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 a formula-by-formula basis and show that there isn't a uniform lower bound, forcing the vc-function to be infinite. In addition we show that Shelah-Spencer graphs do not have a finite dp-rank, in particular they are not dp-minimal.

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

$$\mathrm{vc}^T = \mathrm{vc} : \mathbb{N} \longrightarrow \mathbb{R} \cup \{\infty\}$$

where $\mathrm{vc}(n)$ measures the worst-case complexity of families of definable sets in an n-fold Cartesian power of the underlying set of a model of T (see 1.13 below for a precise definition of vc^T). We always have $\mathrm{vc}(n) \geq n$ for each n, and the simplest possible behavior is $\mathrm{vc}(n) = n$ for all n. Theories with the property that $\mathrm{vc}(1) = 1$ are known to be dp-minimal, i.e., having the smallest possible dp-rank (see Definition 2.1). It is not known whether there can be a dp-minimal theory which doesn't satisfy $\mathrm{vc}(n) = n$ (see [1], diagram in section 5.3).

In this paper, we investigate vc-density of definable sets in Shelah-Spencer graphs. The first major breakthrough in the model-theoretic study for these structures was made in [11]. In our description of Shelah-Spencer graphs we follow closely the treatment in [7]. A Shelah-Spencer graph is a limit of random structures $G(n, n^{-\alpha})$ for an irrational $\alpha \in (0, 1)$. Here $G(n, n^{-\alpha})$ is a random graph on n vertices with edge probability $n^{-\alpha}$. (The model theory of $G(n, n^{-\alpha})$ as $n \longrightarrow \infty$ is much less pleasant if $\alpha \in (0, 1)$ is rational, see [4].)

Our first result is that in Shelah-Spencer graphs

$$\mathrm{vc}(n) = \infty \text{ for each } n.$$

We also show that Shelah-Spencer graphs don't have a finite dp-rank, which in particular implies that they are not dp-minimal. Our second result provides an upper bound on the vc-density of a given formula $\phi(x, y)$:

$$\mathrm{vc}(\phi) \leq D(\phi)$$

where $D(\phi)$ is an explicitly computable expression involving $|y|$ and the number of vertices and edges defined by ϕ. For example let $\phi(x, y)$ be a formula that says that there is an edge between x and y. Our bound gives $\mathrm{vc}(\phi) \leq D(\phi) = \left\lfloor \frac{2}{\alpha} \right\rfloor$. With a more careful computation we can get $\mathrm{vc}(\phi) = \left\lfloor \frac{1}{\alpha} \right\rfloor$ (see 4.9 in [1]).

Section 1 introduces basic facts about VC-dimension and vc-density. More can be found in [1]. Section 2 summarizes notation and basic facts concerning Shelah-Spencer graphs. We direct the reader to [7] for a more in-depth treatment. In Section 3 we introduce key lemmas that will be useful in our proofs. Section 4 computes a lower bound for vc-density to demonstrate that $\mathrm{vc}(n) = \infty$. Here we also do computations involving dp-rank. Section 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 an infinite set X. We call the pair $(X, \mathcal{F})$ a set system.

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 $A \subseteq X$ we say that $\mathcal{F}$ shatters A if $A \cap \mathcal{F} = \mathcal{P}(A)$ (the power set of A).

Definition 1.2. We say $(X, \mathcal{F})$ has VC-dimension 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 notation $\mathrm{VC}(X, \mathcal{F})$, 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 the dual system of a set 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 (see 2.13b in [2])**.** *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 $\pi_{\mathcal{F}} \colon \mathbb{N} \longrightarrow \mathbb{N}$ of $\mathcal{F}$ and the dual shatter function $\pi^*_{\mathcal{F}} \colon \mathbb{N} \longrightarrow \mathbb{N}$ of $\mathcal{F}$ by

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

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

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, see [8], [9])**.** *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. The vc-density of $\mathcal{F}$ quantifies the growth of the shatter function of $\mathcal{F}$:

Definition 1.8. Define the 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 $\Phi(x, y)$ in the language $\mathcal{L}(M)$ of M.

- 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 $\mathrm{VC}(\Phi)$ of Φ to be the VC-dimension of $(M^{|x|}, \mathcal{F}_\Phi)$, similarly for the dual.

- Define the vc-density $\mathrm{vc}(\Phi)$ of Φ 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.

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

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

The functions π^*_Φ and $\pi^*_{\mathcal{F}_\Phi}$ do not have to agree, as one fixes the number of generators of a boolean algebra of sets and the other fixes the size of the parameter set. However, as the following lemma demonstrates, they both give the same asymptotic definition of dual vc-density.

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. With a parameter set B of size n, we get at most $|\Phi|n$ sets $\phi(M^{|x|}, b)$ with $\phi \in \Phi, b \in B$. We check that asymptotically it doesn't matter whether we look at growth of boolean algebra of sets generated by n or by $|\Phi|n$ many sets. We have:

$$\pi^*_{\mathcal{F}_\Phi}(n) \leq \pi^*_\Phi(n) \leq \pi^*_{\mathcal{F}_\Phi}(|\Phi|n).$$

Hence:

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

□

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 M.

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 in a theory T. A theory T is NIP if all the formulas in T are NIP.

In a general combinatorial context (for arbitrary set systems), vc-density can be any real number in $0 \cup [1, \infty)$ (see [3]). 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 <u>vc-function</u>

$$\begin{aligned}
\mathrm{vc}^M(n) &= \sup\{\mathrm{vc}^*(\phi(x,y)) \mid \phi \in \mathcal{L}(M), |x| = n\} \\
&= \sup\{\mathrm{vc}(\phi(x,y)) \mid \phi \in \mathcal{L}(M), |y| = n\} \in \mathbb{R}^{\ge 0} \cup \{+\infty\}.
\end{aligned}$$

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 just whether $\mathrm{vc}(1) < \infty$ implies $\mathrm{vc}(n) < \infty$.

2. REDO!

Dp-rank is a common measure used in study of NIP theories, with dp-minimality being a special case. Those notions originated in [10], and further studied in [6], showing, for example, that dp-rank is additive. Here it is easiest for us to define dp-rank in terms of vc-density over indiscernible sequences.

Definition 2.1. Work in a first-order structure M. Fix a finite collection of formulas $\Phi(x, y)$ in the language $\mathcal{L}(M)$ of M.

