SOME VC-DENSITY COMPUTATIONS IN SHELAH-SPENCER GRAPHS

ANTON BOBKOV

Abstract. We investigate vc-density in Shelah-Spencer graphs. We provide an upper bound on formula-by-formula basis and show that there isn't a uniform lower bound forcing $\text{vc}(n) = \infty$.

VC-density was studied in [1] by Aschenbrenner, Dolich, Haskell, MacPherson, and Starchenko as a natural notion of dimension for NIP theories. In an NIP theory we can define a vc-function

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

Where $\text{vc}(n)$ measures the worst-case complexity of definable sets in an n-dimensional space. Simplest possible behavior is $\text{vc}(n) = n$ for all n. Theories with the property that $\text{vc}(1) = 1$ 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 $\text{vc}(n) = n$.

In this paper, we investigate vc-density of definable sets in Shelah-Spencer graphs. In our description of Shelah-Spencer graphs we follow closely the treatment in [2]. A Shelah-Spencer graph is a limit of random structures $G(n, n^{-\alpha})$ for an irrational $\alpha \in (0, 1)$. $G(n, n^{-\alpha})$ is a random graph on n vertices with edge probability $n^{-\alpha}$.

Our first result is that in Shelah-Spencer graphs

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

which implies that they are not dp-minimal. Our second result is providing an upper bound on a vc-density for a formula ϕ

$$\mathrm{vc}(\phi) \leq K(\phi)\frac{Y(\phi)}{\epsilon(\phi)}$$

where $K(\phi), Y(\phi), \epsilon(\phi)$ are paramters easily computable from the quantifier free form of ϕ.

Chapter 1 introduces basic facts about VC-dimension and vc-density. More can be found in [1]. Chapter 2 summarizes notation and basic facts concerning Shelah-Spencer graphs. We direct the reader to [2] for a more in-depth treatment. In chapter 3 we introduce some measure of dimension for quantifier free formulas as well as proving some elementary facts about it. Chapter 4 computes a lower bound for vc-density to demonstrate that $\mathrm{vc}(n) = \infty$. Chapter 5 computes an upper bound for vc-density on a formula-by-formula basis.

1. VC-dimension and vc-density

Throughout this section we work with a collection $\mathcal{F}$ of subsets of a set X. We call the pair $(X, \mathcal{F})$ a <u>set system</u>.

Definition 1.1.

- Given a subset A of X, we define the set system $(A, A \cap \mathcal{F})$ where $A \cap \mathcal{F} = \{A \cap F \mid F \in \mathcal{F}\}$.
- For $A \subset X$ we say that $\mathcal{F}$ <u>shatters</u> A if $A \cap \mathcal{F} = \mathcal{P}(A)$ (the power set of A).

Definition 1.2. We say $(X, \mathcal{F})$ has <u>VC-dimension</u> n if the largest subset of X shattered by $\mathcal{F}$ is of size n. If $\mathcal{F}$ shatters arbitrarily large subsets of X, we say that $(X, \mathcal{F})$ has infinite VC-dimension. We denote the VC-dimension of $(X, \mathcal{F})$ by $\mathrm{VC}(X, \mathcal{F})$.

Note 1.3. We may drop X from the $\mathrm{VC}(X, \mathcal{F})$, as the VC-dimension doesn't depend on the base set and is determined by $(\bigcup \mathcal{F}, \mathcal{F})$.

Set systems of finite VC-dimension tend to have good combinatorial properties, and we consider set systems with infinite VC-dimension to be poorly behaved.

Another natural combinatorial notion is that of a dual system:

Definition 1.4. For $a \in X$ define $X_a = \{F \in \mathcal{F} \mid a \in F\}$. Let $\mathcal{F}^* = \{X_a \mid a \in X\}$. We call $(\mathcal{F}, \mathcal{F}^*)$ the dual system of $(X, \mathcal{F})$. The VC-dimension of the dual system of $(X, \mathcal{F})$ is referred to as the dual VC-dimension of $(X, \mathcal{F})$ and denoted by $\mathrm{VC}^*(\mathcal{F})$. (As before, this notion doesn't depend on X.)

