VC-DENSITY FOR TREES

ANTON BOBKOV

ABSTRACT. We show that for the theory of infinite trees we have $\mathrm{vc}(n) = n$ for all n. This generalizes a result of Simon showing that the trees are 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 an NIP theory we can define a vc-function

$$\mathrm{vc} : \mathbb{N} \longrightarrow \mathbb{R} \cup \{\infty\},$$

where $\mathrm{vc}(n)$ measures the worst-case complexity of families of definable sets in an n-dimensional space (see 1.13 below for a precise definition of vc^T). 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. It is not known whether there can be a dp-minimal theory which doesn't satisfy $\mathrm{vc}(n) = n$ (see [1], diagram on pg. 41).

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

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

Here we solve this problem by showing that any model of the theory of trees has $\mathrm{vc}(n) = n$.

Section 1 of the paper consists of a basic introduction to the concepts of VC-dimension and vc-density. In Section 2 we introduce proper subdivisions – the main tool that we use to analyze trees. We also prove some basic properties of proper subdivisions. Section 3 introduces the key constructions of proper subdivisions which will be used in the proof. Section 4 presents the proof $\mathrm{vc}(n) = n$ via subdivisions.

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

The language of trees consists of a single binary predicate $\leq$. The theory of trees states that $\leq$ defines a partial order and that for every element a the set $\{x \mid x \leq a\}$ is linearly ordered by $\leq$. For visualization purposes we assume that trees grow upwards, with the smaller elements on the bottom and the larger elements on the top. If $a \leq b$ we will say that a is below b and b is above a.

Definition 0.2. Work in a tree $\boldsymbol{T} = (T, \leq)$. For $x \in T$ let $T^{\leq x} = \{t \in T \mid t \leq x\}$ denote all the elements below x. Two elements a, b are said to be in same connected component if $T^{\leq a} \cap T^{\leq b}$ is non-empty. The meet of two tree elements a, b is the greatest element of $T^{\leq a} \cap T^{\leq b}$ (if one exists) and is denoted by $a \wedge b$.

The 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.

From now on assume that all trees are colored. We allow our trees to be disconnected (so really, we work with forests) or finite unless otherwise stated.

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 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 $A \subset X$ 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 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 (see 2.13b in [**?**])**.** *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

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

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

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, see [?], [?])**.** *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 the vc-density and dual vc-density of $\mathcal{F}$ as

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

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 $\operatorname{VC}(\Phi)$ 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.

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

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

Functions π^*_Φ and $\pi^*_{\mathcal{F}_\Phi}$ are not equal, as one fixes the size of boolean algebra and another 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 parameter set of size n, we get $|\Phi|n$ elements in the boolean algebra. We check that asymptotically it doesn't matter whether we look at growth of boolean algebra of size n or size $|\Phi|n$.

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

$\square$

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 [**?**]). 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

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

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

2. Proper Subdivisions: Definition and Properties

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

Definition 2.1. Define the complexity of a formula by induction:

$$\text{Complexity}(\text{q.f. formula}) = 0$$

$$\text{Complexity}(\exists x \phi(x)) = \text{Complexity}(\phi(x)) + 1$$

$$\text{Complexity}(\phi \wedge \psi) = \max(\text{Complexity}(\phi), \text{Complexity}(\psi))$$

$$\text{Complexity}(\neg\phi) = \text{Complexity}(\phi)$$

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

Definition 2.2. Let $\boldsymbol{B}$ be a structure, $A \subseteq B$ be a finite parameter set, a, b be tuples in $\boldsymbol{B}$, and m, n be natural numbers.

- $\mathrm{tp}^n_{\boldsymbol{B}}(a/A)$ will stand for the set of all A-formulas of complexity $\leq n$ that are true of a in $\boldsymbol{B}$. If $A = \emptyset$ we may also write this as $\mathrm{tp}^n_{\boldsymbol{B}}(a)$. The subscript $\boldsymbol{B}$ will be omitted as well if it is clear from context. Note that if A is finite, there are finitely many such formulas (up to equivalence). The conjunction of those formulas still has complexity $\leq n$ and so we can just associate a single formula to every type $\mathrm{tp}^n_{\boldsymbol{B}}(a/A)$.
- $\boldsymbol{B} \models a \equiv^n_A b$ means that a, b have the same type with complexity $\leq n$ over A in $\boldsymbol{B}$, i.e., $\mathrm{tp}^n_{\boldsymbol{B}}(a/A) = \mathrm{tp}^n_{\boldsymbol{B}}(b/A)$.
- $S^n_{\boldsymbol{B},m}(A)$ will stand for the set of all m-types of complexity $\leq n$ over A:

$$S^n_{\boldsymbol{B},m}(A) = \left\{\mathrm{tp}^n_{\boldsymbol{B}}(a/A) \mid a \in B^m\right\}.$$