- Suppose $A = (a_i)_{i\in\mathbb{N}}$ is an indiscernible sequence with each $a_i \in M^{|x|}$. Let

$$\mathscr{I}(A, \Phi) = \{\phi(\bigcup_{i\in\mathbb{N}} a_i, b) \mid \phi \in \Phi, b \in M^{|y|}\} \subseteq \mathcal{P}(M^{|x|}).$$

  This gives rise to a set system $(M^{|x|}, \mathscr{I}(A, \Phi))$.
- Define

$$\mathrm{vc}_{\mathrm{ind}}(\Phi) = \sup\left\{\mathrm{vc}(\mathscr{I}(A, \Phi)) \mid A = (a_i)_{i\in\mathbb{N}} \text{ is indiscernible}\right\}.$$

- Dp-rank of a theory T is $\leq n$ if $\mathrm{vc}_{\mathrm{ind}}(\phi) \leq n$ for all formulas ϕ.
- A theory T is said to have finite dp-rank if its dp-rank is $\leq n$ for some n.
- A theory T is dp-minimal if its dp-rank ≤ 1.

Refer to [5] for the connection between to the classical definition of dp-rank and the definition given here.

3. Graph Combinatorics

Throughout this paper A, B, C, M (sometimes with decorations) will denote finite graphs, and $\mathbb{D}$ will be used to denote potentially infinite graphs. All graphs are undirected and assymetric. For a graph A the set of its vertices is denoted by $v(A)$, and the set of its edges by $e(A)$. The number of vertices of A will be denoted by $|A|$. Subgraph always means induced subgraph and $A \subseteq \mathbb{D}$ means that A is a subgraph of $\mathbb{D}$. For two subgraphs A, B of a larger graph, the union $A \cup B$ denotes the graph induced by $v(A) \cup v(B)$. Similarly, $A - B$ means a subgraph of A induced by the vertices of $v(A) - v(B)$. For $A \subseteq B \subseteq \mathbb{D}$ and $A \subseteq C \subseteq \mathbb{D}$, graphs B, C are said to be disjoint over A if $v(B) - v(A)$ is disjoint from $v(C) - v(A)$ and there are no edges from $v(B) - v(A)$ to $v(C) - v(A)$ in $\mathbb{D}$.

For the remainder of the paper fix $\alpha \in (0, 1)$, irrational.

Definition 3.1.

- For a graph A let $\dim(A) = |A| - \alpha|e(A)|$. (Note that this may be negative.)
- For A, B with $A \subseteq B$ define $\dim(B/A) = \dim(B) - \dim(A)$.
- We say that $A \leq B$ if $A \subseteq B$ and $\dim(A'/A) > 0$ for all $A \subsetneq A' \subseteq B$.
- Define A to be positive if for all $A' \subseteq A$ we have $\dim(A') \geq 0$.
- We work in theory S_α in the language of graphs axiomatized by the following conditions:
  - Every finite substructure is positive.
  - Given a model $\mathbb{G}$ and graphs $A \leq B$, every embedding $f : A \longrightarrow \mathbb{G}$ extends to an embedding $g : B \longrightarrow \mathbb{G}$.

  (Here an embedding is taken in the model-theoretic sense, of structures in the language of graphs, so each embedding maps edges to edges and nonedges to nonedges.) This theory is complete and stable (see 5.7 and 7.1 in [7]). From now on fix an ambient model $\mathbb{G} \models S_\alpha$. This will be the only infinite graph we work with.

- Given $S \in \mathbb{N}$, a graph $S \subseteq \mathbb{D}$ is called S-strong if for any $R \subseteq T \subseteq \mathbb{D}$ with $|T| - |R| \leq S$ we have $R \leq T$.
- For A, B positive, (A, B) is called a minimal pair if $A \subseteq B$, $\dim(B/A) < 0$ but $\dim(A'/A) \geq 0$ for all proper $A \subseteq A' \subsetneq B$. We call B a minimal extension of A. The dimension of a minimal pair is defined as $|\dim(B/A)|$.
- A sequence $\langle M_i \rangle_{0 \leq i \leq n}$ of finite graphs is called a minimal chain if (M_i, M_{i+1}) is a minimal pair for all $0 \leq i < n$.
- Suppose we have a graph A with vertices $v(A) = \{x_1, \ldots, x_n\}$ with pairwise disjoint x_i. For the variable tuple of vertices $x = (x_1, \ldots, x_n)$ let $\operatorname{diag}_A(x)$ be the atomic diagram of A, i.e. the first-order formula recording whether there is an edge between every pair of vertices. So for a graph $\mathscr{D}$ and a tuple $a = (a_1, \ldots, a_n)$ we have $\mathscr{D} \models \operatorname{diag}_A(a)$ if and only if exists an embedding $f\colon A \longrightarrow \mathscr{D}$ such that $f(x_i) = a_i$.
- Given $A \subseteq B$ let

$$\phi_{A,B}(x) = \operatorname{diag}_A(x) \wedge \exists z \ \operatorname{diag}_B(x, z).$$

  Any graph isomorphic to B is called a witness of $\phi_{A,B}$. Suppose $\mathbb{G} \models \phi_{A,B}(a)$ for some tuple $a = (a_1, \ldots, a_m)$ and we have a finite subgraph $B' \subseteq \mathbb{G}$ with vertices $v(B') = \{b_1, \ldots b_n\}$ such that $b_i = a_i$ with $i = 1, \ldots, m$ and $\mathbb{G} \models \operatorname{diag}_B(b)$. In this case we call such a graph B' a witness of $\phi_{A,B}(a)$.
- A formula $\phi_{A,B}$ is called a basic formula if there is a minimal chain $\langle M_i \rangle_{0 \leq i \leq n}$ such that $A = M_0$ and $B = M_n$. We also denote such a formula by $\phi_{\langle M_i \rangle_{0 \leq i \leq n}}$.

Theorem 3.2 (Quantifier simplication, 5.6 in [7]). *In the theory S_α every formula is equivalent to a boolean combination of basic formulas.*

4. Basic Definitions and Lemmas

We will require the following lemmas from [7]:

Lemma 4.1. *[See 2.3 in [7]] Let $A, B \subseteq \mathbb{D}$. Then*

$$\dim(A \cup B/A) \leq \dim(\mathcal{B}/A \cap B).$$

Moreover,

$$\dim(A \cup B/A) = \dim(\mathcal{B}/A \cap B) - \alpha E,$$

where E is the number of edges connecting the vertices of $B - A$ to the vertices of $A - B$.

