < Y$ and $(n+1)\epsilon_U > Y$.

Pick a trace of $\phi(x,y)$ on $A^{|x|}$ by a parameter b.

$$B = \left\{a \in A^{|x|} \mid \phi(a,b)\right\}$$

Pick $B' \subset B$, ordered $B' = \{a_i\}_{i\in I}$ such that

$$a_i \cap \bigcup_{j<i} a_j \neq \emptyset$$

This is always possible by starting with B and taking away elements one by one. Call such a set a *generating set* of B.

Let $M_i/\{a_i, b\}$ be a witness of $\phi(a_i, b)$ for each $i \in I$. Let $\bar{M} = \bigcup M_i$. Consider $\bar{M}/A$.

Pick $\bar{M}$ such that $\dim(\bar{M}/A)$ is maximized.

$\bar{M} \cap A \leq \bar{M}$ as A is strong. (Make sure M is not too big!) Let $\bar{A} = A - \{a_i\}_{i\in I}$. Suppose $\bar{A} \cap \bar{M} \neq \emptyset$. Then we can abstractly reembed $\mathcal{M}$ over A such that $\bar{A} \cap \bar{M} = \emptyset$. This would increase the dimension, contradicting maximality. Thus we can assume $A \cap \bar{M} = \{a_i\}_{i\in I}$

Let $\bar{M}_j = \bigcup_{i<j} M_i$.

Lemma 7.1. $\dim(\bar{M}_j/A) \leq j \cdot \epsilon_U$

Proof. Proceed by induction. Base case is clear.

For induction case apply lemma to $\bar{M}_j \cup \{a_j\}$ and $M_j/\{a_j, b\}$. There are two cases

1. (1) $M_j \subset \bar{M}' \cup \{a_j\}$. In this case there are edges between $\{a_j\}$ and M_j that contribute to dimension less than $-\epsilon_U$.
2. (2) Otherwise M_j adds extra dimension less than $-\epsilon_U$

□

Thus we have $\dim(\bar{M}/A) = \dim(\bar{M}_n/A) \leq -\epsilon_U n$.

Thus as A is strong we need $|B'|\epsilon_U < Y$. This gives us $|B'| \leq n$. Finally we need to relate $|B'|$ to $|B|$.

Suppose we have $C \subset A^{|x|}$, finite with $|C| = N$. A generating set for a trace has to have size $\leq n$. Thus there are $\binom{N}{n} \leq N^n$ choices for a generating set. A set generated from set of size n can have at most $(x|n|)^{|x|}$ elements. Thus a given set of size n can generate at most

$$2^{(x|n|)^{|x|}}$$

sets. Thus the number of possible traces on C is bounded above by

$$2^{(x|n|)^{|x|}} \cdot N^n = O(N^n)$$

This bounds the vc-density by n.

$$\mathrm{vc}_{\mathscr{A}}(\phi) \geq \left\lfloor \frac{Y}{\epsilon_U} \right\rfloor$$

References

[1] M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson, S. Starchenko, *Vapnik-Chervonenkis density in some theories without the independence property*, I, preprint (2011)

[2] Michael C. Laskowski, *A simpler axiomatization of the Shelah-Spencer almost sure theories*, Israel J. Math. **161** (2007), 157186. MR MR2350161

E-mail address: bobkov@math.ucla.edu