

# VC-DENSITY FOR TREES

ANTON BOBKOV

ABSTRACT. We show that for the theory of infinite trees we have  $vc(n) = n$  for all  $n$ .

VC density was introduced in [1] 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

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

Where  $vc(n)$  measures complexity of definable sets  $n$ -dimensional space. Simplest possible behavior is  $vc(n) = n$  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  $vc(n) = n$ .

In this paper we work with trees viewed as posets. Parigot in [3] showed that such theory is NIP. This result was strengthened by Simon in [2] showing that trees are dp-minimal. [1] has the following problem

**Problem 0.1.** ([1] p. 47) Determine the VC density function of each (infinite) tree.

Here we settle this question by showing that theory of trees has  $vc(n) = n$ .

## 1. PRELIMINARIES

We use notation  $a \in T^n$  for tuples of size  $n$ . For variable  $x$  or tuple  $a$  we denote their arity by  $|x|$  and  $|a|$  respectively.

We work with finite relational languages. Given a formula we can define its complexity  $n$  as the depth of quantifiers used to build up the formula. More precisely

**Definition 1.1.** Define complexity of a formula by induction:

$$\begin{aligned} \text{Complexity}(\text{q.f. formula}) &= 0 \\ \text{Complexity}(\exists x \phi(x)) &= \text{Complexity}(\phi(x)) + 1 \\ \text{Complexity}(\phi \cap \psi) &= \max(\text{Complexity}(\phi), \text{Complexity}(\psi)) \\ \text{Complexity}(\neg \phi) &= \text{Complexity}(\phi) \end{aligned}$$

A simple inductive argument verifies that there are (up to equivalence) only finitely many formulas with fixed complexity and the number of free variables. We will use the following notation for types and Stone Space:

**Definition 1.2.** Let  $n, m$  be naturals,  $\mathbf{B}$  a structure,  $A$  a parameter set and  $a, b$  tuples in  $\mathbf{B}$ .

- $\text{tp}_B^n(a/A)$  will stand for all the  $A$ -formulas of complexity  $\leq n$  that are true of  $a$  in  $B$ . Note that if  $A$  is finite there are finitely many of them (up to equivalence). Conjunction of those formulas would still have complexity  $\leq n$  so we can just associate a single formula describing this type. If  $A =$  we may write  $\text{tp}_B^n(a)$ .  $B$  will be omitted as well if it is clear from context.
- $B \models a \equiv_A^n b$  means that  $a, b$  have the same type of complexity  $n$  over  $A$  in  $B$ , i.e.  $\text{tp}_B^n(a/A) = \text{tp}_B^n(b/A)$
- $S_{B,m}^n(A)$  will stand for all  $m$ -types of complexity  $n$  over  $A$ .

Language for the trees consists of a single binary predicate  $\{\leq\}$ . Theory of trees states that  $\leq$  defines a partial order and for every element  $a$  we have  $\{x \mid x < a\}$  a linear order. For visualization purposes we assume trees grow upwards, with smaller elements on the bottom and larger elements on the top. If  $a \leq b$  we will say that  $a$  is below  $b$  and  $b$  is above  $a$ .

**Definition 1.3.** Work in a tree  $T$ . For  $x \in T$  let  $I(x) = \{t \in T \mid t \leq x\}$  denote all elements below  $x$ . Meet of two tree elements  $a, b$  is the greatest element of  $I(a) \cap I(b)$  (if one exists) and is denoted by  $a \wedge b$ .

Theory of meet trees requires that any two elements in the same connected component have a meet. Colored trees are trees with a finite number of colors added via unary predicates.

For the entire paper when we say tree we mean a colored tree. We allow our trees to be disconnected or finite unless otherwise stated.

For completeness we also present definition of VC function. One should refer to [1] for more details. Suppose we have a collection  $\mathcal{S}$  of subsets of  $X$ . We define a *shatter function*  $\pi_{\mathcal{S}}(n)$

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

Sauer's Lemma asserts that asymptotically  $\pi_{\mathcal{S}}$  is either  $2^n$  or polynomial. In the polynomial case we define VC density of  $\mathcal{S}$  to be power of polynomial that bounds  $\pi_{\mathcal{S}}$ . More formally

$$\text{vc}(\mathcal{S}) = \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{S}}}{\log n}$$

Given a model  $M \models T$  and a formula  $\phi(x, y)$  we define

$$\begin{aligned} \mathcal{S}_{\phi} &= \{\phi(M^{|n|}, b) : b \in M^{|y|}\} \\ \text{vc}(\phi) &= \text{vc}(\mathcal{S}_{\phi}) \end{aligned}$$

One has to check that this definition is independent of realization of  $T$ , see [1], Lemma 3.2. For a theory  $T$  we define the VC function

$$\text{vc}(n) = \sup\{\text{vc}(\phi(x, y)) : |x| = n\}$$

Also, [1], 3.2 tells us that  $\text{vc}(m) \geq m$ .