Lemma 4.2. *[See 4.1 in [7]] Suppose A is a positive graph, with at least $\lceil 1/\alpha \rceil + 2$ vertices. Then for any $\epsilon > 0$ there exists a graph B such that (A, B) is a minimal pair with dimension $\leq \epsilon$. Moreover, every vertex in A is connected to a vertex in $B - A$.*

Lemma 4.3. *[See 4.4 in [7]] Suppose we have $A \subseteq \mathbb{G}$. Then for any integer $S \geq 0$ there exists an embedding $f \colon A \longrightarrow \mathbb{G}$ such that $f(A)$ is S-strong in $\mathbb{G}$.*

Lemma 4.4. *[See 3.8 in [7]] For all $S > 0$ there exists $M = M(S, \alpha) \in \mathbb{N}$ with the following property. Suppose $A \subseteq \mathbb{G}$. Then there exists B with $A \subseteq B \subseteq \mathbb{G}$ such that B is S-strong in $\mathbb{G}$ and $|B| \leq M|A|$.*

We conclude this section by stating a couple of technical lemmas that will be useful in our proofs later. In those lemmas we work in an ambient graph $\mathbb{D}$; all finite graphs that we are dealing with are subgraphs of $\mathbb{D}$.

Lemma 4.5. *Let B be a graph and (A, M) be a minimal pair with $A \subseteq B$ and $\dim(M/A) = -\epsilon$. Then either $M \subseteq B$ or $\dim(M \cup B/B) < -\epsilon$.*

Proof. By Lemma 4.1 we have

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

and as $A \subseteq M \cap B \subseteq M$ we get

$$\dim(M/A) = \dim(M/M \cap B) + \dim(M \cap B/A).$$

In addition we are given $\dim(M/A) = -\epsilon$. If $M \not\subseteq B$ then $A \subseteq M \cap B \subsetneq M$ and by minimality $\dim(M \cap B/A) > 0$. Combining the inequalities above we obtain the desired result:

$$\dim(M \cup B/B) \leq \dim(M/M \cap B) = \dim(M/A) - \dim(M \cap B/A) < -\epsilon.$$

$\square$

Lemma 4.6. *Let B be a graph and $\langle M_i \rangle_{0 \leq i \leq n}$ be a minimal chain with dimensions*

$$\dim(M_{i+1}/M_i) = -\epsilon_i$$

and $M_0 \subseteq B$. Let $\epsilon = \min_{0 \leq i \leq n} \epsilon_i$. Then either $M_n \subseteq B$ or $\dim((M_n \cup B)/B) < -\epsilon$.

Proof. Let $\overline{M}_i = M_i \cup B$. Then:

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

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

Lemma 4.7. *Let (A, M) be a minimal pair with dimension ϵ and $B \subseteq M$. Then $\dim B/(A \cap B) \geq -\epsilon$. Moreover if $B \cup A \neq M$ then $\dim B/(A \cap B) \geq 0$.*

Proof. We have $\dim(B \cup A/A) \leq \dim(B/A \cap B)$ by Lemma 4.1. Note that $A \subseteq B \cup A \subseteq M$. If $B \cup A \neq M$ then we have $\dim(B \cup A/A) \geq 0$ by minimality. If $B \cup A = M$ then we have $\dim(B \cup A/A) = -\epsilon$. $\square$

Lemma 4.8. *Let $\langle M_i \rangle_{0 \leq i \leq n}$ be a minimal chain with dimensions*

$$\dim(M_i/M_{i-1}) = -\epsilon_i.$$

Let

$$\epsilon = \sum_{i=1}^{n} \epsilon_i,$$

and let $B \subseteq M_n$. *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 the previous lemma. Thus

$$\dim(B/M_0 \cap B) = \dim(B_n/B_0) = \sum_{i=1}^{n} \dim(B_{i+1}/B_i) \geq -\epsilon.$$

$\square$

5. Lower bound

Definition 5.1. Suppose $\phi_{A,B}(x, y)$ is a basic formula. Define $\mathcal{X}$ to be the graph on vertices x with edges defined by ϕ (equivalently it is a subgraph of A induced by vertices of x). Similarly define $\mathcal{Y}$. Note that $\mathcal{X}$, $\mathcal{Y}$ are positive as A is positive. Additionally, let $\mathcal{Y}'$ be a subgraph of $\mathcal{Y}$ induced by vertices of $\mathcal{Y}$ that are connected to $W - (X \cup Y)$, where W is a witness of ϕ.

In this section we restrict our attention to the following family of basic formulas $\phi(x, y)$:

- All formulas have $\mathcal{Y}' = \mathcal{Y}$.
- All formulas define no edges between X and Y.
- The minimal chain of $\phi(x, y)$ consists of one step, that is we only have one minimal extension as opposed to a chain of minimal extensions.
- The dimension of that minimal extension is smaller than α.

We obtain a lower bound for the formulas that are boolean combinations of basic formulas of this type written in disjunctive-normal form. First, define $\epsilon_L(\phi)$.

Definition 5.2. For a basic formula $\phi = \phi_{\langle M_i \rangle_{0 \leq i \leq n}}(x, y)$ let

- $\epsilon_i(\phi) = -\dim(M_i/M_{i-1})$.
- $\epsilon_L(\phi) = \sum_1^n \epsilon_i(\phi)$.

Definition 5.3 (Negation)**.** If ϕ is a basic formula, then define

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

Definition 5.4 (Conjunction)**.** Take a finte collection of formulas $\phi_i(x, y)$ where each ϕ_i is a positive or a negative basic formula and $\phi = \bigwedge_i \phi_i$. 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 let $\epsilon_L(\phi) = \infty$. In the case of conflicting definitions the formula would have no realizations. Otherwise let

$$\epsilon_L\left(\bigwedge_i \phi_i\right) = \sum_i \epsilon_L(\phi_i).$$

Definition 5.5 (Disjunction)**.** Take a collection of formulas ψ_i where each instance is a conjunction as above all agreeing on $\mathcal{X}$ and $\mathcal{Y}$. Then

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

Theorem 5.6. *For a formula ψ as above we have*

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

where $Y(\psi)$ is $\dim(Y)$ (as all basic components agree on $\mathcal{Y}$).

Proof. First, work with a formula that is a conjunction of positive basic formulas $\psi = \bigwedge_{i\in I} \phi_i$. Then as we have defined above

$$\epsilon_L(\psi) = \sum_{i\in I} \epsilon_L(\phi_i).$$

If W_i is a witness of ϕ_i, let $S_i = |W_i|$. Let n_1 be the largest natural number such that

$$n_1\epsilon_L(\psi) < Y(\psi).$$

Let ϵ' be the smallest value among $\epsilon_L(\phi_i)$ corresponding to the formula ϕ'. Let n_2 be the largest natural number such that

$$n_1\epsilon_L(\psi) + n_2\epsilon' < Y(\psi).$$

Fix some $N > n_1 + n_2$. Let

$$J = \{0 \leq j < N\} \subseteq \mathbb{N}.$$

Let a_j be a graph isomorphic to $\mathcal{X}$ for each $j \in J$, pairwise disjoint. Let $A = \bigcup_{1\leq j\leq N} a_j$. Let

$$S = |Y| + (n_1 + n_2 + 1)\sum_{i\in I} S_i.$$

By Lemma 4.3 the graph A can be embedded into $\mathbb{G}$ as an S-strong graph. Abusing notation, we identify A with this embedding. Thus we have $A \subseteq \mathbb{G}$, S-strong.

