

# VC-DENSITY IN AN ADDITIVE REDUCT OF THE $P$ -ADIC NUMBERS

ANTON BOBKOV

ABSTRACT. Aschenbrenner et. al. computed a bound  $\text{vc}(n) \leq 2n - 1$  for the vc-density function in the field of  $p$ -adic numbers, but it is not known to be optimal. In this paper we investigate a certain  $P$ -minimal additive reduct of the field of  $p$ -adic numbers and use a cell decomposition result of Leenknegt to compute an optimal bound  $\text{vc}(n) = n$  for that structure.

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

$$\text{vc}^T = \text{vc} : \mathbb{N} \longrightarrow \mathbb{N}$$

where  $\text{vc}(n)$  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  $\text{vc}^T$ ). The simplest possible behavior is  $\text{vc}(n) = n$  for all  $n$ , satisfied, for example, if  $T$  is o-minimal. For  $T = \text{Th}(\mathbb{Q}_p)$ , the paper [1] computes an upper bound for this function to be  $2n - 1$ , and it is not known whether this is optimal. This same bound holds in any reduct of the field of  $p$ -adic numbers, but one may expect that the simplified structure of the reduct would allow a better bound. In [2], Leenknegt provides a cell decomposition result for a certain  $P$ -minimal additive reduct of the field of  $p$ -adic numbers. Using this result, in this paper we improve the bound for the vc-function, showing that in Leenknegt's structure  $\text{vc}(n) = n$ .

Section 1 defines vc-density and states some basic lemmas about it. A more in depth exposition of vc-density can be found in [1]. Section 2 defines and states some

basic facts about the theory of  $p$ -adic numbers. Here we also introduce the reduct which we will be working with. Section 3 sets up basic definitions and lemmas that will be needed for the proof. We define trees and intervals and show how they help with vc-density calculations. Section 4 concludes the proof.

Throughout the paper, variables and tuples of elements will be simply denoted as  $x, y, a, b, \dots$ . We will occasionally write  $\vec{a}$  instead of  $a$  for a tuple in  $\mathbb{Q}_p^n$  to emphasize it as an element of  $\mathbb{Q}_p$ -vector space  $\mathbb{Q}_p^n$ . We denote the arity of a tuple  $x$  of variables by  $|x|$ .

### 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  $\text{VC}(X, \mathcal{F})$ .

**Note 1.3.** We may drop  $X$  from the  $\text{VC}(X, \mathcal{F})$ , 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  $\text{VC}^*(\mathcal{F})$ . (As before, this notion doesn't depend on  $X$ .)

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

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

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  $\pi_{\mathcal{F}}: \mathbb{N} \rightarrow \mathbb{N}$  and the dual shatter function  $\pi_{\mathcal{F}}^*: \mathbb{N} \rightarrow \mathbb{N}$  of  $\mathcal{F}$  by

$$\begin{aligned} \pi_{\mathcal{F}}(n) &= \max \{ |A \cap \mathcal{F}| \mid A \subset X \text{ and } |A| = n \} \\ \pi_{\mathcal{F}}^*(n) &= \max \{ \text{atoms}(B) \mid B \subset \mathcal{F}, |B| = n \} \end{aligned}$$

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

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

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

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

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

$$\begin{aligned} \text{vc}(\mathcal{F}) &= \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{F}}(n)}{\log n} \in \mathbb{R}^{\geq 0} \cup \{+\infty\}, \\ \text{vc}^*(\mathcal{F}) &= \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{F}}^*(n)}{\log n} \in \mathbb{R}^{\geq 0} \cup \{+\infty\}. \end{aligned}$$

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

We will also refer to the vc-density and VC-dimension of a single formula  $\phi$  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  $\Phi$ -type over  $B$  is a maximal consistent collection of functions of the form  $\phi(x, b)$  or  $\neg\phi(x, b)$  where  $\phi \in \Phi$  and  $b \in B$ .