## 2. PROPER SUBDIVISIONS: DEFINITION AND PROPERTIES

**Definition 2.1.** Let  $\mathbf{A}, \mathbf{B}, \mathbf{T}$  be models in some (possibly different) finite relational languages. If  $\mathbf{A}, \mathbf{B}$  partition  $\mathbf{T}$  (i.e.  $\mathbf{T} = \mathbf{A} \sqcup \mathbf{B}$ ) we say that  $(\mathbf{A}, \mathbf{B})$  is a *subdivision* of  $\mathbf{T}$ .

**Definition 2.2.**  $(\mathbf{A}, \mathbf{B})$  subdivision of  $\mathbf{T}$  is called *n-proper* if for all  $p, q \in \mathbb{N}$ , for all  $a_1, a_2 \in A^p$  and  $b_1, b_2 \in B^q$  we have

$$\begin{aligned} \mathbf{A} &\models a_1 \equiv_n a_2 \\ \mathbf{B} &\models b_1 \equiv_n b_2 \end{aligned}$$

then

$$\mathbf{T} \models a_1 b_1 \equiv_n a_2 b_2$$

**Definition 2.3.**  $(\mathbf{A}, \mathbf{B})$  subdivision of  $\mathbf{T}$  is called *proper* if it is *n-proper* for all  $n \in \mathbb{N}$ .

**Lemma 2.4.** Consider a subdivision  $(\mathbf{A}, \mathbf{B})$  of  $\mathbf{T}$ . If it is 0-proper then it is proper.

*Proof.* Prove the subdivision is *n-proper* for all  $n$  by induction. Case  $n = 0$  is given by the assumption. Suppose  $n = k + 1$  and we have  $\mathbf{T} \models \exists x \phi^k(x, a_1, b_1)$  where  $\phi^k$  is some formula of complexity  $k$ . Let  $a \in \mathbf{T}$  witness the existential claim i.e.  $\mathbf{T} \models \phi^k(a, a_1, b_1)$ .  $a \in A$  or  $a \in B$ . Without loss of generality assume  $a \in A$ . Let  $\mathbf{p} = \text{tp}_{\mathbf{A}}^k(a, a_1)$ . Then we have

$$\mathbf{A} \models \exists x \text{tp}_{\mathbf{A}}^k(x, a_1) = \mathbf{p}$$

Formula  $\text{tp}_{\mathbf{A}}^k(x, a_1) = \mathbf{p}$  is of complexity  $k$  so  $\exists x \text{tp}_{\mathbf{A}}^k(x, a_1) = \mathbf{p}$  is of complexity  $k + 1$  by inductive hypothesis we have

$$\mathbf{A} \models \exists x \text{tp}_{\mathbf{A}}^k(x, a_2) = \mathbf{p}$$

Let  $a'$  witness this existential claim so that

$$\begin{aligned} \text{tp}_{\mathbf{A}}^k(a', a_2) &= \mathbf{p} \\ \text{tp}_{\mathbf{A}}^k(a', a_2) &= \text{tp}_{\mathbf{A}}^k(a, a_1) \\ \mathbf{A} &\models a' a_2 \equiv_k a a_1 \end{aligned}$$

by inductive assumption we have

$$\begin{aligned} \mathbf{T} &\models a a_1 b_1 \equiv_k a' a_2 b_2 \\ \mathbf{T} &\models \phi^k(a', a_2, b_2) \quad \text{as } \mathbf{T} \models \phi^k(a, a_1, b_1) \\ \mathbf{T} &\models \exists x \phi^k(x, a_2, b_2) \end{aligned}$$

□

We use this lemma for colored trees. Suppose we have  $\mathbf{T}$  to be a model of a colored tree in some language  $\mathcal{L} = \{\leq, \dots\}$  and  $\mathbf{A}, \mathbf{B}$  be in some languages  $\mathcal{L}_A, \mathcal{L}_B$  which will be expands of  $\mathcal{L}$ , with  $\mathbf{A}, \mathbf{B}$  substructures of  $\mathbf{T}$  as reducts to  $\mathcal{L}$ . In this case we'll refer to  $(\mathbf{A}, \mathbf{B})$  as a *proper subdivision* (of  $\mathbf{T}$ ).

**Example 2.5.** Suppose a tree consists of two connected components  $C_1, C_2$ . Then  $(C_1, \leq)$  and  $(C_2, \leq)$  form a proper subdivision.