Let J_1, J_2 be disjoint subsets of J, of sizes n_1, n_2 respectively. Let b be a graph isomorphic to $\mathcal{Y}$. For each $i \in I, j \in J_1$ let W_{ij} be a witness of $\phi_i(a_j, b)$. (Note that then $(a_j \cup b, W_{ij})$ is a minimal pair.) For each $j \in J_1$ let W_j be a union of $\{W_{ij}\}_{i\in I}$ disjoint over $a_j \cup b$. For each $j \in J_2$ let W_j be a witness of $\phi'(a_j, b)$. Let W' be a union of $\{W_j\}_{j\in J_1\cup J_2}$ disjoint over b. Let W be a union of W' and A disjoint over $\{a_j\}_{j\in J_1\cup J_2}$.

Claim 5.7. *We have $A \leq W$.*

Proof. Consider some $A \subsetneq B \subseteq W$. We need to show $\dim(B/A) > 0$. Let $\overline{A} = A\cup b$. We have

$$\dim(B/A) = \dim(B/B \cap \overline{A}) + \dim(B \cap \overline{A}/A).$$

Let $B_{ij} = B \cap W_{ij}$. Let $B_j = B \cap W_j$. To unify indices, relabel all the graphs above as $\{B_k\}_{k \in K}$ for some index set K. By the construction of W we have

$$\dim(B/B \cap \overline{A}) = \sum_{k \in K} \dim(B_k/B_k \cap \overline{A}).$$

Fix k. We have $B_k \subseteq W_k$, where W_k is a minimal extension of $M_0^k = a \cup b$ for some $a \in A$. Let ϵ_k be the dimension of this minimal extension. We have $\dim(B_k/B_k \cap \overline{A}) = \dim(B_k/a \cup (B \cap b))$.

Case 1: $B \cap b = b$. Then $M_0^k \subseteq B_k \subseteq W_k$ and

$$\dim(B_k/a \cup (B \cap b)) = \dim(B_k/M_0^k).$$

By minimality of (M_0^k, B_k) we have $\dim(B_k/M_0^k) \geq -\epsilon_k$. Thus

$$\dim(B/B \cap \overline{A}) \geq -\sum_{k \in K} \epsilon_k = -\left(n_1 \epsilon_L(\psi) + n_2 \epsilon'\right).$$

In addition

$$\dim(B \cap \overline{A}/A) = \dim(b) = Y(\psi).$$

Combining the two, we get

$$\dim(B/A) \geq Y(\psi) - \left(n_1 \epsilon_L(\psi) + n_2 \epsilon'\right),$$

which is positive by the construction of n_1, n_2 as needed.

Case 2: $B \cap b \subsetneq b$.

Claim 5.8. *We have* $\dim(B_k/B_k \cap \overline{A}) > 0$.

Proof. Recall that $\dim(B_k/B_k \cap \overline{A}) = \dim(B_k/a \cup (B \cap b))$. First, suppose that $B_k \cup M_0^k \neq W_k$. Then by Lemma 4.7 we get the required inequality. Thus we may assume that $B_k \cup M_0^k = W_k$. By Lemma 4.1 we have

$$\dim(B_k \cup M_0^k/M_0^k) = \dim(B_k/B_k \cap M_0^k) - \alpha E,$$

where E is the number of edges connecting the vertices of $B_k - M_0^k = B_k \cup M_0^k - M_0^k$ to the vertices of $M_0^k - B_k = M_0^k - B_k \cap M_0^k$. Noting that $B_k \cup M_0^k = W_k$, $\dim W_k/M_0^k = -\epsilon_k$, and $B_k \cap M_0^k = a \cup (B \cap b)$ we may rewrite the equality above as

$$\dim(B_k/a \cup (B \cap b)) = \alpha E - \epsilon,$$

and E is the number of edges connecting the vertices of $W_k - M_0^k$ to the vertices of $M_0^k - a \cup (B \cap b)$. As $\mathcal{Y} = \mathcal{Y}'$ and $B \cap b \subsetneq b$ we must have $E \geq 1$. But then as $\alpha > \epsilon$ we have $\dim(B_k/a \cup (B \cap b)) > 0$ as needed. □

Now, recall that

$$\dim(B/A) = \dim(B \cap \overline{A}/A) + \sum_{k \in K} \dim(B_k/B_k \cap \overline{A}).$$

By the claim above each of $\dim(B_k/B_k \cap \overline{A}) > 0$, thus

$$\dim(B/A) > \dim(B \cap \overline{A}/A).$$

In addition

$$\dim(B \cap \overline{A}/A) = \dim(B \cap b) \geq 0,$$

as b is postive. Thus $\dim(B/A) > 0$ as needed. □

As $A \leq W$ and $A \subseteq \mathbb{G}$, we can embed W into $\mathbb{G}$ over A. Abusing notation again, we identify W with its embedding $A \leq W \subseteq \mathbb{G}$. In particular, now we have $b \in \mathbb{G}$. Also note that

$$\dim(W/A) = Y(\psi) - (n_1 \epsilon_L(\psi) + n_2 \epsilon')\,,$$

$$|W| - |A| \leq |b| + (n_1 + n_2) \sum_{i \in I} S_i.$$

Lemma 5.9. *We have*

$$\{a_j\}_{j\in J_1} \subseteq \psi(A,b) \subseteq \{a_j\}_{j\in J_1\cup J_2}.$$

Proof. First inclusion $\{a_j\}_{j\in J_1} \subseteq \psi(A,b)$ is immediate from the construction of W, as W_{ij} witnesses that $\phi_i(a_j,b)$ holds. For the second inclusion, suppose that there is $a \in A - \{a_j\}_{j\in J_1\cup J_2}$ such that $\psi(a,b)$ holds. Let $W' \subseteq \mathbb{G}$ be a witness of $\phi_1(a,b)$. First, note that the case $W' \subseteq W$ is impossible as there are no edges between a and $W-a$, but there are edges between a and $W'-a$. Thus assume $W' \not\subseteq W$. As $(a\cup b, W')$ is minimal, by Lemma 4.5 we have $\dim(W'\cup W/W) < -\epsilon_1$. Therefore

$$\dim(W'\cup W/A) = \dim(W'\cup W/W) + \dim(W/A) < Y(\psi) - (n_1\epsilon_L(\psi) + n_2\epsilon') - \epsilon_1,$$

which is negative by the construction of n_1, n_2. Thus $A \not\leq W\cup W'$, as then it would have a positive dimension. Additionally,

$$|W'\cup W| - |A| \leq |W'-W| + |W| - |A| \leq S_1 + |b| + (n_1+n_2)\sum_{i\in I} S_i \leq S,$$

but then this contradicts that A is S-strong, as then we would have $A \leq W\cup W'$. $\square$

In the construction of W we have chosen indices J_1, J_2 arbitrarily. In particular, suppose we let J_2 to be the last n_2 indices of J and J_1 an arbitrary n_1-element subset of the first $N-n_2$ elements of J. Each of those choices would then yield a different trace $\psi(A,b)$ by the lemma above. Thus $\psi(A, M^{|y|}) \geq \binom{N-n_2}{n_1}$ and therefore $\mathrm{vc}(\psi) \geq n_1$. By the definition of n_1 we have $n_1 = \left\lfloor \frac{Y(\psi)}{\epsilon_L(\psi)} \right\rfloor$, so this proves the theorem for ψ.