Definition 2.3.
- Let $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{T}$ be structures in some (possibly different) finite relational languages. If the underlying sets A, B of $\boldsymbol{A}, \boldsymbol{B}$ partition the underlying set T of $\boldsymbol{T}$ (i.e. $T = A \sqcup B$), then we say that $(\boldsymbol{A}, \boldsymbol{B})$ is a subdivision of $\boldsymbol{T}$.

- A subdivision $(\boldsymbol{A}, \boldsymbol{B})$ of $\boldsymbol{T}$ is called <u>n-proper</u> if given $p, q \in \mathbb{N}$, $a_1, a_2 \in A^p$ and $b_1, b_2 \in B^q$ with

$$\boldsymbol{A} \models a_1 \equiv_n a_2$$

$$\boldsymbol{B} \models b_1 \equiv_n b_2$$

we have

$$\boldsymbol{T} \models a_1 b_1 \equiv_n a_2 b_2.$$

- A subdivision $(\boldsymbol{A}, \boldsymbol{B})$ of $\boldsymbol{T}$ is called <u>proper</u> if it is n-proper for all $n \in \mathbb{N}$.

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

Proof. We prove that the subdivision is n-proper for all n by induction. The case $n = 0$ is given by the assumption. Suppose we have $\boldsymbol{T} \models \exists x\, \phi^n(x, a_1, b_1)$ where ϕ^n is some formula of complexity n. Let $a \in T$ witness the existential claim, i.e., $\boldsymbol{T} \models \phi^n(a, a_1, b_1)$. We can have $a \in A$ or $a \in B$. Without loss of generality assume $a \in A$. Let $\boldsymbol{p} = \mathrm{tp}^n_{\boldsymbol{A}}(a, a_1)$. Then we have

$$\boldsymbol{A} \models \exists x \ \mathrm{tp}^n_{\boldsymbol{A}}(x, a_1) = \boldsymbol{p}$$

(with $\mathrm{tp}^n_{\boldsymbol{A}}(x, a_1) = \boldsymbol{p}$ a shorthand for $\phi_{\boldsymbol{p}}(x)$ where $\phi_{\boldsymbol{p}}$ is a formula that determines the type $\boldsymbol{p}$). The formula $\mathrm{tp}^n_{\boldsymbol{A}}(x, a_1) = \boldsymbol{p}$ is of complexity $\leq n$ so $\exists x \ \mathrm{tp}^n_{\boldsymbol{A}}(x, a_1) = \boldsymbol{p}$ is of complexity $\leq k + 1$. By the inductive hypothesis we have

$$\boldsymbol{A} \models \exists x \ \mathrm{tp}^n_{\boldsymbol{A}}(x, a_2) = \boldsymbol{p}\,.$$

Let a' witness this existential claim, so that $\mathrm{tp}^n_{\boldsymbol{A}}(a', a_2) = \boldsymbol{p}$, hence $\mathrm{tp}^n_{\boldsymbol{A}}(a', a_2) = \mathrm{tp}^n_{\boldsymbol{A}}(a, a_1)$, that is, $\boldsymbol{A} \models a' a_2 \equiv_n a a_1$. By the inductive hypothesis we therefore have $\boldsymbol{T} \models a a_1 b_1 \equiv_n a' a_2 b_2$; in particular $\boldsymbol{T} \models \phi^n(a', a_2, b_2)$ as $\boldsymbol{T} \models \phi^n(a, a_1, b_1)$, and $\boldsymbol{T} \models \exists x \phi^n(x, a_2, b_2)$. $\square$

This lemma is general, but we will use it specifically applied to (colored) trees. Suppose $\boldsymbol{T}$ is a (colored) tree in some language $\mathcal{L} = \{\leq, \dots\}$ expanding the language of trees by finitely many predicate symbols. Suppose $\boldsymbol{A}, \boldsymbol{B}$ are some structures in languages $\mathcal{L}_A, \mathcal{L}_B$ which expand $\mathcal{L}$, with the $\mathcal{L}$-reducts of $\boldsymbol{A}, \boldsymbol{B}$ substructures of $\boldsymbol{T}$. Furthermore suppose that $(\boldsymbol{A}, \boldsymbol{B})$ is 0-proper. Then by the previous lemma $(\boldsymbol{A}, \boldsymbol{B})$ is a proper subdivision of $\boldsymbol{T}$. From now on all the subdivisions we work with will be of this form.