**Lemma 1.11.**

$$\text{vc}^*(\Phi) = \text{degree of polynomial growth of } \pi_\Phi^*(n) = \limsup_{n \rightarrow \infty} \frac{\log \pi_\Phi^*(n)}{\log n}$$

*Proof.*

$$\begin{aligned} \pi_{\mathcal{F}_\Phi}^*(n) &\leq \pi_\Phi^*(n) \leq \pi_{\mathcal{F}_\Phi}^*(|\Phi|n) \\ \text{vc}^*(\Phi) &\leq \limsup_{n \rightarrow \infty} \frac{\log \pi_\Phi^*(n)}{\log n} \leq \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{F}_\Phi}^*(|\Phi|n)}{\log n} = \\ &= \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{F}_\Phi}^*(|\Phi|n)}{\log |\Phi|n} \frac{\log |\Phi|n}{\log n} = \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{F}_\Phi}^*(|\Phi|n)}{\log |\Phi|n} \leq \\ &\leq \limsup_{n \rightarrow \infty} \frac{\log \pi_{\mathcal{F}_\Phi}^*(n)}{\log n} = \text{vc}^*(\Phi) \end{aligned}$$

□

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

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

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

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

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

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

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

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

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

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

## 2. $P$ -ADIC NUMBERS

The field  $\mathbb{Q}_p$  of  $p$ -adic numbers is often studied in the language of Macintyre

$$\mathcal{L}_{Mac} = \{0, 1, +, -, \cdot, |, \{P_n\}_{n \in \mathbb{N}}\}$$

which is a language  $\{0, 1, +, -, \cdot\}$  of fields together with unary predicates  $P_n$  interpreted in  $\mathbb{Q}_p$  so as to satisfy

$$P_n x \leftrightarrow \exists y \, y^n = x$$

and a divisibility relation where  $a|b$  holds in  $\mathbb{Q}_p$  when  $\text{val } a \leq \text{val } b$ .

Note that  $P_n \setminus \{0\}$  is a multiplicative subgroup of  $\mathbb{Q}_p$  with finitely many cosets.

**Theorem 2.1** (Macintyre '76). *The  $\mathcal{L}_{Mac}$ -structure  $\mathbb{Q}_p$  has quantifier elimination.*

There is also a cell decomposition result:

**Definition 2.2.** Define  $k$ -cell recursively as follows. 0-cell is a singleton subset of  $\mathbb{Q}_p$ . A  $(k+1)$ -cell is a subset of  $\mathbb{Q}_p^{k+1}$  of the following form:

$$\{(x, t) \in D \times \mathbb{Q}_p \mid \text{val } a_1(x) \square_1 \text{val}(t - c(x)) \square_2 \text{val } a_2(x), t - c(x) \in \lambda P_n\}$$

where  $D$  is a  $k$ -cell,  $a_1(x), a_2(x), c(x)$  are definable functions  $D \rightarrow \mathbb{Q}_p$ ,  $\square_i$  is  $<, \leq$  or no condition, and  $\lambda \in \mathbb{Q}_p$ .

**Theorem 2.3** (Denef '84). *Any definable subset of  $\mathbb{Q}_p^n$  defined by an  $\mathcal{L}_{Mac}$ -formula decomposes into a finite disjoint union of  $n$ -cells.*

In [1], Aschenbrenner, Dolich, Haskell, Macpherson, and Starchenko show that this structure satisfies  $\text{vc}(n) \leq 2n - 1$ , however it is not known whether this bound is optimal.

In [2], Leenknegt analyzes the reduct of  $\mathbb{Q}_p$  to the language

$$\mathcal{L}_{aff} = \left\{ 0, 1, +, -, \{\bar{c}\}_{c \in \mathbb{Q}_p}, |, \{Q_{m,n}\}_{m,n \in \mathbb{N}} \right\}$$

where  $\bar{c}$  denotes a scalar multiplication by  $c$ ,  $a|b$  as above stands for  $\text{val } a \leq \text{val } b$ , and  $Q_{m,n}$  is a unary predicate interpreted as

$$Q_{m,n} = \bigcup_{k \in \mathbb{Z}} p^{km} (1 + p^n \mathbb{Z}_p).$$

Note that  $Q_{m,n} \setminus \{0\}$  is a subgroup of the multiplicative group of  $\mathbb{Q}_p$  with finitely many cosets. One can check that these extra relation symbols are definable in the  $\mathcal{L}_{Mac}$ -structure  $\mathbb{Q}_p$ . The paper [2] provides a cell decomposition result with the following cells:

**Definition 2.4.** A 0-cell is a singleton subset of  $\mathbb{Q}_p$ . A  $(k+1)$ -cell is a subset of  $\mathbb{Q}_p^{k+1}$  of the following form:

$$\{(x, t) \in D \times \mathbb{Q}_p \mid \text{val } a_1(x) \square_1 \text{val}(t - c(x)) \square_2 \text{val } a_2(x), t - c(x) \in \lambda Q_{m,n}\}$$

where  $D$  is a  $k$ -cell, called the base of the cell,  $a_1(x), a_2(x), c(x)$  are polynomials of degree  $\leq 1$ , called the defining polynomials  $\square_1, \square_2$  is  $<$  or no condition, and  $\lambda \in \mathbb{Q}_p$ . We call  $\mathbb{Q}_{m,n}$  the defining predicate.

**Theorem 2.5** (Leenknecht '12). *Any definable subset of  $\mathbb{Q}_p^n$  defined by an  $\mathcal{L}_{aff}$ -formula decomposes into a finite disjoint union of  $n$ -cells.*

Moreover, [2] shows that  $(\mathbb{Q}_p, \mathcal{L}_{aff})$  is a  $P$ -minimal reduct, that is, the one-dimensional definable sets of  $(\mathbb{Q}_p, \mathcal{L}_{aff})$  coincide with the one-dimensional definable sets in the full structure  $(\mathbb{Q}_p, \mathcal{L}_{Mac})$ .

The main result of this paper is the computation of the vc-function for this structure:

**Theorem 2.6.**  *$(\mathbb{Q}_p, \mathcal{L}_{aff})$  has  $vc(n) = n$ .*

### 3. KEY LEMMAS AND DEFINITIONS

To show that  $vc(n) = n$  it suffices to bound  $vc^*(\phi) \leq |x|$  for every  $\mathcal{L}_{aff}$ -formula  $\phi(x; y)$ . Fix such a formula  $\phi(x; y)$ . Instead of working with it directly, we simplify it using quantifier elimination. The required quantifier elimination result can be easily obtained from cell decomposition:

**Lemma 3.1.** *Any formula  $\phi(x; y)$  in  $(\mathbb{Q}_p, \mathcal{L}_{aff})$  can be written as a boolean combination of formulas from the following collection*

$$\begin{aligned} \Phi(x; y) = & \{ \text{val}(p_i(x) - c_i(y)) < \text{val}(p_j(x) - c_j(y)) \}_{i,j \in I} \cup \\ & \{ p_i(x) - c_i(y) \in \lambda_k \mathbb{Q}_{m,n} \}_{i \in I, k \in K} \end{aligned}$$

where  $I, K$  are finite index sets, each  $p_i$  is a degree  $\leq 1$  polynomial in  $x$  without a constant term, each  $c_i$  is a degree  $\leq 1$  polynomial in  $y$ , and  $\lambda_k \in \mathbb{Q}_p$ .

*Proof.* Let  $l = |x| + |y|$ . Partition the subset of  $\mathbb{Q}_p^l$  defined by  $\phi$  to obtain  $\mathcal{D}^l$ , a collection of  $l$ -cells. Let  $\mathcal{D}^{l-1}$  be the collection of the bases of the cells in  $\mathcal{D}^l$ . Similarly, construct by induction  $\mathcal{D}^j$  for each  $0 \leq j < l$ , where  $\mathcal{D}^j$  is the collection



of  $j$ -cells which are the bases of cells in  $\mathcal{D}^{j+1}$ .

$$m = \prod \{m' \mid Q_{m',n'} \text{ is the defining predicate of a cell in } \mathcal{D}^j \text{ for } 0 \leq j \leq l\}$$

$$n = \max \{n' \mid Q_{m',n'} \text{ is the defining predicate of a cell in } \mathcal{D}^j \text{ for } 0 \leq j \leq l\}$$