Lemma 1.5. *A set system $(X, \mathcal{F})$ has finite VC-dimension if and only if its dual system has finite VC-dimension. More precisely*

$$\mathrm{VC}^*(\mathcal{F}) \leq 2^{1+\mathrm{VC}(\mathcal{F})}.$$

For a more refined notion of complexity of $(X, \mathcal{F})$ we look at the traces of our family on finite sets:

Definition 1.6. Define the shatter function $\pi_{\mathcal{F}} \colon \mathbb{N} \longrightarrow \mathbb{N}$ and the dual shatter function $\pi^*_{\mathcal{F}} \colon \mathbb{N} \longrightarrow \mathbb{N}$ of $\mathcal{F}$ by

$$\pi_{\mathcal{F}}(n) = \max\{|A \cap \mathcal{F}| \mid A \subset X \text{ and } |A| = n\}$$

$$\pi^*_{\mathcal{F}}(n) = \max\{\mathrm{atoms}(B) \mid B \subset \mathcal{F}, |B| = n\}$$

where $\mathrm{atoms}(B)$ = number of atoms in the Boolean algebra of sets generated by B. Note that the dual shatter function is precisely the shatter function of the dual system: $\pi^*_{\mathcal{F}} = \pi_{\mathcal{F}^*}$.

A simple upper bound is $\pi_{\mathcal{F}}(n) \leq 2^n$ (same for the dual). If the VC-dimension of $\mathcal{F}$ is infinite then clearly $\pi_{\mathcal{F}}(n) = 2^n$ for all n. Conversely we have the following remarkable fact:

Theorem 1.7 (Sauer-Shelah '72)**.** *If the set system $(X, \mathcal{F})$ has finite VC-dimension d then $\pi_{\mathcal{F}}(n) \leq \binom{n}{\leq d}$ for all n, where $\binom{n}{\leq d} = \binom{n}{d} + \binom{n}{d-1} + \ldots + \binom{n}{1}$.*

Thus the systems with a finite VC-dimension are precisely the systems where the shatter function grows polynomially. Define the vc-density of $\mathcal{F}$ to quantify the growth of the shatter function of $\mathcal{F}$:

Definition 1.8. Define vc-density and dual vc-density of $\mathcal{F}$ as

$$\mathrm{vc}(\mathcal{F}) = \limsup_{n\to\infty} \frac{\log \pi_{\mathcal{F}}(n)}{\log n} \in \mathbb{R}^{\geq 0} \cup \{+\infty\},$$
$$\mathrm{vc}^*(\mathcal{F}) = \limsup_{n\to\infty} \frac{\log \pi^*_{\mathcal{F}}(n)}{\log n} \in \mathbb{R}^{\geq 0} \cup \{+\infty\}.$$

Generally speaking a shatter function that is bounded by a polynomial doesn't itself have to be a polynomial. Proposition 4.12 in [1] gives an example of a shatter function that grows like $n \log n$ (so it has vc-density 1).

So far the notions that we have defined are purely combinatorial. We now adapt VC-dimension and vc-density to the model theoretic context.

Definition 1.9. Work in a first-order structure M. Fix a finite collection of formulas $\Phi(x, y)$.

- For $\phi(x, y) \in \mathcal{L}(M)$ and $b \in M^{|y|}$ let

$$\phi(M^{|x|}, b) = \{a \in M^{|x|} \mid \phi(a, b)\} \subseteq M^{|x|}.$$

- Let $\Phi(M^{|x|}, M^{|y|}) = \{\phi(M^{|x|}, b) \mid \phi_i \in \Phi, b \in M^{|y|}\} \subseteq \mathcal{P}(M^{|x|})$.
- Let $\mathcal{F}_\Phi = \Phi(M^{|x|}, M^{|y|})$, giving rise to a set system $(M^{|x|}, \mathcal{F}_\Phi)$.
- Define the VC-dimension of Φ, $\mathrm{VC}(\Phi)$ to be the VC-dimension of $(M^{|x|}, \mathcal{F}_\Phi)$, similarly for the dual.
- Define the vc-density of Φ, $\mathrm{vc}(\Phi)$ to be the vc-density of $(M^{|x|}, \mathcal{F}_\Phi)$, similarly for the dual.