**Example 2.6.** Work with a (colored) tree  $\mathbf{T}$  in language  $\mathcal{L} = \{\leq, C_1, \dots, C_n\}$  with colors interpreted by sets  $S_1, \dots, S_n$ . Fix  $a \in T$ . Let  $B = \{t \in T \mid a < t\}$ ,  $S = \{t \in T \mid t \leq a\}$ ,  $A = T - B$ . Let  $\mathcal{L}_A = \mathcal{L} \cup \{C\}$  - just  $\mathcal{L}$  expanded by an extra unary predicate. Consider structure  $\mathbf{A}$  in language  $\mathcal{L}_A$  with universe  $A$ , colors  $C_1 \dots C_n$  interpreted appropriately and  $C$  interpreted by  $S$ . Also let  $\mathbf{B}$  be a structure in language  $\mathcal{L}$  and universe  $B$ , considered as substructure of  $\mathbf{T}$ . Then  $(\mathbf{A}, \mathbf{B})$  is a proper subdivision of  $\mathbf{T}$ .

**Definition 2.7.** For  $\phi(x, y)$ ,  $A \subseteq T^{|x|}$  and  $B \subseteq T^{|y|}$

- Let  $\phi(A, b) = \{a \in A \mid \phi(a, b)\} \subseteq A$
- Let  $\phi(A, B) = \{\phi(A, b) \mid b \in B\} \subseteq \mathcal{P}(A)$

$\phi(A, B)$  is a collection of subsets of  $A$  definable by  $\phi$  with parameters from  $B$ . We notice the following bound when  $A, B$  are parts of a proper subdivision.

**Corollary 2.8.** Suppose  $\phi(x, y)$  is a formula of complexity  $n$ . Let  $\mathbf{A}, \mathbf{B}$  be a proper subdivision of  $\mathbf{T}$  and  $b_1, b_2 \in B^{|y|}$ . Then if  $\text{tp}_{\mathbf{B}}^n(b_1) = \text{tp}_{\mathbf{B}}^n(b_2)$  then  $\phi(A^{|x|}, b_1) = \phi(A, b_2)$ . Thus  $|\phi(A^{|x|}, B^{|y|})|$  is bounded by  $|S_{\mathbf{B}}^n(y)|$

*Proof.* Take some  $a \in A^{|x|}$ . We have  $(\mathbf{B}, b_1) \equiv_n (\mathbf{B}, b_2)$  and (trivially)  $(\mathbf{A}, a) \equiv_n (\mathbf{A}, a)$ . Thus by the Lemma 2.4 we have  $(\mathbf{T}, a, b_1) \equiv_n (\mathbf{T}, a, b_2)$  so  $\phi(a, b_1) \iff \phi(a, b_2)$ . Since  $a$  was arbitrary we have  $\phi(A^{|x|}, b_1) = \phi(A^{|x|}, b_2)$ .  $\square$

We note that the number of such types can be bounded uniformly.

**Definition 2.9.** Fix a (finite relational) language  $\mathcal{L}_B$ , and  $n, |y|$ . Let  $N = N(n, |y|, \mathcal{L}_B)$  be smallest number such that for any structure  $\mathbf{B}$  in  $\mathcal{L}_B$  we have  $|S_{\mathbf{B}, |y|}^n| \leq N$ . This number is well-defined as there are a finite number (up to equivalence) of possible formulas of complexity  $\leq n$  with  $|y|$  free variables. Note the following easy inequalities

$$\begin{aligned} n \leq m &\Rightarrow N(n, |y|, \mathcal{L}_B) \leq N(m, |y|, \mathcal{L}_B) \\ |y| \leq |z| &\Rightarrow N(n, |y|, \mathcal{L}_B) \leq N(n, |z|, \mathcal{L}_B) \\ \mathcal{L}_A \subseteq \mathcal{L}_B &\Rightarrow N(n, |y|, \mathcal{L}_A) \leq N(n, |y|, \mathcal{L}_B) \\ N(n, |y|, \mathcal{L}_B) \cdot N(n, |z|, \mathcal{L}_B) &\leq N(n, |y| + |z|, \mathcal{L}_B) \end{aligned}$$

### 3. PROPER SUBDIVISIONS: CONSTRUCTIONS

From now on work in meet trees. First, we describe several constructions of proper subdivisions that are needed for the proof.

**Definition 3.1.** We say that  $E(b, c)$  if  $b$  and  $c$  are connected

$$E(b, c) \Leftrightarrow \exists x (b \geq x) \wedge (c \geq x)$$

Given a tree element  $a$  we can look at all elements above  $a$ , i.e.  $\{x \mid x \geq a\}$ . We can think about it as a closed cone above  $a$ . Connected component of that cone can be thought of an open cones. With that interpretation the notation  $E_a(b, c)$  means that  $b$  and  $c$  are in the same open cone over  $a$ . A more compact way to write it down is through meet notation:

$$E_a(b, c) \Leftrightarrow E(b, c) \text{ and } (b \wedge c) > a$$

Fix a language for colored trees  $\mathcal{L} = \{\leq, C_1, \dots, C_n\} = \{\leq, \vec{C}\}$ . In the following four definitions  $\mathbf{B}$ -structures are going to be in the same language  $\mathcal{L}_B = \mathcal{L} \cup \{U\}$  with  $U$  a unary predicate. It is not always necessary to have this predicate but for the sake of uniformity we keep it.  $\mathbf{A}$ -structures will have different  $\mathcal{L}_A$  languages (those are not as important in later applications). In all subdivisions defined below colors  $\vec{C}$  are interpreted by colors in  $\mathbf{T}$  restricted to the appropriate substructure.