Now consider a formula which is a conjunction consisting of negative basic formulas $\psi = \bigwedge_{i\in I} \neg\phi_i$. Let $\bar{\psi} = \bigwedge_{i\in I} \phi_i$. Do the construction above for $\bar{\psi}$ and suppose its trace is $X \subseteq A$ for some b. Then over b the same construction gives trace $(A-X)$ for ψ. Thus we get as many traces as above, and the same bound.

Finally consider a formula which is a disjunction of formulas considered above $\theta = \bigvee_{k \in K} \psi_k$. Choose the one with the smallest ϵ_L, say ψ_k, and repeat the construction above for ψ_k. Any trace we obtain is automatically a trace for θ, and thus we get as many traces as above, and the same bound. □

Corollary 5.10. *The vc-function is infinite in the theory of Shelah-Spencer random graphs:*

$$\mathrm{vc}^{S_\alpha}(n) = \infty.$$

Proof. Let A be a graph consisting of $\lceil 1/\alpha \rceil + 2 + n$ disconnected vertices. Fix $\epsilon > 0$. By Lemma 4.2, there exists B such that (A, B) is minimal with dimension $\leq \epsilon$. Consider a basic formula $\psi_{A,B}(x, y)$ where $|x| = \lceil 1/\alpha \rceil + 2$ and $|y| = n$. Then by the theorem above $\mathrm{vc}^{S_\alpha}(n) \geq \mathrm{vc}(\psi_{A,B}) \geq \left\lfloor \frac{n}{\epsilon} \right\rfloor$. As ϵ was arbitrary, this number can be made arbitrarily large, giving $\mathrm{vc}^{S_\alpha}(n) = \infty$ as needed. □

Corollary 5.11. *The theory of Shelah-Spencer random graphs doesn't have finite dp-rank. In particular it is not dp-minimal.*

Proof. Suppose that the ambient model $\mathbb{G}$ is $\aleph_1$-saturated. We would like to modify the proof of Theorem 5.6 such that A is indiscernible. Note that as S_α is stable, all indiscernible sequences are totally indiscernible. Note that in the proof we can construct sets $A = \{a_j\}_{j \in J}$ of arbitrary length. Moreover for every finite $J' \subseteq J$, the set $A = \{a_j\}_{j \in J'}$ is still S-strong. Thus we can find an infinite set $A = \{a_j\}_{j \in \mathbb{N}}$ indiscernible and S-strong. Repeating the construction of the corollary above, we can obtain a formula with an arbitrarily large vc-density over the indiscernible sequence A. □

6. Upper bound

Consider a basic formula $\phi(x,y)$ associated to a minimal chain $\langle M_i \rangle_{0\leq i\leq n_\phi}$ with dimensions $\dim(M_{i+1}/M_i) = -\epsilon_i$. Define

$$\epsilon(\phi) = \min\{\epsilon_i\}_{0\leq i\leq n_\phi}$$
$$K(\phi) = |M_{n_\phi}|.$$

Now consider a finite collection of basic formulas

$$\Phi = \Phi(x,y) = \{\phi_i(x,y)\}_{i\in I}\,.$$

Define

$$\epsilon(\Phi) = \min\{\epsilon(\phi_i)\}_{i\in I}\cup\{\alpha\}\,, (\text{ so })\epsilon(\Phi) > 0$$
$$K(\Phi) = \max\{K(\phi_i)\}_{i\in I}\,.$$

Theorem 6.1. *If ϕ is a boolean combination of formulas from Φ, then*

$$\mathrm{vc}(\phi) \leq \left\lfloor |y|\frac{K(\Phi)}{\epsilon(\Phi)}\right\rfloor.$$

We first reduce Theorem 6.1 to a combinatorial statement (Theorem 6.6 below), the proof of which takes up the rest of this section.

Let

$$S = \left\lceil \left(\frac{|y|}{\epsilon(\phi)} + 1\right) K(\phi)\right\rceil.$$

Suppose we have a finite parameter set $A_0 \subseteq \mathbb{G}^{|x|}$ with $|A_0| = N_0$. We would like to bound $\left|\phi(A_0, \mathbb{G}^{|y|})\right|$ in terms of Φ and N_0. Let $A_1 \subseteq \mathbb{G}$ consist of the components of the tuples of A_0 (so $A_1 \subseteq A_0^{|x|}$). Then $|A_1| \leq |x|N_0$. Using Lemma 4.4 let A be a graph $A_0 \subseteq A \subseteq \mathbb{G}$, S-strong in $\mathbb{G}$. Let $N = |A|$. We have $N \leq |x|N_0M$ (where $M = M(S,\alpha)$ is the constant from Lemma 4.4). As $A_0 \subseteq A^{|x|}$ we have

$$\left|\phi(A_0, \mathbb{G}^{|y|})\right| \leq \left|\phi(A^{|x|}, \mathbb{G}^{|y|})\right|.$$

Therefore it suffices to bound $\left|\phi(A^{|x|}, \mathbb{G}^{|y|})\right|$ uniformly in $\Phi, |A|$.

Definition 6.2. For $A \subseteq \mathbb{G}^{|x|}, B \subseteq \mathbb{G}^{|y|}, b \in \mathbb{G}^{|y|}$ define

$$\Phi(A, b) = \{(a, i) \in A \times I \mid \mathbb{G} \models \phi_i(a, b)\} \subseteq A \times I,$$

$$\Phi(A, B) = \{\Phi(A, b) \mid b \in B\} \subseteq \mathcal{P}(A \times I).$$

Lemma 6.3. *For $A \subseteq \mathbb{G}^{|x|}, B \subseteq \mathbb{G}^{|y|}$ if ϕ is a boolean combination of formulas from Φ then $|\phi(A, B)| \leq |\Phi(A, B)|$.*

Proof. Clear, as for all $a \in A, b \in B$ the set

$$\Phi(a, b) = \{i \in I \mid \mathbb{G} \models \phi_i(a, b)\}$$

determines the truth value of $\phi(a, b)$. □

Thus it suffices to bound $\left|\Phi(A^{|x|}, \mathbb{G}^{|y|})\right|$ in terms of $\Phi, |A|$. Below we fix $A \subseteq \mathbb{G}$.