We will also refer to the vc-density and VC-dimension of a single formula ϕ viewing it as a one element collection $\Phi = \{\phi\}$.

Counting atoms of a Boolean algebra in a model theoretic setting corresponds to counting types, so it is instructive to rewrite the shatter function in terms of types.

Definition 1.10.

$$\pi^*_\Phi(n) = \max\{\text{number of } \Phi\text{-types over } B \mid B \subset M, |B| = n\}$$

Here a Φ-type over B is a maximal consistent collection of functions of the form $\phi(x, b)$ or $\neg\phi(x, b)$ where $\phi \in \Phi$ and $b \in B$.

Lemma 1.11.

$$\mathrm{vc}^*(\Phi) = \textit{degree of polynomial growth of } \pi^*_\Phi(n) = \limsup_{n\to\infty} \frac{\log \pi^*_\Phi(n)}{\log n}$$

Proof.

$$\begin{aligned}
&\pi^*_{\mathcal{F}_\Phi}(n) \le \pi^*_\Phi(n) \le \pi^*_{\mathcal{F}_\Phi}(|\Phi|n) \\
&\mathrm{vc}^*(\Phi) \le \limsup_{n\to\infty} \frac{\log \pi^*_\Phi(n)}{\log n} \le \limsup_{n\to\infty} \frac{\log \pi^*_{\mathcal{F}_\Phi}(|\Phi|n)}{\log n} = \\
&= \limsup_{n\to\infty} \frac{\log \pi^*_{\mathcal{F}_\Phi}(|\Phi|n)}{\log |\Phi|n} \frac{\log |\Phi|n}{\log n} = \limsup_{n\to\infty} \frac{\log \pi^*_{\mathcal{F}_\Phi}(|\Phi|n)}{\log |\Phi|n} \le \\
&\le \limsup_{n\to\infty} \frac{\log \pi^*_{\mathcal{F}_\Phi}(n)}{\log n} = \mathrm{vc}^*(\Phi)
\end{aligned}$$

□

One can check that the shatter function and hence VC-dimension and vc-density of a formula are elementary notions, so they only depend on the first-order theory of the structure.

NIP theories are a natural context for studying vc-density. In fact we can take the following as the definition of NIP:

Definition 1.12. Define ϕ to be NIP if it has finite VC-dimension. A theory T is NIP if all the formulas are NIP.

In a general combinatorial context for arbitrary set systems, vc-density can be any real number in $0 \cup [1, \infty)$. Less is known if we restrict our attention to NIP theories. Proposition 4.6 in [1] gives examples of formulas that have non-integer rational vc-density in an NIP theory, however it is open whether one can get an irrational vc-density in this model-theoretic setting.

Instead of working with a theory formula by formula, we can look for a uniform bound for all formulas:

Definition 1.13. For a given NIP structure M, define the vc-function

$$\begin{aligned}\mathrm{vc}^M(n) &= \sup\{\mathrm{vc}^*(\phi(x,y)) \mid \phi \in \mathcal{L}(M), |x| = n\} \\ &= \sup\{\mathrm{vc}(\phi(x,y)) \mid \phi \in \mathcal{L}(M), |y| = n\}\end{aligned}$$

As before this definition is elementary, so it only depends on the theory of M. We omit the superscript M if it is understood from the context. One can easily check the following bounds:

Lemma 1.14 (Lemma 3.22 in [1])**.** *We have* $\mathrm{vc}(1) \geq 1$ *and* $\mathrm{vc}(n) \geq n\,\mathrm{vc}(1)$.

However, it is not known whether the second inequality can be strict or even whether $\mathrm{vc}(1) < \infty$ implies $\mathrm{vc}(n) < \infty$.

2. Graph Combinatorics

We denote graph by $\mathcal{A}$, set of its vertices by A.

Definition 2.1. Fix $\alpha \in (0, 1)$, irrational.