**Definition 3.2.** Fix  $c_1 < c_2$  in  $T$ . Let

$$\begin{aligned} B &= \{b \in T \mid E_{c_1}(c_2, b) \wedge \neg(b \geq c_2)\} \\ A &= T - B \\ S_1 &= \{t \in T \mid t < c_1\} \\ S_2 &= \{t \in T \mid t < c_2\} \\ S_B &= S_2 - S_1 \\ T_A &= \{t \in T \mid c_2 \leq t\} \end{aligned}$$

Define structures  $\mathbf{A}_{c_2}^{c_1} = (A, \leq, \vec{C}, S_1, T_A)$  and  $\mathbf{B}_{c_2}^{c_1} = (B, \leq, \vec{C}, S_B)$  where  $\mathcal{L}_A$  is expansion of  $\mathcal{L}$  by two unary predicates (and  $\mathcal{L}_B$  as defined above). Note that  $c_1, c_2 \notin B$ .

**Definition 3.3.** Fix  $c$  in  $T$ . Let

$$\begin{aligned} B &= \{b \in T \mid \neg(b \geq c) \wedge E(b, c)\} \\ A &= T - B \\ S_1 &= \{t \in T \mid t < c\} \end{aligned}$$

Define structures  $\mathbf{A}_c = (A, \leq, \vec{C})$  and  $\mathbf{B}_c = (B, \leq, \vec{C}, S_1)$  where  $\mathcal{L}_A = \mathcal{L}$  (and  $\mathcal{L}_B$  as defined above). Note that  $c \notin B$ . (cf example 2.6).

**Definition 3.4.** Fix  $c$  in  $T$  and  $S \subseteq T$  a finite subset. Let

$$\begin{aligned} B &= \{b \in T \mid (b > c) \text{ and for all } s \in S \text{ we have } \neg E_c(s, b)\} \\ A &= T - B \\ S_1 &= \{t \in T \mid t \leq c\} \end{aligned}$$

Define structures  $\mathbf{A}_S^c = (A, \leq, \vec{C}, S_1)$  and  $\mathbf{B}_S^c = (B, \leq, \vec{C}, B)$  where  $\mathcal{L}_A$  is expansion of  $\mathcal{L}$  by a single unary predicate (and  $U \in \mathcal{L}_B$  vacuously interpreted by  $B$ ). Note that  $c \notin B$  and  $S \cap B = \emptyset$ .

**Definition 3.5.** Fix  $S \subseteq T$  a finite subset. Let

$$\begin{aligned} B &= \{b \in T \mid \text{for all } s \in S \text{ we have } \neg E(s, b)\} \\ A &= T - B \end{aligned}$$

Define structures  $\mathbf{A}_S = (A, \leq)$  and  $\mathbf{B}_S = (B, \leq, \vec{C}, B)$  where  $\mathcal{L}_A = \mathcal{L}$  (and  $U \in \mathcal{L}_B$  vacuously interpreted by  $B$ ). Note that  $S \cap B = \emptyset$ .

**Lemma 3.6.** *Pairs of structures defined above are all proper subdivisions.*

*Proof.* We only show this holds for the first definition  $\mathbf{A} = \mathbf{A}_{c_2}^{c_1}$  and  $\mathbf{B} = \mathbf{B}_{c_2}^{c_1}$ . Other cases follow by a similar argument.  $A, B$  partition  $T$  by definition, so it is a