Definition 6.4.

- For all $i \in I, a \in A^{|x|}, b \in \mathbb{G}^{|y|}$ if $\phi_i(a, b)$ holds fix $W^i_{a,b} \subseteq \mathbb{G}$, a witness of $\phi_i(a, b)$.
- For $b \in \mathbb{G}^{|y|}$ let
$$W_b = \bigcup \left\{ W^i_{a,b} \mid a \in A^{|x|}, i \in I, \mathbb{G} \models \phi_i(a, b) \right\}.$$
- Suppose A, B are subgraphs of $\mathbb{D}$ such that $v(A), v(B)$ are disjoint. Then define $\mathscr{E}(A, B)$ to be the number of edges between the vertices in $v(A)$ and the vertices in $v(B)$.
- For $C, B \subseteq \mathbb{G}$ define the <u>boundary</u> of C over B
$$\partial(C, B) = \{b \in B \mid \mathscr{E}(b, C - B) \neq 0\} \subseteq B.$$
- For $b \in \mathbb{G}^{|y|}$ let $\partial_b = \partial(W_b, A) \subseteq A$.
- For $b \in \mathbb{G}^{|y|}$ let $\overline{W}_b = (W_b - A) \cup \partial_b$.

- For $b_1, b_2 \in \mathbb{G}^{|y|}$ we say that $b_1 \sim b_2$ if $\partial_{b_1} = \partial_{b_2}$, $b_1 \cap A = b_2 \cap A$, and there exists a graph isomorphism from $\overline{W}_{b_1} \cup b_1$ to $\overline{W}_{b_2} \cup b_2$ (here, and from now on, we treat tuples as subgraphs) that fixes ∂_{b_1} and maps b_1 to b_2. One easily checks that this defines an equivalence relation.
- For $b \in \mathbb{G}^{|y|}$ define $\mathscr{I}_b$ to be the $\sim$-equivalence class of b.

Lemma 6.5. *For $b_1, b_2 \in \mathbb{G}^{|y|}$ if $b_1 \sim b_2$ then $\Phi(A^{|x|}, b_1) = \Phi(A^{|x|}, b_2)$.*

Proof. Fix a graph isomorphism $\bar{f} \colon \overline{W}_{b_1} \cup b_1 \longrightarrow \overline{W}_{b_2} \cup b_2$. Extend it to an isomorphism $f \colon W_{b_1} \cup A \longrightarrow W_{b_2} \cup A$ by being an identity map on the new vertices. Suppose $\mathbb{G} \models \phi_i(a, b_1)$ for some $a \in A^{|x|}$. Then $f(W^i_{a,b_1})$ is a witness of $\phi_i(a, b_2)$ (though not necessarily equal to W^i_{a,b_2}) and thus $\mathbb{G} \models \phi_i(a, b_2)$. As a was arbitrary, this proves the equality of the traces. $\square$

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

Theorem 6.6. *Suppose we have $b \in \mathbb{G}^{|y|}$. Let $Y = |b - A|$. Then*

$$|\partial_b| \leq \left\lfloor Y \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor$$
$$|\overline{W}_b| \leq \left\lfloor 3Y \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor$$

From this theorem we get the desired result:

Corollary 6.7. *(Theorem 6.1) If ϕ is a boolean combination of formulas from Φ, then* $\mathrm{vc}(\phi) \leq \left\lfloor |y| \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor$.

Proof of Theorem 6.1 is based on Theorem 6.6. We count possible equivalence classes of $\sim$. This amounts to bounding the possibilities for ∂_b, $b \cap A$, and the number of isomorphism classes of W_b. Fix $b \in \mathbb{G}^{|y|}$ and let $Y = |b - A|$. Let $D = \left\lfloor Y \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor$, and $D' = \left\lfloor 3Y \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor$. By the first part of Theorem 6.6 there are $\binom{N}{D}$ choices for boundary ∂_b. By the second part of Theorem 6.6 there are at most D' vertices in $\overline{W}_b$. Thus to determine the graph $\overline{W}_b$ we need to choose how many vertices it has and then decide where edges go. This amounts to at most $D'2^{(D')^2}$ choices. Additionally there are $\binom{N}{|y|-Y}$ choices for $b \cap A$. Let $D_1(\Phi) = \frac{K(\Phi)}{\epsilon(\Phi)}, D(\Phi) = \left\lfloor |y| \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor$. In total this gives us at most

$$\binom{N}{D} \cdot \binom{N}{|y|-Y} \cdot D'2^{(D')^2} = O(N^{D+|y|-Y}) =$$
$$= O(N^{\left\lfloor Y \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor + |y| - Y}) = O(N^{\lfloor |y| D_1(\Phi) \rfloor}) = O(N^{D(\Phi)})$$

choices (second to last inequality uses $\frac{K(\Phi)}{\epsilon(\Phi)} \geq 1$). By Lemma 6.5 we have $\left|\Phi(A^{|x|}, \mathbb{G}^{|y|})\right| = O(N^{D(\Phi)})$. Recall that

$$\left|\phi(A_0, \mathbb{G}^{|y|})\right| \leq \left|\Phi(A^{|x|}, \mathbb{G}^{|y|})\right|.$$

Therefore we have

$$\left|\phi(A_0, \mathbb{G}^{|y|})\right| = O(N^{D(\Phi)}) = O((|x| N_0 M)^{D(\Phi)}) =$$
$$= O((|x|M)^{D(\Phi)} N_0^{D(\Phi)}) = O(N_0^{D(\Phi)}).$$

As A_0 was arbitrary, this shows that $\mathrm{vc}(\phi) \leq \left\lfloor |y| \frac{K(\Phi)}{\epsilon(\Phi)} \right\rfloor$ as needed. (Note that throughout this proof we are using the fact that quantities $\boldsymbol{K}_\alpha(\Phi), \epsilon(\Phi), M$ are completely determined by Φ, thus independent from A_0.) □

Proof. (of Theorem 6.6) The graph W_b is a union of witnesses of the form $W_{a,b}$ for some $a \in A^{|x|}, b \in \mathbb{G}^{|y|}$. Enumerate all of them as $\{W_j\}_{1 \leq j \leq J}$. Define $M_j = \bigcup_{k=1}^{j} W_k$ for $1 \leq j \leq J$ and let $M_0 = b, M_{-1} = \emptyset$. Let $\overline{A} = A \cup b$.