This way if  $a, a'$  are in the same coset of  $Q_{m',n'}$  then they are in the same coset of  $Q_{m,n}$ . Choose  $\{\lambda_k\}_{k \in K}$  to go over all the cosets of  $Q_{m,n}$ . Let  $q_i(x, y)$  enumerate all of the defining polynomials  $a_1(x), a_2(x), t - c(x)$  that show up in the cells of  $\mathcal{D}^j$  for any  $j$ . All if those are all polynomials of degree  $\leq 1$  in variables  $x, y$ . We can split each of them as  $q_i(x, y) = p_i(x) - c_i(y)$  where the constant term goes into  $c_i$ . This gives us the appropriate finite collection of formulas  $\Phi$ . From the cell decomposition it is easy to see that when  $a, a'$  have the same  $\Phi$ -type, then they have the same  $\phi$ -type. Thus  $\phi$  can be written as a boolean combination of formulas from  $\Phi$ .  $\square$

**Lemma 3.2.** *Let  $\Phi(x; y)$  be a finite collection of formulas. If  $\phi$  can be written as a boolean combination of formulas from  $\Phi$  then*

$$\text{vc}^*(\Phi) \leq r \implies \text{vc}^*(\phi) \leq r \text{ for all } r \in \mathbb{R}$$

*Proof.* If  $a, a'$  have the same  $\Phi$ -type over  $B$ , then they have the same  $\phi$ -type over  $B$ , where  $B$  is some parameter set. Therefore the number of  $\phi$ -types is bounded by the number of  $\Phi$ -types. The bound follows from Lemma 1.11.  $\square$

For the remainder of the paper fix  $\Phi(x; y)$  to be the collection of formulas defined by Lemma 3.1. By the previous lemma, to show that  $\text{vc}^*(\phi) \leq |x|$ , it suffices to bound  $\text{vc}^*(\Phi) \leq |x|$ . More precisely, it is sufficient to show that if there is a parameter set  $B$  of size  $N$  then the number of  $\Phi$ -types over  $B$  is  $O(N^{|x|})$ . Fix such a parameter set  $B$  and work with it from now on. We will compute a bound for the number of  $\Phi$ -types over  $B$ .

Consider a set  $T = T(\Phi, B) = \{c_i(b) \mid b \in B, i \in I\} \subset \mathbb{Q}_p$ . In this definition  $B$  is the parameter set that we have fixed and  $c_i(b)$  come from the collection of formulas  $\Phi$  from the quantifier elimination above. View  $T$  as a tree as follows:

**Definition 3.3.**

- For  $c \in \mathbb{Q}_p, \alpha \in \mathbb{Z}$  define a ball

$$B(c, \alpha) = \{c' \in \mathbb{Q}_p \mid \text{val}(c' - c) > \alpha\}.$$

We also let  $B(c, -\infty) = \mathbb{Q}_p$  and  $B(c, +\infty) = \emptyset$ .

- Define a collection of balls  $\mathcal{B} = \{B(t_1, \text{val}(t_1 - t_2))\}_{t_1, t_2 \in T}$ . Note that  $\mathcal{B}$  is a (directed) boolean algebra of sets in  $\mathbb{Q}_p$ . We refer to the atoms in that algebra as intervals.
- Let's introduce some notation for the intervals. For  $t \in T$  and  $\alpha_L, \alpha_U \in \mathbb{Z} \cup \{-\infty, +\infty\}$  define

$$I(t, \alpha_L, \alpha_U) = B(t, \alpha_L) \setminus \bigcup \{B(t', \alpha_U) \mid t' \in T, \text{val}(t' - t) \geq \alpha_U\}$$

(this is sometimes referred to as the swiss cheese construction). One can check that every interval is of the form  $I(t, \alpha_L, \alpha_U)$  for some values of  $t, \alpha_L, \alpha_U$ . Quantities  $\alpha_L, \alpha_U$  are uniquely determined by the interval, while  $t$  might not be.