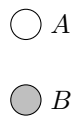


FIGURE 1. Proper subdivision for  $(A, B) = (A_{c_2}^{c_1}, B_{c_2}^{c_1})$

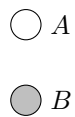
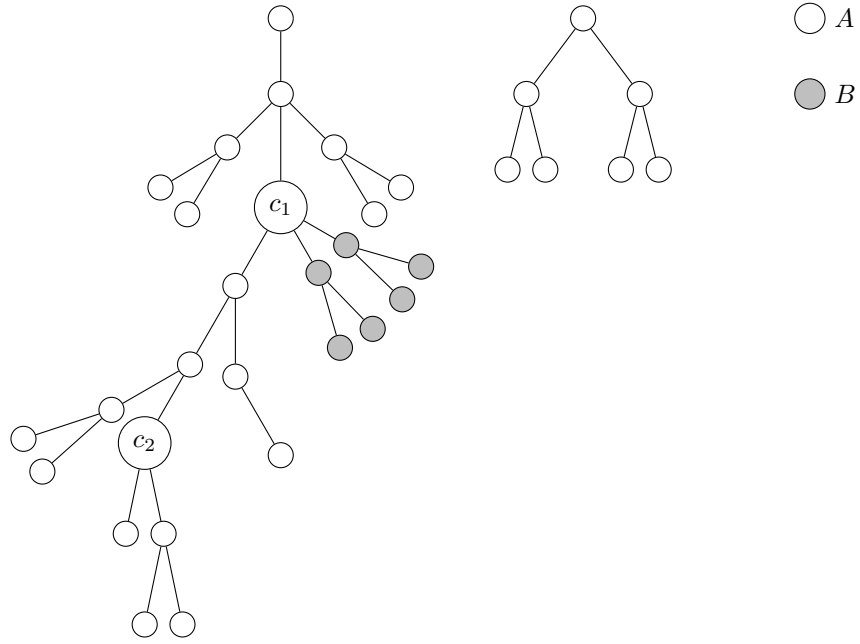
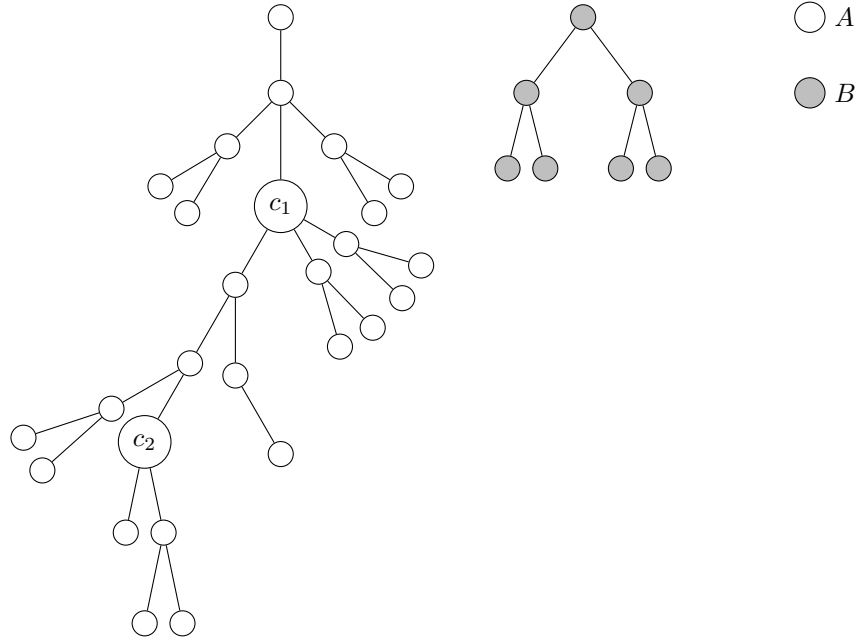


FIGURE 2. Proper subdivision for  $(A, B) = (A_{c_1}, B_{c_1})$

FIGURE 3. Proper subdivision for  $(A, B) = (A_G^{c_1}, B_G^{c_1})$  for  $G = \{c_2\}$ FIGURE 4. Proper subdivision for  $(A, B) = (A_G, B_G)$  for  $G = \{c_1, c_2\}$

subdivision. To show that it is proper by Lemma 2.4 we only need to check that it is 0-proper. Suppose we have

$$\begin{aligned} a &= (a_1, a_2, \dots, a_p) \in A^p \\ a' &= (a'_1, a'_2, \dots, a'_p) \in A^p \\ b &= (b_1, b_2, \dots, b_q) \in B^q \\ b' &= (b'_1, b'_2, \dots, b'_q) \in B^q \end{aligned}$$

with  $(\mathbf{A}, a) \equiv_0 (\mathbf{A}, a')$  and  $(\mathbf{B}, b) \equiv_0 (\mathbf{B}, b')$ . We need to make sure that  $ab$  has the same quantifier free type as  $a'b'$ . Any two elements in  $T$  can be related in the four following ways

$$\begin{aligned} x &= y \\ x &< y \\ x &> y \\ x, y &\text{ are incomparable} \end{aligned}$$

We need to check that the same relations hold for pairs of  $(a_i, b_j), (a'_i, b'_j)$  for all  $i, j$ .