Example 2.5. Suppose a tree consists of two connected components C_1, C_2. Then those components $(C_1, \leq)$, $(C_2, \leq)$ viewed as substructures form a proper subdivision. To see this we only need to show that this subdivision is 0-proper. But that is immediate as any $c_1 \in C_1$ and $c_2 \in C_2$ are incomparable.

Example 2.6. Fix a tree $\boldsymbol{T}$ in the language $\{\leq\}$, and let $a \in T$. Let $B = \{t \in T \mid a < t\}$, $S = \{t \in T \mid t \leq a\}$, $A = T - B$. Then $(A, \leq, S)$ and $(B, \leq)$ form a proper subdivision, where $\mathcal{L}_A$ has a unary predicate interpreted by S. To see this, again, we show that the subdivision is 0-proper. The only time $a \in A$ and $b \in B$ are comparable is when $a \in S$, and this is captured by the language. (See proof of Lemma 3.7 for more details.)

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

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

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

Corollary 2.8. *Let $\boldsymbol{A}, \boldsymbol{B}$ be a proper subdivision of $\boldsymbol{T}$ and $\phi(x, y)$ be a formula of complexity n. Then $|\phi(A^{|x|}, B^{|y|})| \leq |S^n_{\boldsymbol{B},|y|}|$.*

Proof. Take some $a \in A^{|x|}$ and $b_1, b_2 \in B^{|y|}$ with $\mathrm{tp}^n_{\boldsymbol{B}}(b_1) = \mathrm{tp}^n_{\boldsymbol{B}}(b_2)$. We have $\boldsymbol{B} \models b_1 \equiv_n b_2$ and (trivially) $\boldsymbol{A} \models a \equiv_n a$. Thus we have $\boldsymbol{T} \models ab_1 \equiv_n ab_2$, so $\boldsymbol{T} \models$

$\phi(a, b_1) \leftrightarrow \phi(a, b_2)$. Since a was arbitrary we have $\phi(A^{|x|}, b_1) = \phi(A^{|x|}, b_2)$ as different traces can only come from parameters of different types. Thus $|\phi(A^{|x|}, B^{|y|}) \leq |S^n_{\boldsymbol{B},|y|}|$. □

We note that the size of the type space $|S^n_{\boldsymbol{B},|y|}|$ can be bounded uniformly:

Definition 2.9. Fix a (finite relational) language $\mathcal{L}_B$. Let $N = N(n, m, \mathcal{L}_B)$ be smallest integer such that for any structure $\boldsymbol{B}$ in $\mathcal{L}_B$ we have $|S^n_{\boldsymbol{B},m}| \leq N$. This integer exists as there is a finite number (up to logical equivalence) of possible formulas of complexity $\leq n$ with m free variables. Note that $N(n, m, \mathcal{L}_B)$ is increasing in all variables:

$$n \leq n', m \leq m', \mathcal{L}_B \subseteq \mathcal{L}'_B \Rightarrow N(n, m, \mathcal{L}_B) \leq N(n', m', \mathcal{L}'_B)$$

3. Proper Subdivisions: Constructions

Throughout this section, $\boldsymbol{T}$ denotes a colored meet tree. First, we describe several constructions of proper subdivisions that are needed for the proof.

Definition 3.1. We use $E(b, c)$ to express that b and c are in the same connected component:

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

Definition 3.2. Given an element a of the tree $\boldsymbol{T}$ we call the set of all elements above a, i.e. the set $T^{>a} = \{x \mid x > a\}$, the closed cone above a. Connected components of that cone can be thought of as open cones above a. With that interpretation in mind, the notation $E_a(b, c)$ means that b and c are in the same open cone above a. More formally:

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

Fix a language $\mathcal{L}$ for a colored tree $\mathcal{L} = \{\leq, C_1, \ldots C_n\}$. Put $\vec{C} = \{C_1, \ldots C_n\}$. Given $A \subseteq T$ we put $\vec{C} \cap A = \{C_1 \cap A, \ldots, C_n \cap A\}$. In that case $(A, \leq, \vec{C} \cap A)$ is a substructure of $\boldsymbol{T}$ (here as usual by abuse of notation $\leq$ denotes the restriction of the ordering $\leq$ of T to A.) For the following four definitions fix $\mathcal{L}_B = \mathcal{L} \cup \{U\}$ be an expansion of language $\mathcal{L}$ by a unary predicate. We refer to the figures below for illustrations of these definitions.