- Intervals are a natural construction for trees, however we will require a more refined notion to make Lemma 3.12 below work. Define a larger collection of balls

$$\mathcal{B}' = \mathcal{B} \cup \{B(c_i(b), \text{val}(c_j(b) - c_k(b)))\}_{i,j,k \in I, b \in B}.$$

Similar to the previous definition, we define a subinterval to be an atom of the boolean algebra generated by  $\mathcal{B}'$ . Subintervals refine intervals. Moreover, as before, each subinterval can be written as  $I(t, \alpha_L, \alpha_U)$  for some

values of  $t, \alpha_L, \alpha_U$ . As before,  $\alpha_L, \alpha_U$  are uniquely determined by the subinterval, while  $t$  might not be.

Subintervals are fine enough to make Lemma 3.12 below work while coarse enough to be  $O(N)$  small.

**Lemma 3.4.**

- *There are at most  $2|T| = 2N|I| = O(N)$  different intervals.*
- *There are at most  $2|T| + |B| \cdot |I|^3 = O(N)$  different subintervals.*

*Proof.* Each new element in the tree  $T$  adds at most two intervals to the total count, so by induction there can be at most  $2|T|$  many intervals. Each new ball in  $\mathcal{B}' \setminus \mathcal{B}$  adds at most one subinterval to the total count, so by induction there are at most  $|\mathcal{B}' \setminus \mathcal{B}|$  more subintervals than there are intervals.  $\square$

**Definition 3.5.** Suppose  $a \in \mathbb{Q}_p$  lies in an interval  $I(t, \alpha_L, \alpha_U)$ . Define T-valuation of  $a$  to be  $\text{T-val}(a) = \text{val}(a - t)$ .

This a natural notion having the following properties:

**Lemma 3.6.**

- (a) *T-val( $a$ ) is well-defined, independent of choice of  $t$  to represent the interval.*
- (b) *If  $a \in \mathbb{Q}_p$  lies in a subinterval  $I(t, \alpha_L, \alpha_U)$  (as opposed to an interval), then  $\text{T-val}(a) = \text{val}(a - t)$  as well (this works for any refinement of intervals).*
- (c) *If  $a \in \mathbb{Q}_p$  lies in a (sub)interval  $I(t, \alpha_L, \alpha_U)$  then  $\alpha_L < \text{T-val}(a) \leq \alpha_U$ .*
- (d) *For any  $a \in \mathbb{Q}_p$  lying in a (sub)interval  $I(t, \alpha_L, \alpha_U)$  and  $t' \in T$* 
  - *If  $\text{val}(t - t') \geq \alpha_U$ , then  $\text{val}(a - t') = \text{T-val}(a)$ .*
  - *If  $\text{val}(t - t') \leq \alpha_L$ , then  $\text{val}(a - t') = \text{val}(t - t') (\leq \alpha_L < \text{T-val}(a))$ .*

*Proof.* (a)-(c) are clear. For (d) fix  $t' \in T$  and suppose  $a \in \mathbb{Q}_p$  lies in a subinterval  $I(t, \alpha'_L, \alpha'_U)$ . This subinterval lies inside of an interval  $I(t, \alpha_L, \alpha_U)$  for some choice

of  $\alpha_L, \alpha_U$  and by the definition of intervals (or more specifically  $\mathcal{B}$ )

$$\text{val}(t - t') \geq \alpha_U \iff \text{val}(t - t') \geq \alpha'_U$$

$$\text{val}(t - t') \geq \alpha_L \iff \text{val}(t - t') \geq \alpha'_L.$$

Therefore without loss of generality we may assume that  $a \in \mathbb{Q}_p$  lies in an interval  $I(t, \alpha_L, \alpha_U)$ . By (c) and the definition of intervals one of the three following cases has to hold.

Case 1:  $\text{val}(t - t') \geq \alpha_U$  and  $\text{T-val}(a) < \alpha_U$ .

$$\text{val}(t - t') \geq \alpha_U > \text{T-val}(a) = \text{val}(a - t)$$

thus  $\text{val}(a - t') = \text{val}(a - t) = \text{T-val}(a)$  as needed.

Case 2:  $\text{val}(t - t') \geq \alpha_U$  and  $\text{T-val}(a) = \alpha_U$ .

$$\text{T-val}(a) = \text{val}(a - t) = \text{val}(t - t') \geq \alpha_U$$

thus  $\text{val}(a - t') \geq \alpha_U$ . The interval  $I(t, \alpha_L, \alpha_U)$  excludes the ball  $B(t', \alpha_U)$ , so  $a \notin B(t', \alpha_U)$ , that is  $\text{val}(a - t') \leq \alpha_U$ . Combining this with the previous inequality we get that  $\text{val}(a - t') = \alpha_U = \text{T-val}(a)$  as needed.