- For a finite graph $\mathcal{A}$ let

$$\delta(\mathcal{A}) = |A| - \alpha e(\mathcal{A})$$

  where $e(\mathcal{A})$ is the number of edges in $\mathcal{A}$.
- For finite $\mathcal{A}, \mathcal{B}$ with $\mathcal{A} \subseteq \mathcal{B}$ define $\delta(\mathcal{B}/\mathcal{A}) = \delta(\mathcal{B}) - \delta(\mathcal{A})$.
- We say that $\mathcal{A} \leq \mathcal{B}$ if $\mathcal{A} \subseteq \mathcal{B}$ and $\delta(\mathcal{A}'/\mathcal{B}) > 0$ for all $\mathcal{A} \subseteq \mathcal{A}' \subsetneq \mathcal{B}$.
- We say that finite $\mathcal{A}$ is positive if for all $\mathcal{A}' \subseteq \mathcal{A}$ we have $\delta(\mathcal{A}') \geq 0$.
- We work in theory S_α axiomatized by
  - Every finite substructure is positive.

  - For a model $\mathscr{M}$ given $\mathcal{A} \leq \mathcal{B}$ every embedding $f : \mathcal{A} \longrightarrow \mathscr{M}$ extends to $g : \mathcal{B} \longrightarrow \mathscr{M}$.

- For $\mathcal{A}, \mathcal{B}$ positive, $(\mathcal{A}, \mathcal{B})$ is called a minimal pair if $\mathcal{A} \subseteq \mathcal{B}$, $\delta(\mathcal{B}/\mathcal{A}) < 0$ but $\delta(\mathcal{A}'/\mathcal{A}) \geq 0$ for all proper $\mathcal{A} \subseteq \mathcal{A}' \subsetneq \mathcal{B}$.
- $\langle \mathcal{A}_i \rangle_{i \leq m}$ is called a minimal chain if $(\mathcal{A}_i, \mathcal{A}_i + 1)$ is a minimal pair (for all $i < m$).
- For a positive $\mathcal{A}$ let $\delta_{\mathcal{A}}(\bar{x})$ be the atomic diagram of $\mathcal{A}$. For positive $\mathcal{A} \subset \mathcal{B}$ let

$$\Psi_{\mathcal{A},\mathcal{B}}(\bar{x}) = \delta_{\mathcal{A}}(\bar{x}) \wedge \exists \bar{y}\ \delta_{\mathcal{B}}(\bar{x}, \bar{y}).$$

  Such formula is called a chain-minimal extension formula if in addition we have that there is a minimal chain starting at $\mathcal{A}$ and ending in $\mathcal{B}$. Denote such formulas as $\Psi_{\langle \mathcal{M}_i \rangle}$.

Theorem 2.2 (5.6 in [2])**.** *S_α admits quantifier elimination down to boolean combination of chain-minimal extension formulas.*

3. Basic Definitions and Lemmas

Fix tuples $x = (x_1, \dots x_n), y = (y_1, \dots, y_m)$. We refer to chain-minimal extension formulas as basic formulas. Let $\phi_{\langle \mathcal{M}_i \rangle}(x, y)$ be a basic formula.

Definition 3.1. Define $\mathcal{X}$ to be the graph on vertices $\{x_i\}$ with edges as defined by $\phi_{\langle \mathcal{M}_i \rangle}$. Similarly define $\mathcal{Y}$. We define those abstractly, i.e. on a new set of vertices disjoint from $\mathscr{M}$.

Note that $\mathcal{X}$, $\mathcal{Y}$ are positive as they are subgraphs of $\mathcal{M}_0$. As usual X, Y will refer to vertices of those graphs.

We restrict our attention to formulas that define no edges between X and Y.

Note 3.2. We can handle edges between x and y as separate elements of the minimal chain extension.