Definition 3.3. Fix $c_1 < c_2$ in T. Let

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

Define structures $\boldsymbol{A}_{c_2}^{c_1} = (A, \leq, \vec{C} \cap A, T^{<c_1}, T^{\geq c_2})$ in a language expanding $\mathcal{L}$ by two unary predicate symbols and $\boldsymbol{B}_{c_2}^{c_1} = (B, \leq, \vec{C} \cap B, S_B)$ with language $\mathcal{L}_B$ (as defined above). Note that $c_1, c_2 \notin B$.

Definition 3.4. Fix c in T. Let

$$\begin{aligned} B &= \{b \in T \mid \neg(b \geq c) \wedge E(b, c)\}, \\ A &= T - B, \\ T^{<c_1} &= \{t \in T \mid t < c\}. \end{aligned}$$

Define structures $\boldsymbol{A}_c = (A, \leq, \vec{C} \cap A)$ with language $\mathcal{L}$ and $\boldsymbol{B}_c = (B, \leq, \vec{C} \cap B, T^{<c_1})$ with language $\mathcal{L}_B$. Note that $c \notin B$. (cf. example 2.6).

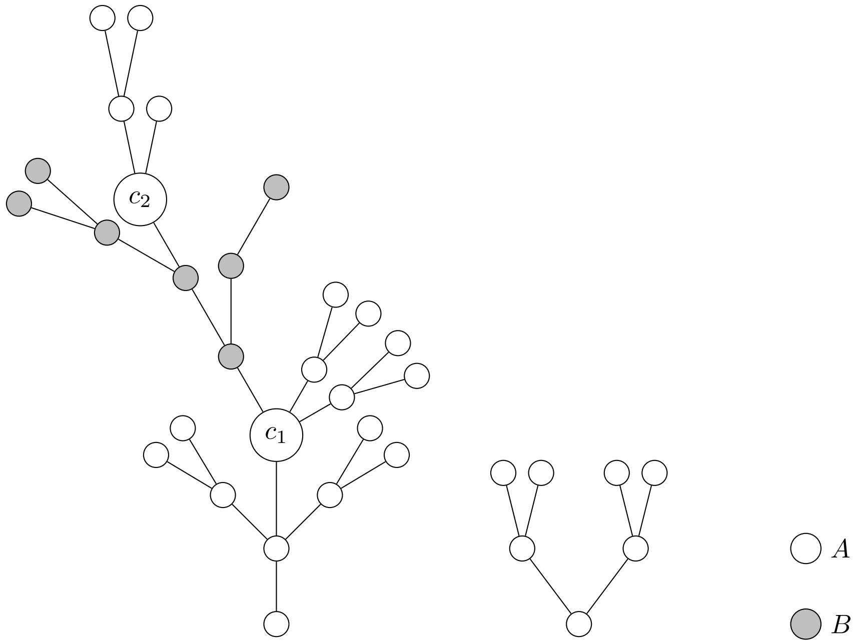

FIGURE 1. Proper subdivision $(\boldsymbol{A}, \boldsymbol{B}) = (\boldsymbol{A}_{c_2}^{c_1}, \boldsymbol{B}_{c_2}^{c_1})$