Case 3:  $\text{val}(t - t') \leq \alpha_L$

$$\text{val}(t - t') \leq \alpha_L < \text{T-val}(a) = \text{val}(a - t)$$

thus  $\text{val}(a - t') = \text{val}(a - t) = \text{T-val}(a)$  as needed. □

**Definition 3.7.** Suppose  $a \in \mathbb{Q}_p$  lies in a subinterval  $I(t, \alpha_L, \alpha_U)$ . We say that  $a$  is far from boundary if

$$\alpha_L + n \leq \text{T-val}(a) \leq \alpha_U - n.$$

Otherwise we say that it is close to boundary.

**Definition 3.8.** Suppose  $a_1, a_2 \in \mathbb{Q}_p$  lie in the same subinterval  $I(t, \alpha_L, \alpha_U)$ . We say  $a_1, a_2$  have the same subinterval type if one of the following holds:

- Both  $a_1, a_2$  are far from boundary and  $a_1 - t, a_2 - t$  are in the same  $Q_{m,n}$  coset.
- Both  $a_1, a_2$  are close to boundary and

$$\text{T-val}(a_1) = \text{T-val}(a_2) \leq \text{val}(a_1 - a_2) - n.$$

**Definition 3.9.** For  $c \in \mathbb{Q}_p$  and  $\alpha, \beta \in \mathbb{Z}$  define  $c \upharpoonright [\alpha, \beta) \in (\mathbb{Z}/p\mathbb{Z})^{\beta-\alpha}$  to be the record of the coefficients of  $c$  for the valuations between  $[\alpha, \beta)$ . More precisely write  $c$  in its power series form

$$c = \sum_{\gamma \in \mathbb{Z}} c_\gamma p^\gamma \text{ with } c_\gamma \in \mathbb{Z}/p\mathbb{Z}$$

Then  $c \upharpoonright [\alpha, \beta)$  is just  $(c_\alpha, c_{\alpha+1}, \dots, c_{\beta-1})$ .

The following lemma is an adaptation of Lemma 7.4 in [1].

**Lemma 3.10.** Fix  $m, n \in \mathbb{N}$ . For any  $x, y, c \in \mathbb{Q}_p$ , if

$$\text{val}(x - c) = \text{val}(y - c) \leq \text{val}(x - y) - n,$$

then  $x - c, y - c$  are in the same coset of  $Q_{m,n}$ .

*Proof.* Call  $a, b \in \mathbb{Q}_p$  similar if  $\text{val } a = \text{val } b$  and

$$a \upharpoonright [\text{val } a, \text{val } a + n) = b \upharpoonright [\text{val } b, \text{val } b + n)$$

If  $a, b$  are similar then

$$a \in Q_{m,n} \iff b \in Q_{m,n}$$

Moreover for any  $\lambda \in \mathbb{Q}_p^\times$ , if  $a, b$  are similar then so are  $\lambda a, \lambda b$ . Thus if  $a, b$  are similar, then they belong to the same coset of  $Q_{m,n}$ . Conditions of the lemma force  $x - c, y - c$  to be similar, thus belonging to the same coset.  $\square$

**Lemma 3.11.** For each subinterval there are at most  $K = K(Q_{m,n})$  many subinterval types (with  $K$  not dependent on  $B$  on the subinterval).

*Proof.* Let  $a, a' \in \mathbb{Q}_p$  lie in the same subinterval  $I(t, \alpha_L, \alpha_U)$ .

Suppose  $a, a'$  are far from boundary. Then they have the same subinterval type if  $a - t, a' - t$  are in the same  $Q_{m,n}$ -coset. Number of such subinterval types is bounded by the number of  $Q_{m,n}$ -cosets.