- It is impossible that  $a_i = b_j$  as they come from disjoint sets.
- Suppose  $a_i < b_j$ . This forces  $a_i \in S_1$  thus  $a'_i \in S_1$  and  $a'_i < b'_j$
- Suppose  $a_i > b_j$ . This forces  $b_j \in S_B$  and  $a \in T_A$ , thus  $b'_j \in S_B$  and  $a'_i \in T_A$  so  $a'_i > b'_j$
- Suppose  $a_i$  and  $b_j$  are incomparable. Two cases are possible:
  - $b_j \notin S_B$  and  $a_i \in T_A$ . Then  $b'_j \notin S_B$  and  $a'_i \in T_A$  making  $a'_i, b'_j$  incomparable
  - $b_j \in S_B$ ,  $a_i \notin T_A$ ,  $a_i \notin S_1$ . Similarly this forces  $a'_i, b'_j$  incomparable

Also we need to check that  $ab$  has the same colors as  $a'b'$ . That follows directly from definition.  $\square$

#### 4. MAIN PROOF

Basic idea for the proof is that we are able to divide our parameter space into  $O(n)$  many pieces. Each of  $q$  parameters can come from any of those  $O(n)$  partitions giving us  $O(n)^q$  many choices for parameter configuration. When every parameter coming from a fixed partition the number of definable sets is constant and in fact is uniformly bounded by some  $N$ . This gives us  $NO(n)^q = O(n^q)$  possibilities for different definable sets.

First, we generalize Corollary 2.8. (This is only required for computing vc-density for formulas  $\phi(x, y)$  with  $|y| > 1$ )

**Lemma 4.1.** *Consider a finite collection  $(\mathbf{A}_i, \mathbf{B}_i)_{i \leq n}$  where each  $(\mathbf{A}_i, \mathbf{B}_i)$  is a proper subdivision or a singleton:  $B_i = \{b_i\}$  with  $A_i = T$ . Also assume that all  $\mathbf{B}_i$  have the same language  $\mathcal{L}_B$ . Let  $A = \bigcap_{i \in I} A_i$ . Fix a formula  $\phi(x, y)$  of complexity  $m$ . Let  $N = N(m, |y|, \mathcal{L}_B)$  as in Definition 2.9. Consider any  $B \subseteq T^{|y|}$  of the form*

$$B = B_1^{i_1} \times B_2^{i_2} \times \dots \times B_n^{i_n} \text{ with } i_1 + i_2 + \dots + i_n = |y|$$



(some of the indexes can be zero). Then we have the following bound

$$\phi(A^{|x|}, B) \leq N$$

*Proof.* We show this result by counting types. Suppose we have

$$\begin{aligned} b_1, b'_1 &\in B_1^{i_1} \text{ with } b_1 \equiv_m b'_1 \text{ in } B_1 \\ b_2, b'_2 &\in B_2^{i_2} \text{ with } b_2 \equiv_m b'_2 \text{ in } B_2 \\ &\dots \\ b_n, b'_n &\in B_n^{i_n} \text{ with } b_n \equiv_m b'_n \text{ in } B_n \end{aligned}$$

Then we have

$$\phi(A^{|x|}, b_1, b_2, \dots, b_n) \Leftrightarrow \phi(A^{|x|}, b'_1, b'_2, \dots, b'_n)$$

This is easy to see by applying Corollary 2.8 one by one for each tuple. This works if  $B_i$  is part of a proper subdivision; if it is a singleton then the implication is trivial as  $b_i = b'_i$ . This shows that  $\phi(A^{|x|}, B)$  only depends on the choice of types for the tuples

$$|\phi(A^{|x|}, B)| \leq |S_{B_1, i_1}^m| \cdot |S_{B_2, i_2}^m| \cdot \dots \cdot |S_{B_n, i_n}^m|$$

Now for each type space we have inequality

$$|S_{B_j, i_j}^m| \leq N(m, i_j, \mathcal{L}_B)$$

(For singletons  $|S_{B_j, i_j}^m| = 1 \leq N(m, i_j, \mathcal{L}_B)$ ). Thus we have

$$\begin{aligned} |\phi(A^{|x|}, B)| &\leq N(m, i_1, \mathcal{L}_B) \cdot N(m, i_2, \mathcal{L}_B) \cdot \dots \cdot N(m, i_n, \mathcal{L}_B) \\ &\leq N(m, i_1 + i_2 + \dots + i_n, \mathcal{L}_B) \quad (\text{See Definition 2.9}) \\ &\leq N(m, y, \mathcal{L}_B) = N \end{aligned}$$

as needed.  $\square$

For subdivisions to work out properly we will need to work with subsets closed under meets. We observe that closure under meets doesn't add too many new elements.

**Lemma 4.2.** *Suppose  $S \subseteq T$  is a non-empty finite subset of size  $n$  in a meet tree and  $S' = \text{cl}(S)$  its closure under meets. Then  $|S'| \leq 2n$ .*

*Proof.* We can partition  $S$  into connected components and prove the result separately for each component. Thus we may assume elements of  $S$  lie in the same connected component. We prove the claim by induction on  $n$ . Base case  $n = 1$  is clear. Suppose we have  $S$  of size  $k$  with closure of size at most  $2k - 1$ . Take a new point  $s$ , and look at its meets with all the elements of  $S$ . Pick the smallest one,  $s'$ . Then  $S \cup \{s, s'\}$  is closed under meets.  $\square$

Putting all of those results together we are able to compute vc-density of formulas in meet trees.