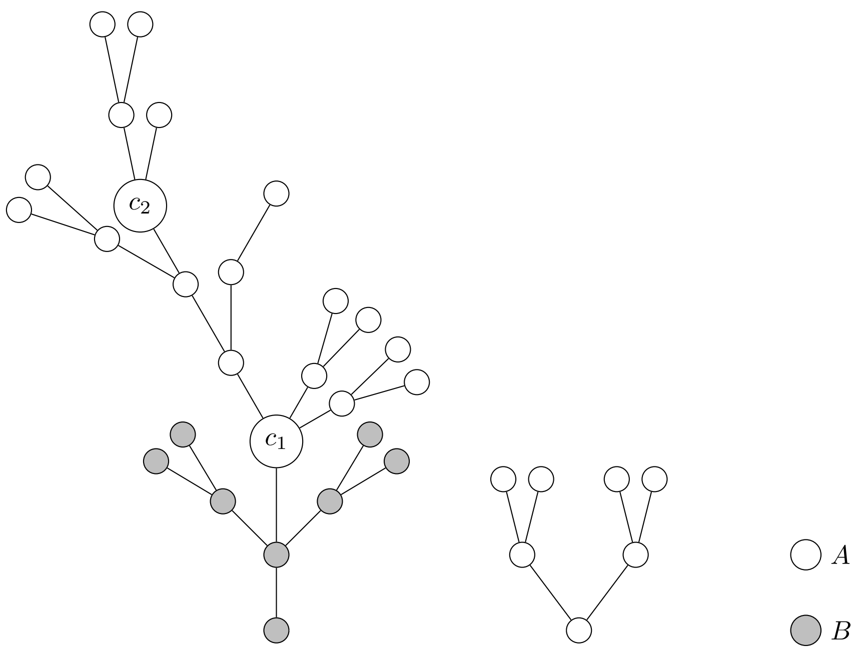

FIGURE 2. Proper subdivision $(\boldsymbol{A}, \boldsymbol{B}) = (\boldsymbol{A}_{c_1}, \boldsymbol{B}_{c_1})$

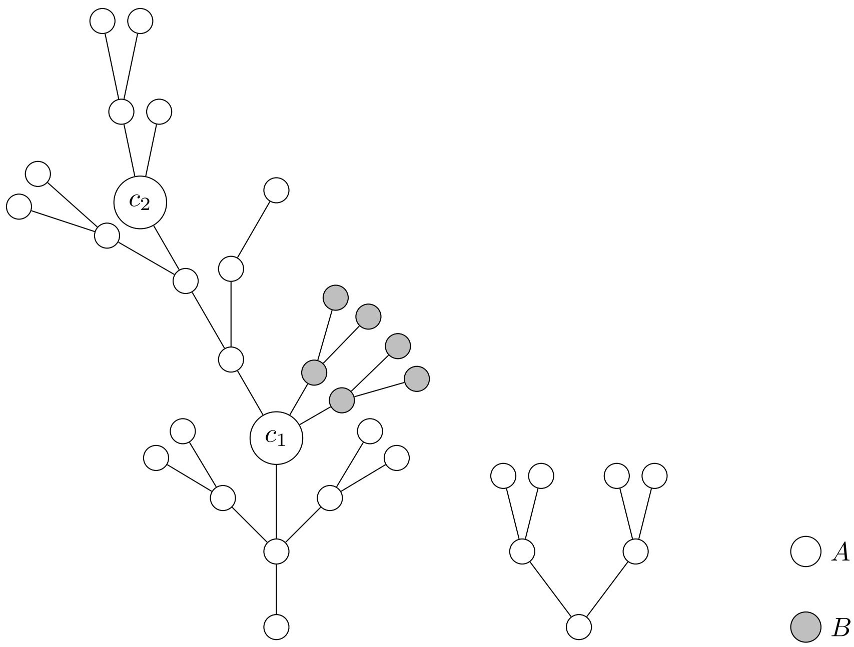

FIGURE 3. Proper subdivision $(\boldsymbol{A}, \boldsymbol{B}) = (\boldsymbol{A}_S^{c_1}, \boldsymbol{B}_S^{c_1})$ for $S = \{c_2\}$