Definition 6.8. For $0 \leq j \leq J$ define:

- Let $v_j = 1$ if new vertices are added to M_j outside of $\overline{A}$, that is if $M_j - \overline{A} \neq M_{j-1} - \mathcal{B}$, and let it be 0 otherwise.
- Let $E_j = \partial(A - W_j, M_j - A)$.
- Let

$$m_j = \sum_{k=0}^{j} (v_j + |E_j|).$$

Lemma 6.9. *For $0 \leq j \leq J$ we have*

$$|\partial(M_j, A)| \leq |E_0| + m_j K(\Phi)$$

Proof. Proceed by induction on $j = 0, \ldots, J$. The base case $j = 0$ is clear. For the inductive step suppose that

$$|\partial(M_{j-1}, A)| \leq m_{j-1} K(\Phi)$$

holds. Let

$$\begin{aligned} \delta_1 &= \partial(M_j, A) - \partial(M_{j-1}, A) = \\ &= \{a \in A \mid \mathscr{E}(a, M_j - A) \neq 0 \text{ and } \mathscr{E}(a, M_{j-1} - A) = 0\}. \end{aligned}$$

If $M_j - A = M_{j-1} - A$ then $\delta_1 = \emptyset$ and we are done as m_j is increasing. Suppose not. We have $|\delta_1| = |\delta_1 \cap W_j| + |\delta_1 - W_j|$, and

$$\delta_1 - W_j = \{a \in A - W_j \mid \mathscr{E}(a, M_j - A) \neq 0 \text{ and } \mathscr{E}(a, M_{j-1} - A) = 0\}.$$

But then it's clear that $\delta_1 - W_j \subseteq E_j$ as

$$\begin{aligned} &W_j - M_{j-1} - A \subseteq M_j - A, \\ &(W_j - M_{j-1} - A) \cap (M_{j-1} - A) = \emptyset. \end{aligned}$$

As $b \in M_{j-1}$ and $M_j - A \neq M_{j-1} - A$, then $M_j - \overline{A} \neq M_{j-1} - \overline{A}$, and thus $v_j = 1$. Therefore we have

$$\begin{aligned} |\delta_1| &= |\delta_1 \cap W_j| + |\delta_1 - W_j| \leq |W_j| + |E_j| \leq \\ &\leq K(\Phi) + |E_j| \leq (v_j + |E_j|)K(\Phi) \leq (m_j - m_{j-1})K(\Phi), \end{aligned}$$

as needed. $\square$

Lemma 6.10. *For* $0 \le j \le J$ *we have*

$$|M_j - \overline{A}| \le \sum_{k=0}^{j} v_k K(\Phi)$$

Proof. Proceed by induction on $j = 0, \dots, J$. The base case $j = 0$ is clear. For the inductive step suppose that

$$|M_{j-1} - \overline{A}| \le \sum_{k=0}^{j-1} v_k K(\Phi)$$

holds. If $M_j - \overline{A} = M_{j-1} - \overline{A}$ then the inequality is immediate as $v_j \ge 0$. Therefore assume this is not the case, so $v_j = 1$ and $|M_j - A| - |M_{j-1} - A| \le |W_j| \le v_j K(\Phi)$, and so we get the required inequality.

□

Lemma 6.11. *For* $0 \le j \le J$ *we have*

$$\dim(M_j \cup \overline{A}/\overline{A}) \le -m_j \epsilon(\Phi),$$

Proof. Proceed by induction on $j = 0, \dots, J$. The base case $j = 0$ is clear. For the inductive step suppose that

$$\dim(M_{j-1} \cup \overline{A}/\overline{A}) \le -m_{j-1} \epsilon(\Phi)$$

holds. We have

$$\dim(M_j \cup \overline{A}/\overline{A}) = \dim(M_j \cup \overline{A}/M_{j-1} \cup \overline{A}) + \dim(M_{j-1} \cup \overline{A}/\overline{A}) \le$$

$$\le \dim(M_j \cup \overline{A}/M_{j-1} \cup \overline{A}) - m_{j-1} \epsilon(\Phi).$$

Let $\overline{M}_{j-1} = M_{j-1} \cup \overline{A}$. By Lemma 4.1

$$\dim(M_j \cup \overline{A}/M_{j-1} \cup \overline{A}) = \dim(W_j \cup \overline{M}_{j-1}/\overline{M}_{j-1}) = \dim(W_j/W_j \cap \overline{M}_{j-1}) - e\alpha$$

where e is the number of edges connecting the vertices of $\overline{M}_{j-1} - W_j$ to the vertices of $W_j - \overline{M}_{j-1}$. Recall that $E_j = \partial(A - W_j, M_j - A)$. We have $A - W_j \subseteq \overline{M}_{j-1} - W_j$ (as $A \subseteq \overline{M}_{j-1}$) and $W_j - M_{j-1} - A = W_j - \overline{M}_{j-1}$ (as for $j > 1$, we have $b \subseteq M_{j-1}$). Thus $|E_j| \leq e$, and we get

$$\dim(M_j \cup \overline{A}/M_{j-1} \cup \overline{A}) \leq \dim(W_j/W_j \cap \overline{M}_{j-1}) - |E_j|\alpha.$$

If $W_j \subseteq \overline{M}_{j-1}$ then $\dim(W_j/W_j \cap \overline{M}_{j-1}) = 0$. If not, then by Lemma 4.6 we have $\dim(W_j/W_j \cap \overline{M}_{j-1}) \leq -\epsilon(\Phi)$. Either way, we have $\dim(W_j/W_j \cap \overline{M}_{j-1}) \leq -v_j\epsilon(\Phi)$. Using this and the fact that $\epsilon(\Phi) \leq \alpha$, we obtain

$$\dim(M_j \cup \overline{A}/M_{j-1} \cup \overline{A}) \leq -v_j\epsilon(\Phi) - |E_j|\epsilon(\Phi) = -(m_j - m_{j-1})\epsilon(\Phi).$$

Finally,

$$\begin{aligned}\dim(M_j \cup \overline{A}/\overline{A}) &\leq \dim(M_j \cup \overline{A}/M_{j-1} \cup \overline{A}) - m_{j-1}\epsilon(\Phi) \leq \\ &\leq -(m_j - m_{j-1})\epsilon(\Phi) - m_{j-1}\epsilon(\Phi) = -m_j\epsilon(\Phi),\end{aligned}$$

as needed. □

(Proof of Theorem 6.6 continued) For any $0 \leq j \leq J$ we have

$$\begin{aligned}\dim(M_j \cup A/A) &= \dim(\overline{A}/A) + \dim(M_j \cup \overline{A}/\overline{A}) \\ &\leq Y - |E_0|\alpha + \dim(M_j \cup \overline{A}/\overline{A}).\end{aligned}$$