Definition 3.3. For a basic formula $\phi = \phi_{\langle \mathcal{M}_i \rangle_{i \leq k}}(x, y)$ let

- $\epsilon_i(\phi) = -\dim(M_i/M_{i-1})$.
- $\epsilon_L(\phi) = \sum_{[1..k]} \epsilon_i(\phi)$.
- $\epsilon_U(\phi) = \min_{[1..k]} \epsilon_i(\phi)$.
- Let $\mathcal{Y}'$ be a subgraph of $\mathcal{Y}$ induced by vertices of $\mathcal{Y}$ that are connected to $M_k - (X \cup Y)$.
- Let $Y(\phi) = \dim(\mathcal{Y}')$. In particular if $\mathcal{Y} = \mathcal{Y}'$ and $\mathcal{Y}$ is disconnected then $Y(\phi)$ is just the arity of the tuple y.

We conclude this section by stating a couple of technical lemmas that will be useful in our proofs later.

Lemma 3.4. *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 3.5. *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. □

Lemma 3.6. *Suppose we have a minimal chain M_n with dimensions $-\epsilon_i$. Let ϵ be the sum of all ϵ_i. Suppose we have some B with $B \subseteq M_n$. Then $\dim B/(M_0 \cap B) \geq -\epsilon$.*

Proof. Let $B_i = B \cap M_i$. We have $\dim B_{i+1}/B_i \geq \dim M_{i+1}/M_i$ by minimality. $\dim B/(M_0 \cap B) = \dim B_n/B_0 = \sum \dim B_{i+1}/B_i \geq -\epsilon$. □

4. Lower bound

As a simplification for our lower bound computation we assume that all the basic formulas involved we have $\mathcal{Y}' = \mathcal{Y}$ (see Definition 3.3).

We work with formulas that are boolean combinations of basic formulas written in disjunctive-conjunctive form. First, we extend our definition of ϵ.

Definition 4.1 (Negation)**.** If ϕ is a basic formula, then define

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

Definition 4.2 (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 $\epsilon_L(\phi) = \infty$. We don't have a lower bound for that case. If different formulas define $\mathcal{X}$ or $\mathcal{Y}$ differently then $\epsilon_L(\phi) = \infty$. In that case of the conflicting definitions would make the formula have no realizations. Otherwise

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

Definition 4.3 (Disjunction)**.** Take a collection of formulas ψ_i where each instance is a conjunction of positive and negative instances of basic formulas that agree on $\mathcal{X}$ and $\mathcal{Y}$.

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

Theorem 4.4. *For a formula ϕ as above*

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

where $Y(\phi)$ is $Y(\psi)$ for ψ one the basic components of ϕ (all basic componenets agree on $\mathcal{Y}$).

Proof. First work with a formula that is a conjunction of positive basic formulas

$$\psi = \bigwedge_{j \leq J} \phi_j.$$

Then as we defined above

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

Let ϕ be one of the basic formulas in ψ with a chain $\langle M_i \rangle_{i \leq k}$. Let $K_\phi = |M_k|$ i.e. the size of the extension. Let K be the largest such size among all ϕ_i.

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

Label $\mathcal{Y}$ by an tuple b.

Pick parameter set $A \subset \mathscr{M}$ such that

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

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

Fix n arbitrary elements out of A, label them a_j.

For each ϕ_i, a_j pick an abstract realization M_{ij} of ϕ_i over (a_j, b) (abstract meaning disjoint from $\mathscr{M}$).

Let $\bar{M}$ be an abstract disjoint union of all those realizations.

Claim 4.5. $(A \cap \bar{M}) \leq \bar{M}$.

Proof. Consider some $(A \cap \bar{M}) \subseteq B \subseteq \bar{M}$. Let $B_{ij} = B \cap M_{ij} \subseteq M_{ij}$. Then B_{ij}'s are disjoint over $\bar{A} = A \cup b$. In particular $\dim B/(\bar{A} \cap B) = \sum \dim B_{ij}/(\bar{A} \cap B_{ij})$. $\dim B_{ij}/\bar{A} \geq -\epsilon_L(\phi_i)$ by Lemma 3.6. Thus $\dim B/(\bar{A} \cap B) \geq -n\epsilon(\psi)$. Thus $\dim B/(A \cap B) \geq \dim(B) - n\epsilon(\psi)$. By construction we have $Y - n\epsilon_L(\psi) > 0$ as needed. □

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

Lemma 4.6. *There are arbitrarily large sets with properties of A.*

Proof. A is positive, as each of its disjoint components is positive. Thus $0 \leq A$. We apply proposition 4.4 in Laskoswki paper, embedding A into our structure $\mathscr{M}$ while avoiding all nonpositive extensions of size at most $N \cdot I \cdot K$. □

This shows

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

Now consider the formula which is a 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.