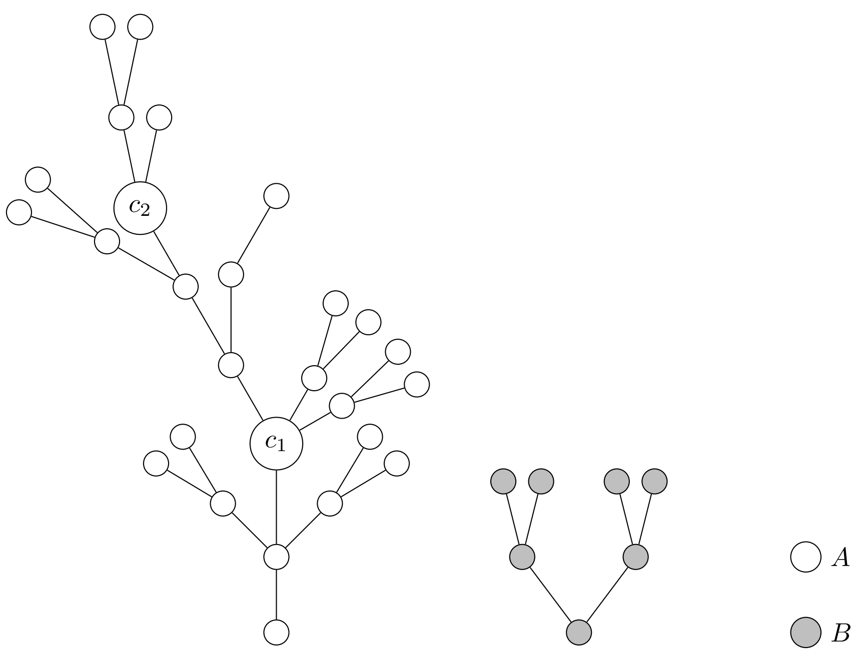

FIGURE 4. Proper subdivision $(\boldsymbol{A}, \boldsymbol{B}) = (\boldsymbol{A}_S, \boldsymbol{B}_S)$ for $S = \{c_1, c_2\}$

Definition 3.5. Fix c in T and a finite subset $S \subseteq T$. Let

$$\begin{aligned} B &= \{b \in T \mid (b > c) \text{ and for all } s \in S \text{ we have } \neg E_c(s,b)\}, \\ A &= T - B, \\ T^{\leq c_1} &= \{t \in T \mid t \leq c\}. \end{aligned}$$

Define structures $\boldsymbol{A}_S^c = (A, \leq, \vec{C} \cap A, T^{\leq c_1})$ and $\boldsymbol{B}_S^c = (B, \leq, \vec{C} \cap B, B)$ both with language $\mathcal{L}_B$. Note that $c \notin B$ and $S \cap B = \emptyset$.

Definition 3.6. Fix a finite subset $S \subseteq T$. 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 $\boldsymbol{A}_S = (A, \leq, \vec{C} \cap A)$ with language $\mathcal{L}$ and $\boldsymbol{B}_S = (B, \leq, \vec{C} \cap B, B)$ with language $\mathcal{L}_B$. (cf. example 2.5)

Note that we forced the structures $\boldsymbol{B}_{c_2}^{c_1}, \boldsymbol{B}_c, \boldsymbol{B}_S^c, \boldsymbol{B}_S$ to have the same language $\mathcal{L}_B$. This is done for uniformity to simplify lemma 4.1. By comparison, the corresponding $\boldsymbol{A}$ structures have different languages.

Lemma 3.7. *The pairs of structures defined above are all proper subdivisions of* $\boldsymbol{T}$.

Proof. We only show this holds for the pair $(\boldsymbol{A}, \boldsymbol{B}) = (\boldsymbol{A}_{c_2}^{c_1}, \boldsymbol{B}_{c_2}^{c_1})$. The other cases follow by a similar argument. The sets A, B partition T by definition, so (A, B) is a subdivision of $\boldsymbol{T}$. To show that it is proper, by Lemma 2.4 we only need to check that it is 0-proper. Suppose we have

$$a = (a_1, a_2, \ldots, a_p) \in A^p,$$

$$a' = (a'_1, a'_2, \ldots, a'_p) \in A^p,$$

$$b = (b_1, b_2, \ldots, b_q) \in B^q,$$

$$b' = (b'_1, b'_2, \ldots, b'_q) \in B^q.$$

with $\boldsymbol{A} \models a \equiv_0 a'$ and $\boldsymbol{B} \models b \equiv_0 b'$. We need to show that ab has the same quantifier-free type in $\boldsymbol{T}$ as $a'b'$. Any two elements in T can be related in the four following ways:

$$x = y,$$

$$x < y,$$

$$x > y, \text{ or}$$

$$x, y \text{ are incomparable.}$$

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

- It is impossible that $a_i = b_j$ as they come from disjoint sets.
- Suppose $a_i < b_j$. This forces $a_i \in T^{<c_1}$ thus $a'_i \in T^{<c_1}$ and $a'_i < b'_j$.
- Suppose $a_i > b_j$. This forces $b_j \in S_B$ and $a \in T^{\geq c_2}$, thus $b'_j \in S_B$ and $a'_i \in T^{\geq c_2}$, so $a'_i > b'_j$.
- Suppose a_i and b_j are incomparable. Two cases are possible:
  - $b_j \notin S_B$ and $a_i \in T^{\geq c_2}$. Then $b'_j \notin S_B$ and $a'_i \in T^{\geq c_2}$ making a'_i, b'_j incomparable.
  - $b_j \in S_B$, $a_i \notin T^{\geq c_2}$, $a_i \notin T^{<c_1}$. Similarly this forces a'_i, b'_j to be incomparable.