Lemma 6.11 gives us

$$\dim(M_j \cup \overline{A}/\overline{A}) \leq -m_j\epsilon(\Phi).$$

Thus

$$\dim(M_j \cup A/A) \leq Y - |E_0|\alpha - m_j\epsilon(\Phi).$$

Suppose j is an index such that

$$Y - |E_0|\alpha - m_j\epsilon(\Phi) \geq 0,$$

$$Y - |E_0|\alpha - m_{j+1}\epsilon(\Phi) < 0$$

if one exists. Then

$$m_j \leq \frac{Y - |E_0|\alpha}{\epsilon(\Phi)}.$$

By Lemma 6.10 we have

$$\begin{aligned}|M_{j+1} - A| &\leq \left(\sum_{k=1}^{j+1} v_k\right) K(\Phi) \leq (m_j + 1)K(\Phi)\\ &\leq \left(\frac{Y - |E_0|\alpha}{\epsilon(\Phi)} + 1\right) K(\Phi) \leq S.\end{aligned}$$

This is a contradiction, as A is S-strong and $\dim(M_{j+1} \cup A/A)$ is negative. Thus $Y - |E_0|\alpha - m_j\epsilon(\Phi) \geq 0$ for all $j \leq J$. In particular $Y - |E_0|\alpha - m_J\epsilon(\Phi) \geq 0$, so $m_J \leq \frac{Y-|E_0|\alpha}{\epsilon(\Phi)}$. Noting that $M_J = W_b$, Lemma 6.9 gives us

$$|\partial_b| = |\partial(W_b, A)| \leq |E_0| + m_J K(\Phi) \leq |E_0| + K(\Phi)\frac{Y - |E_0|\alpha}{\epsilon(\Phi)}.$$

As $K(\Phi) \geq 1$ and $\epsilon(\Phi) \geq \alpha$, we get

$$|\partial_b| \leq K(\Phi)\frac{Y}{\epsilon(\Phi)} \leq Y\frac{K(\Phi)}{\epsilon(\Phi)}.$$

As $|\partial_b|$ is an integer we have $|\partial_b| \leq \left\lfloor Y\frac{K(\Phi)}{\epsilon(\Phi)}\right\rfloor$. But this is precisely the first inequality we need to prove. For the second inequality, Lemma 6.10 gives us

$$\begin{aligned}\left|W_b - \overline{A}\right| &\leq Y + \left(\sum_{k=0}^{J} v_k\right) K(\Phi) \leq Y + m_J K(\Phi) \leq\\ &\leq Y + K(\Phi)\frac{Y}{\epsilon(\Phi)} \leq 2Y\frac{K(\Phi)}{\epsilon(\Phi)} \leq 2Y\frac{K(\Phi)}{\epsilon(\Phi)}.\end{aligned}$$

As $\left|W_b - \overline{A}\right|$ is an integer we have $\left|W_b - \overline{A}\right| \leq \left\lfloor 2Y\frac{K(\Phi)}{\epsilon(\Phi)}\right\rfloor$. Thus we have

$$|\overline{W}_b| \leq |W_b - A| + |\partial_b| \leq \left\lfloor 3Y\frac{K(\Phi)}{\epsilon(\Phi)}\right\rfloor,$$

as needed. This ends the proof of Theorem 6.6. □

7. Conclusion

This paper computes upper and lower bounds for certain types of formulas in Shelah-Spencer graphs. The bounds are not tight: in the best case scenario for a basic formula $\phi(x, y)$ defining a minimal extension of dimension ϵ we have

$$\frac{|y|}{\epsilon} \leq \mathrm{vc}(\phi) \leq K\frac{|y|}{\epsilon},$$

where K is the number of vertices in the minimal extension. Thus there is a multiple of K gap between lower and upper bounds. It is this author's hope that a refinement of presented techniques can yield better estimates of the vc-density. One potential direction towards this goal is to have a closer study on how multiple minimal extensions can intersect without increasing overall dimension.

One direction for future work is to ask what these bounds on vc-density can tell about structure in large finite random graphs, along the lines of results in [**?**].

Note that this paper doesn't answer the question whether there can be exotic values for vc-density of individual formulas, such as non-integer or irrational values. A better bound can help address this.

Open Question 7.1. *In Shelah-Spencer graphs can a formula have non-integer or irrational vc-density?*

Another observation is that while $\mathrm{vc}(n) = \infty$ there seems to be a good structural behavior of the vc-density for individual formulas. This suggests that perhaps the vc-function is not the best tool to describe behaviour of the definable sets in Shelah-Spencer graphs, and some more refined measure might be required. One potential way to do this is to separate the formulas based on values of $K(\phi), \epsilon(\phi)$. Once those are bounded, vc-density seems to be well-behaved. The author hopes to explore this further in his future work.

References

[1] Matthias Aschenbrenner, Alf Dolich, Deirdre Haskell, Dugald Macpherson, and Sergei Starchenko. Vapnik-chervonenkis density in some theories without the independence property, i. *Transactions of the American Mathematical Society*, 368(8):5889–5949, 2016.

[2] Patrick Assouad. Densité et dimension. *Annales de l'institut Fourier*, 33(3):233–282, 1983.

[3] Patrick Assouad. Observations sur les classes de vapnik-cervonenkis et la dimension combinatoire de blei. In *Seminaire d'Analyse Harmonique, 1983–1984*, volume 85-2 of *Publications Mathématiques d'Orsay*, pages 92–112. Université de Paris-Sud, Département de Mathématiques, Orsay, 1985.

[4] Justin Brody and Michael C Laskowski. On rational limits of shelah–spencer graphs. *The Journal of Symbolic Logic*, 77(02):580–592, 2012.

[5] Vincent Guingona and Cameron Donnay Hill. Local dp-rank and vc-density over indiscernible sequences. *MLQ Math. Log. Q.*, 60(1-2):59–65, 2014.

[6] Itay Kaplan, Alf Onshuus, and Alexander Usvyatsov. Additivity of the dp-rank. *Trans. Amer. Math. Soc.*, 365(11):5783–5804, 2013.

[7] Michael C. Laskowski. A simpler axiomatization of the shelah-spencer almost sure theories. *Israel Journal of Mathematics*, 161(1):157–186, 2007.

[8] N Sauer. On the density of families of sets. *Journal of Combinatorial Theory, Series A*, 13(1):145–147, 1972.

[9] Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41(1):247–261, 1972.

[10] Saharon Shelah. Strongly dependent theories. *Israel J. Math.*, 204(1):1–83, 2014.

[11] Saharon Shelah and Joel Spencer. Zero-one laws for sparse random graphs. *Journal of the American Mathematical Society*, 1(1):97–115, 1988.

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