Suppose  $a, a'$  are close to boundary and

$$\text{T-val}(a) - \alpha_L = \text{T-val}(a') - \alpha_L < n$$

$$a \upharpoonright [\text{T-val}(a), \text{T-val}(a) + n] = a' \upharpoonright [\text{T-val}(a'), \text{T-val}(a') + n]$$

Then  $a, a'$  have the same subinterval type. Such subinterval type is thus determined by  $\text{T-val}(a) - \alpha_L$  and  $a \upharpoonright [\text{T-val}(a), \text{T-val}(a) + n]$ , therefore there are at most  $np^n$  many such types.

A similar argument works for  $a$  with  $\alpha_U - \text{T-val}(a) \leq n$ .

Adding those up we get that there are at most

$$K = (\text{number of } Q_{m,n} \text{ cosets}) + 2np^n$$

many subinterval types. □

The following lemma relates tree notions to  $\Phi$ -types.

**Lemma 3.12.** *Suppose  $d, d' \in \mathbb{Q}_p^{|x|}$  satisfy the following three conditions*

- *For all  $i \in I$   $p_i(d)$  and  $p_i(d')$  are in the same subinterval.*
- *For all  $i \in I$   $p_i(d)$  and  $p_i(d')$  have the same subinterval type.*
- *For all  $i, j \in I$ ,  $\text{T-val}(p_i(d)) > \text{T-val}(p_j(d))$  iff  $\text{T-val}(p_i(d')) > \text{T-val}(p_j(d'))$ .*

*Then  $d, d'$  have the same  $\Phi$ -type over  $B$ .*

*Proof.* There are two kinds of formulas in  $\Phi$  (see Lemma 3.1). First we show that  $d, d'$  agree on formulas of the form  $p_i(x) - c_i(y) \in \lambda_k Q_{m,n}$ . It is enough to show that for every  $i \in I, b \in B$  we have  $p_i(d) - c_i(b), p_i(d') - c_i(b)$  are in the same  $Q_{m,n}$ -coset. Fix such  $i, b$ . For brevity let  $a = p_i(d), a' = p_i(d')$  and  $Q = Q_{m,n}$ . We want to show that  $a - c_i(b), a' - c_i(b)$  are in the same  $Q$ -coset.

Suppose  $a, a'$  are close to boundary. Then  $\text{T-val}(a) = \text{T-val}(a') \leq \text{val}(a - a') - n$ .

Using Lemma 3.6d, we have

$$\text{val}(a - c_i(b)) = \text{val}(a' - c_i(b)) \leq \text{T-val}(a) \leq \text{val}(a - a') - n$$

Lemma 3.10 shows that  $a - c_i(b), a' - c_i(b)$  are in the same  $Q$ -coset.

Now, suppose both  $a, a'$  are far from boundary. Label their interval as  $I(t, \alpha_L, \alpha_U)$ .

Then we have

$$\alpha_L + n \leq \text{val}(a - t) \leq \alpha_U - n$$

$$\alpha_L + n \leq \text{val}(a' - t) \leq \alpha_U - n$$

(as being far from the subinterval's boundary also makes  $a, a'$  far from interval's boundary). We have either  $\text{val}(t - c_i(b)) \geq \alpha_U$  or  $\text{val}(t - c_i(b)) \leq \alpha_L$  (as otherwise it would contradict the definition of intervals, or more specifically  $\mathcal{B}$ ).

Suppose it is the first case  $\text{val}(t - c_i(b)) \geq \alpha_U$ . Then using Lemma 3.6d

$$\text{val}(a - c_i(b)) = \text{val}(a - t) \leq \alpha_U - n \leq \text{val}(t - c_i(b)) - n.$$

So by Lemma 3.10 we have  $a - c_i(b), a - t$  are in the same  $Q$ -coset. By a parallel argument we have  $a' - c_i(b), a' - t$  are in the same  $Q$ -coset. As  $a, a'$  have the same subinterval type,  $a - t, a' - t$  are in the same  $Q$ -coset. Thus by transitivity we get that  $a - c_i(b), a' - c_i(b)$  are in the same  $Q$ -coset.