Also we need to check that ab has the same colors as $a'b'$. But that is immediate as having the same color in a substructure means having the same color in the tree. $\square$

4. Main proof

The basic idea for the proof is as follows. Suppose we have a formula with q parameters over a parameter set of size n. We are able to split our parameter space into $O(n)$ many partitions. 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 is coming from a fixed partition the number of definable sets is constant and in fact is uniformly bounded above by some N. This gives us at most $N \cdot O(n)^q$ possibilities for different definable sets.

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

Lemma 4.1. *Consider a finite collection $(\boldsymbol{A}_i, \boldsymbol{B}_i)_{i \leq n}$ satisfying the following properties:*

- *$(\boldsymbol{A}_i, \boldsymbol{B}_i)$ is either a proper subdivision of $\boldsymbol{T}$ or $A_i = T$ and $B_i = \{b_i\}$,*
- *all $\boldsymbol{B}_i$ have the same language $\mathcal{L}_B$,*
- *sets $\{B_i\}_{i \leq n}$ are pairwise disjoint.*

Let $A = \bigcap_{i \in I} A_i$. Fix a formula $\phi(x, y)$ of complexity m . Let $N = N(m, |y|, \mathcal{L}_B)$ be as in Definition 2.9. Consider any $B \subseteq T^{|y|}$ of the form

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

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

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

Proof. We show this result by counting types.

Claim 4.2. *Suppose we have*

$$b_1, b_1' \in B_1^{i_1} \text{ with } b_1 \equiv_m b_1' \text{ in } \boldsymbol{B}_1,$$

$$b_2, b_2' \in B_2^{i_2} \text{ with } b_2 \equiv_m b_2' \text{ in } \boldsymbol{B}_2,$$

$$\dots$$

$$b_n, b_n' \in B_n^{i_n} \text{ with } b_n \equiv_m b_n' \text{ in } \boldsymbol{B}_n\,.$$

Then

$$\phi(A^{|x|}, b_1, b_2, \dots b_n) \iff \phi(A^{|x|}, b_1', b_2', \dots b_n').$$

Proof. Define $\bar{b}_i = (b_1, \dots, b_i, b_{i+1}', \dots, b_n') \in B$ for $i \in [0..n]$. (That is, a tuple where first i elements are without prime, and elements after that are with a prime.) We have $\phi(A^{|x|}, \bar{b}_i) \iff \phi(A^{|x|}, \bar{b}_{i+1})$ as either $(\boldsymbol{A}_{i+1}, \boldsymbol{B}_{i+1})$ is m-proper or $\boldsymbol{B}_{i+1}$ is a singleton, and the implication is trivial. (Notice that $b_i \in \boldsymbol{A}_j$ for $j \neq i$ by disjointness assumption.) Thus, by induction we get $\phi(A^{|x|}, \bar{b}_0) \iff \phi(A^{|x|}, \bar{b}_n)$ as needed. □

Thus $\phi(A^{|x|}, B)$ only depends on the choice of the types for the tuples:

$$|\phi(A^{|x|}, B)| \leq |S^m_{\boldsymbol{B}_1, i_1}| \cdot |S^m_{\boldsymbol{B}_2, i_2}| \cdot \dots \cdot |S^m_{\boldsymbol{B}_n, i_n}|$$

Now for each type space we have an inequality

$$|S^m_{\boldsymbol{B}_j, i_j}| \leq N(m, i_j, \mathcal{L}_B) \leq N(m, |y|, \mathcal{L}_B) \leq N$$

(For singletons $|S^m_{\boldsymbol{B}_j, i_j}| = 1 \leq N$). Only non-zero indices contribute to the product and there are at most $|y|$ of those (by the equality $i_1 + i_2 + \ldots + i_n = |y|$). Thus we have

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

as needed. □

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

Lemma 4.3. *Suppose $S \subseteq T$ is a finite subset of size $n \geq 1$ in a meet tree and S' is its closure under meets. Then $|S'| \leq 2n - 1$.*

Proof. We can partition S into connected components and prove the result separately for each component. Thus we may assume all elements of S lie in the same connected component. We prove the claim by induction on n. The 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. □

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

Theorem 4.4. *Let $\boldsymbol{T}$ be an infinite (colored) meet tree and $\phi(x, y)$ a formula with $|x| = p$ and $|y| = q$. Then $\mathrm{vc}(\phi) \leq q$.*