Finally consider a formula which is a disjunction of formulas considered above. Choose the one with the smallest ϵ_L, this yields the lower bound for the entire formula. □

Claim 4.7 (4.1 in [2])**.** *We can find a minimal extension M with arbitrarily small dimension.*

Corollary 4.8. *This shows that the vc-function is infinite in Shelah-Spencer random graphs.*

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

In particular, this implies that Shelah-Spencer graphs are not dp-minimal.

5. Upper bound

We bound the number of types of some finite collection of formulas $\Psi(\vec{x}, \vec{y}) = \{\phi_i(\vec{x}, \vec{y})\}_{i \in I}$ over a parameter set B of size N, where ϕ_i is a basic formula.

Fix a formula ϕ from our collection. Suppose it defines a minimal chain extension over $\{x, y\}$. Record the size of that extension as $K(\phi)$ and its total dimension $\epsilon(\phi) = \epsilon_U(\phi)$. Define dimension of that formula $D(\phi) = |\vec{y}| \frac{K(\phi)}{\epsilon(\phi)}$ Define dimension of the entire collection as $D(\Psi) = \max_{i \in I} D(\phi_i)$

In general we have parameter set $B \subset \mathscr{M}^{|y|}$, however without loss of generality we may work with a parameter set $B^{|y|}$, with $B \subset \mathscr{M}$.

Let $S = \lfloor D(\Psi) \rfloor$.

For our proof to work we also need B to be S-strong. We can achieve this by taking (the unique) S-strong closure of B. If size of B is N then the size of its closure is $O(N)$. So without loss of generality we can assume that B is S-strong.

Definition 5.1. A <u>witness</u> of a is a union of realizations of the existential formulas $\phi_i(a, b)$ for all i, b so that the formula holds.

Definition 5.2. For sets C, B define the boundary of C over B

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

Definition 5.3. For each a pick some $\bar{M}_a$ to be its witness. Define two quantities

- ∂_a is the boundary $\partial(\bar{M}_a, B \cup a)$
- Suppose G_1, G_2 are some subgraphs of our model and $a_1 \subset G_1, a_2 \subset G_2$ finite tuples of vertices. Call $f\colon (G_1, a_1) \longrightarrow (G_2, a_2)$ a ∂-isomorphism if it is a graph isomorphism, f and f^{-1} are constant on B, and $f(a_1) = a_2$.
- Define $\mathscr{I}_a$ as the ∂-isomorphism class of $(\bar{M}_a, a)$.

Lemma 5.4. *If $\mathscr{I}_{a_1} = \mathscr{I}_{a_2}$ then a_1, a_2 have the same Ψ-type over B.*

Proof. Fix a ∂-isomorphism $f\colon (\bar{M}_{a_1}, a_1) \longrightarrow (\bar{M}_{a_1}, a_2)$. Suppose we have $\phi(a_1, b)$ for some $b \in B$. Pick witness of this existential formula $M_1 \subset \bar{M}_{a_1}$. Then $f(M_1)$ is a witness for $\phi(a_2, b)$. □

Thus to bound the number of traces it is sufficient to bound the number of possibilities for $\mathscr{I}_a$.