**Theorem 4.3.** *Let  $T$  be an infinite (colored) meet tree and  $\phi(x, y)$  a formula with  $|x| = p$  and  $|y| = q$ . Then  $\text{vc}(\phi) \leq q$ .*

*Proof.* Pick a finite subset of  $S_0 \subset T^p$  of size  $n$ . Let  $S_1 \subset T$  consist of coordinates of  $S_0$ . Let  $S \subset T$  be a closure of  $S_1$  under meets. Using Lemma 4.2 we have  $|S_2| \leq 2|S_1| \leq 2p|S_0| = 2pn = O(n)$ . We have  $S_0 \subseteq S^p$ , so  $|\phi(S_0, T^q)| \leq |\phi(S^p, T^q)|$ . Thus it is enough to show  $|\phi(S^p, T^q)| = O(n^q)$ .

Label  $S = \{c_i\}_{i \in I}$  with  $|I| \leq 2pn$ . For every  $c_i$  we construct two partitions in the following way. We have  $c_i$  is either minimal in  $S$  or it has a predecessor in  $S$  (greatest element less than  $c$ ). If it is minimal construct  $(\mathbf{A}_{c_i}, \mathbf{B}_{c_i})$ . If there is a predecessor  $p$  construct  $(\mathbf{A}_{c_i}^p, \mathbf{B}_{c_i}^p)$ . For the second subdivision let  $G$  be all elements in  $S$  greater than  $c_i$  and construct  $(\mathbf{A}_G^c, \mathbf{B}_G^c)$ . So far we have constructed two subdivisions for every  $i \in I$ . Additionally construct  $(\mathbf{A}_S, \mathbf{B}_S)$ . We end up with a finite collection of proper subdivisions  $(\mathbf{A}_j, \mathbf{B}_j)_{j \in J}$  with  $|J| = 2|I| + 1$ . Before we proceed we note the following two lemmas describing our partitions.

**Lemma 4.4.** *For all  $j \in J$  we have  $S \subseteq A_j$ . Thus  $S \subseteq \bigcap_{j \in J} A_j$  and  $S^p \subseteq \bigcap_{j \in J} (A_j)^p$*

*Proof.* Check this for each possible choices of partition. Cases for partitions of the type  $\mathbf{A}_S, \mathbf{A}_G^c, \mathbf{A}_c$  are easy. Suppose we have partition  $(\mathbf{A}, \mathbf{B}) = (\mathbf{A}_{c_2}^{c_1}, \mathbf{B}_{c_2}^{c_1})$ . We need to show that  $B \cap S = \emptyset$ . By construction we have  $c_1, c_2 \notin B$ . Suppose we have some other  $c \in S$  with  $c \in B$ . We have  $E_{c_1}(c_2, c)$  i.e. there is some  $b$  such that  $(b > c_1), (b \leq c_2)$  and  $(b \leq c)$ . Consider the meet  $(c \wedge c_2)$ . We have  $(c \wedge c_2) \geq b > c_1$ . Also as  $\neg(c \geq c_2)$  we have  $(c \wedge c_2) < c_2$ . To summarize  $c_2 > (c \wedge c_2) > c_1$ . But this contradicts our construction as  $S$  is closed under meets, so  $(c \wedge c_2) \in S$  and  $c_1$  is supposed to be a predecessor of  $c_2$  in  $S$ .  $\square$

**Lemma 4.5.**  *$\{B_j\}_{j \in J}$  partition  $T - S$  i.e.  $T = \bigsqcup_{j \in J} B_j \sqcup S$*

*Proof.* This more or less follows from the choice of partitions. Pick any  $b \in S - T$ . Take all elements in  $S$  greater than  $b$  and take the minimal one  $a$ . Take all elements in  $S$  less than  $b$  and take the maximal one  $c$  (possible as  $S$  is closed under meets). Also take all elements in  $S$  incomparable to  $b$  and denote them  $G$ . If both  $a$  and  $c$  exist we have  $b \in \mathbf{B}_c^a$ . If only upper bound exists we have  $b \in \mathbf{B}_G^a$ . If only lower bound exists we have  $b \in \mathbf{B}_c$ . If neither exists we have  $b \in \mathbf{B}_G$ .  $\square$

**Note 4.6.** Those two lemmas imply  $S = \bigcap_{j \in J} A_j$

**Note 4.7.** For one-dimensional case  $q = 1$  we don't need to do any more work. We have partitioned parameter space into  $|J| = O(n)$  many pieces and over each piece the number of definable sets is uniformly bounded. By Corollary 2.8 we have that  $|\phi((A_j)^p, B_j)| \leq N$  for any  $j \in J$  (letting  $N = N(n_\phi, q, \mathcal{L} \cup \{S\})$  where  $n_\phi$  is

complexity of  $\phi$  and  $S$  is a unary predicate). Compute

$$\begin{aligned}
|\phi(S^p, T)| &= \left| \bigcup_{j \in J} \phi(S^p, B_j) \cup \phi(S^p, S) \right| \leq \\
&\leq \sum_{j \in J} |\phi(S^p, B_j)| + |\phi(S^p, S)| \leq \\
&\leq \sum_{j \in J} |\phi((A_j)^p, B_j)| + |S| \leq \\
&\leq \sum_{j \in J} N + |I| \leq \\
&\leq (4pn + 1)N + 2pn = (4pN + 2p)n + N = O(n)
\end{aligned}$$