Proof. Pick a finite subset of $S_0 \subset T^p$ of size n. Let $S_1 \subset T$ consist of the components of the elements of S_0. Let $S \subset T$ be the closure of S_1 under meets. Using Lemma 4.3 we have $|S| \leq 2|S_1| \leq 2p|S_0| = 2pn = O(n)$. 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 that c_i is either minimal in S or it has a predecessor in S (greatest element less than c). If it is minimal, construct $(\boldsymbol{A}_{c_i}, \boldsymbol{B}_{c_i})$. If there is a predecessor p, construct $(\boldsymbol{A}^p_{c_i}, \boldsymbol{B}^p_{c_i})$. For the second subdivision let G be all the elements in S greater than c_i and construct $(\boldsymbol{A}^c_G, \boldsymbol{B}^c_G)$. So far we have constructed two subdivisions for every $i \in I$. Additionally construct $(\boldsymbol{A}_S, \boldsymbol{B}_S)$. We end up with a finite collection of proper subdivisons $(\boldsymbol{A}_j, \boldsymbol{B}_j)_{j \in J}$ with $|J| = 2|I| + 1$. Before we proceed, we note the following two lemmas describing our partitions.

Lemma 4.5. *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 choice of partition. Cases for partitions of the type $\boldsymbol{A}_S, \boldsymbol{A}^c_G, \boldsymbol{A}_c$ are easy. Suppose we have a partition $(\boldsymbol{A}, \boldsymbol{B}) = (\boldsymbol{A}^{c_1}_{c_2}, \boldsymbol{B}^{c_1}_{c_2})$. 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. □

Lemma 4.6. *$\{B_j\}_{j \in J}$ is a disjoint partition of $T - S$ 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 the elements in S greater than b and take the minimal one a. Take all the elements in S less than b and take the maximal one c (possible as S is closed under meets). Also take all the elements in S incomparable to b and denote them G. If both a and c exist we have $b \in \boldsymbol{B}^a_c$. If only the upper bound exists we have $b \in \boldsymbol{B}^a_G$. If only the lower bound exists we have $b \in \boldsymbol{B}_c$. If neither exists we have $b \in \boldsymbol{B}_G$. □

Note 4.7. Those two lemmas imply $S = \bigcap_{j \in J} A_j$.

Note 4.8. For one-dimensional case $q = 1$ we don't need to do any more work. We have partitioned the 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_ϕ is the complexity of ϕ 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 $|J| = O(n)$ many partitions to pick each parameter from giving us $|J|^q = O(n^q)$ choices for the parameter configuration, each giving a 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 $\mathrm{vc}(\phi) \leq q$ as needed. The rest of the proof is stating this idea formally.

First, we extend our collection of subdivisions $(\boldsymbol{A}_j,\boldsymbol{B}_j)_{j\in J}$ by the following singleton sets. For each $c_i \in S$ let $B_i = \{c_i\}$ and $A_i = T$ and add $(\boldsymbol{A}_i,\boldsymbol{B}_i)$ to our collection with $\mathcal{L}_B$ the language of B_i interpreted arbitrarily. We end up with a new collection $(\boldsymbol{A}_k,\boldsymbol{B}_k)_{k\in K}$ indexed by some K with $|K| = |J| + |I|$ (we added $|S|$ new pairs). Now $\{B_k\}_{k\in K}$ partitions 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

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

Then we have the following identity

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

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. $(\boldsymbol{A}_k,\boldsymbol{B}_k)_{k\in K}$ satisfies the requirements of the lemma and $B_{\vec{k}}$ looks like B in the lemma after possibly permuting some variables in ϕ. Applying the lemma we get

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

with N only depending on q and complexity of ϕ. We complete our computation

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

□

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

We get the general result for the trees that aren't necessarily meet trees via an easy application of interpretability.

Corollary 4.10. *In the theory of infinite (colored) trees we have $vc(n)=n$ for all n.*

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

This settles the question of *vc*-function for trees. Lacking a quantifier elimination result, a lot is still not known. One can try to adapt these techniques to compute the vc-density of a fixed formula, and see if it can take non-integer values. It is also not known whether trees have VC 1 property (see [1] 5.2 for the definition). Our techniques can be used to show that VC 2 property holds but this doesn't give the optimal vc-function.

One can also try to apply similar techniques to more general classes of partially ordered sets. For example, vc-density values are not known for lattices. Similarly, dropping the order, one can look at nicely behaved families of graphs, such as planar graphs or flat graphs. Those are known to be dp-minimal, so one would expect a simple vc-function. It is this author's hope that the techniques developed in this paper can be adapted to yield fruitful results for a more general class of structures.

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] 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, 841-853

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