Theorem 5.5.

$$|\partial_a| \leq D(\Psi)$$

$$|\bar{M}_b - \bar{A}| \leq D(\Psi)$$

Corollary 5.6.

$$\mathrm{vc}(\phi) \leq K(\phi)\frac{Y(\phi)}{\epsilon(\phi)}$$

Proof. We count possible ∂-isomorphism classes $\mathscr{I}_b$. Let $W = K(\phi)\frac{Y(\phi)}{\epsilon(\phi)}$. If the parameter set A is of size N then there are $\binom{N}{W}$ choices for boundary ∂_b. On top of the boundary there are at most W extra vertices and $(2W)^2$ extra edges. Thus there are at most

$$W \cdot 2^{(2W)^2}$$

configurations up to a graph isomorphism. In total this gives us

$$\binom{N}{W} \cdot W \cdot 2^{(2W)^2} = O(N^W)$$

options for ∂-isomorphism classes. By Lemma 5.4 there are at most $O(N^W)$ many traces, giving the required bound. □

Proof. (of Theorem 5.5) Fix some b-trace A_b. Enumerate $A_b = \{a_1, \ldots, a_I\}$.

Let $M_i/\{a_i, b\}$ be a witness of $\phi(a_i, b)$ for each $i \leq I$. Let $\bar{M}_i = \bigcup_{j<i} M_j$. Let $\bar{M} = \bigcup M_i$, a witness of A_b

Claim 5.7.

$$\left|\partial(M_i M, \bar{A}) - \partial(M, \bar{A})\right| \leq |M_i| = K(\phi)$$

$$\dim(M_i M \bar{A}/M \bar{A}) > -\epsilon(\phi)$$

Definition 5.8. $(j-1, j)$ is called a jump if some of the following conditions happen

- New vertices are added outside of $\bar{A}$ i.e.

$$\bar{M}_j - \bar{A} \neq \bar{M}_{j-1} - \bar{A}$$

- New vertices are added to the boundary, i.e.

$$\partial(\bar{M}_j, \bar{A}) \neq \partial(\bar{M}_{j-1}, \bar{A})$$

Definition 5.9. We now let m_i count all jumps below i

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

Lemma 5.10.

$$\dim(\bar{M}_i/\bar{A}) \leq -m_i \cdot \epsilon(\phi)$$

$$|\partial(\bar{M}_i, \bar{A})| \leq m_i \cdot K(\phi)$$

$$|\bar{M}_j - \bar{A}| \leq m_i \cdot K(\phi)$$

Proof. (of Lemma 5.10) 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 3.5 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 $\leq \epsilon_U$.

This ends the proof for Lemma 5.10. □

(Proof of Theorem 5.5 continued) First part of lemma 5.10 implies that we have $\dim(\bar{M}/\bar{A}) \leq -m_I \cdot \epsilon(\phi)$. The requirement of A to be S-strong forces

$$m_I \cdot \epsilon(\phi) < Y(\phi)$$

$$m_I < \frac{Y(\phi)}{\epsilon(\phi)}$$

Applying the rest of 5.10 gives us

$$|\partial(\bar{M}, A)| \leq m_I \cdot K(\phi) \leq \frac{K(\phi)Y(\phi)}{\epsilon(\phi)}$$

$$|\bar{M} \cap A| \leq m_I \cdot K(\phi) \leq \frac{K(\phi)Y(\phi)}{\epsilon(\phi)}$$

as needed. This ends the proof for Theorem 5.5. □

So far we have computed an upper bound for a single basic formula ϕ.

To bound an arbitrary formula, write it as a boolean combination of basic formulas ϕ_i (via quantifier elimination) It suffices to bound vc-density for collection of formulas $\{\phi_i\}$ to obtain a bound for the original formula.

In general work with a collection of basic formulas $\{\phi_i\}_{i\in I}$. The proof generalizes in a straightforward manner. Instead of $A^{|x|}$ we now work with $A^{|x|} \times I$ separating traces of different formulas. Formula with the largest quantity $Y(\phi)\frac{K(\phi)}{\epsilon(\phi)}$ contributes the most to the vc-density. Thus we have

$$\Phi = \{\phi_i\}_{i\in I}$$

$$\mathrm{vc}(\Phi) \leq \max_{i\in I} Y(\phi_i)\frac{K(\phi_i)}{\epsilon_{\phi_i}}$$

References

[1] M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson, S. Starchenko, *Vapnik-Chervonenkis density in some theories without the independence property*, I, Trans. Amer. Math. Soc. 368 (2016), 5889-5949

[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