Basic idea for the general case  $q \geq 1$  is that we have  $q$  parameters and  $|J| = O(n)$  partitions to pick each parameter from giving us  $|J|^q = O(n^q)$  choices for parameter configuration, each giving uniformly constant number of definable subsets of  $S$ . (If every parameter is picked from a fixed partition, Lemma 4.1 provides a uniform bound). This yields  $\text{vc}(\phi) \leq q$  as needed. The rest of the proof is stating this idea formally.

First, we extend our collection of subdivisions  $(\mathbf{A}_j, \mathbf{B}_j)_{j \in J}$  by the following singleton sets. For each  $c_i \in S$  let  $B_i = \{c_i\}$  and  $A_i = T$  and add  $(\mathbf{A}_i, \mathbf{B}_i)$  to our collection with  $\mathcal{L}_B$  the language of  $B_i$  interpreted arbitrarily. We end up with a new collection  $(\mathbf{A}_k, \mathbf{B}_k)_{k \in K}$  indexed by some  $K$  with  $|K| = |J| + |I|$  (we added  $|S|$  new pairs). Now we have that  $B_k$  partition  $T$ , so  $T = \bigsqcup_{k \in K} B_k$  and  $S = \bigcap_{j \in J} A_j = \bigcap_{k \in K} A_k$ . For  $(k_1, k_2, \dots, k_q) = \vec{k} \in K^q$  denote

$$B_{\vec{k}} = B_{k_1} \times B_{k_2} \times \dots \times B_{k_q}$$

Then we have the following identity

$$T^q = \left( \bigsqcup_{k \in K} B_k \right)^q = \bigsqcup_{\vec{k} \in K^q} B_{\vec{k}}$$

Thus we have that  $\{B_{\vec{k}}\}_{\vec{k} \in K^q}$  partition  $T^q$ . Compute

$$\begin{aligned}
|\phi(S^p, T^q)| &= \left| \bigcup_{\vec{k} \in K^q} \phi(S^p, B_{\vec{k}}) \right| \leq \\
&\leq \sum_{\vec{k} \in K^q} |\phi(S^p, B_{\vec{k}})|
\end{aligned}$$

We can bound  $|\phi(S^p, B_{\vec{k}})|$  uniformly using Lemma 4.1.  $(\mathbf{A}_k, \mathbf{B}_k)_{k \in K}$  satisfies the requirements of the lemma and  $B_{\vec{k}}$  looks like  $B$  in the lemma after possibly permuting some variables in  $\phi$ . Applying the lemma we get

$$|\phi(S^p, B_{\vec{k}})| \leq N$$

with  $N$  only depending on  $q$  and complexity of  $\phi$ . We complete our computation

$$\begin{aligned}
|\phi(S^p, T^q)| &\leq \sum_{\vec{k} \in K^q} |\phi(S^p, B_{\vec{k}})| \leq \\
&\leq \sum_{\vec{k} \in K^q} N \leq \\
&\leq |K^q| N^q \leq \\
&\leq (|J| + |I|)^q N \leq \\
&\leq (4pn + 1 + 2pn)^q N = N(6p + 1/n)^q n^q = O(n^q)
\end{aligned}$$

□

**Corollary 4.8.** *In the theory of infinite (colored) meet trees we have  $vc(n) = n$  for all  $n \in \mathbb{N}^+$ .*

We get the general result via an easy application of interpretability.

**Corollary 4.9.** *In the theory of infinite (colored) trees we have  $vc(n) = n$  for all  $n$ .*

*Proof.* Let  $\mathbf{T}'$  be a tree. We can embed it in a larger tree that is closed under meets  $\mathbf{T}' \subset \mathbf{T}$ . Expand  $\mathbf{T}$  by an extra color and interpret it by coloring the subset  $\mathbf{T}'$ . Thus we can interpret  $\mathbf{T}'$  in  $T^1$ . By Corollary 3.17 in [1] we get that  $vc^{\mathbf{T}'}(n) \leq vc^{\mathbf{T}}(1 \cdot n) = n$  thus  $vc^{\mathbf{T}'}(n) = n$  as well. □

#### 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] P. Simon, *On dp-minimal ordered structures*, J. Symbolic Logic 76 (2011), no. 2, 448-460
- [3] Michel Parigot. Théories d'arbres. *Journal of Symbolic Logic*, 47, 1982.  
E-mail address: bobkov@math.ucla.edu