For the second case, suppose  $\text{val}(t - c_i(b)) \leq \alpha_L$ . Then using Lemma 3.6d

$$\text{val}(a - c_i(b)) = \text{val}(t - c_i(b)) \leq \alpha_L \leq \text{val}(a - t) - n$$

so by Lemma 3.10 we have  $a - c_i(b), t - c_i(b)$  are in the same  $Q$ -coset. By a parallel argument we have  $a' - c_i(b), t - c_i(b)$  are in the same  $Q$ -coset. Thus by transitivity we get that  $a - c_i(b), a' - c_i(b)$  are in the same  $Q$ -coset.

Next, we need to show that  $d, d'$  agree on formulas of the form  $\text{val}(p_i(x) - c_i(y)) < \text{val}(p_j(x) - c_j(y))$  (again, referring to the presentation in Lemma 3.1). Fix  $i, j \in I, b \in B$ . We would like to show the following equivalence:

(3.1)

$$\text{val}(p_i(d) - c_i(b)) < \text{val}(p_j(d) - c_j(b)) \iff \text{val}(p_i(d') - c_i(b)) < \text{val}(p_j(d') - c_j(b))$$

Suppose  $p_i(d), p_i(d')$  are in the subinterval  $I(t_i, \alpha_i, \beta_i)$  and  $p_j(d), p_j(d')$  are in the subinterval  $I(t_j, \alpha_j, \beta_j)$ . Lemma 3.6d yields 4 following cases.

Case 1:

$$\begin{aligned} \text{val}(p_i(d) - c_i(b)) &= \text{val}(p_i(d') - c_i(b)) = \text{val}(t_i - c_i(b)) \\ \text{val}(p_j(d) - c_j(b)) &= \text{val}(p_j(d') - c_j(b)) = \text{val}(t_j - c_j(b)) \end{aligned}$$

Then it is clear that the equivalence (3.1) holds.

Case 2:

$$\begin{aligned} \text{val}(p_i(d) - c_i(b)) &= \text{T-val}(p_i(d)) \text{ and } \text{val}(p_i(d') - c_i(b)) = \text{T-val}(p_i(d')) \\ \text{val}(p_j(d) - c_j(b)) &= \text{T-val}(p_j(d)) \text{ and } \text{val}(p_j(d') - c_j(b)) = \text{T-val}(p_j(d')) \end{aligned}$$

Then the equivalence (3.1) holds by the third condition of the lemma that order of T-valuations is preserved.

Case 3:

$$\begin{aligned} \text{val}(p_i(d) - c_i(b)) &= \text{val}(p_i(d') - c_i(b)) = \text{val}(t_i - c_i(b)) \\ \text{val}(p_j(d) - c_j(b)) &= \text{T-val}(p_j(d)) \text{ and } \text{val}(p_j(d') - c_j(b)) = \text{T-val}(p_j(d')) \end{aligned}$$



If  $p_j(d), p_j(d')$  are close to boundary, then  $\text{T-val}(p_j(d)) = \text{T-val}(p_j(d'))$  and the equivalence (3.1) clearly holds. Suppose then that  $p_j(d), p_j(d')$  are far from boundary.

$$\alpha_j + n \leq \text{T-val}(p_j(d)), \text{T-val}(p_j(d')) \leq \beta_j - n$$

$$\alpha_j < \text{T-val}(p_j(d)), \text{T-val}(p_j(d')) < \beta_j$$

and  $\text{val}(t_i - c_i(b))$  lies outside of the  $(\alpha_j, \beta_j)$  by the definition of subinterval (more specifically definition of  $\mathcal{B}'$ ). Therefore (3.1) has to hold. (Note that we always have  $\text{T-val}(p_j(d)), \text{T-val}(p_j(d')) \in (\alpha_j, \beta_j]$  by Lemma 3.6c, so we only need the far from boundary condition to avoid the edge case of equality to  $\beta_j$ .)

Case 4:

$$\text{val}(p_i(d) - c_i(b)) = \text{T-val}(p_i(d)) \text{ and } \text{val}(p_i(d') - c_i(b)) = \text{T-val}(p_i(d'))$$

$$\text{val}(p_j(d) - c_j(b)) = \text{val}(p_j(d') - c_j(b)) = \text{val}(t_j - c_j(b))$$

Similar to case 3 (switching  $i, j$ ). □

**Note 3.13.** This gives us an upper bound on the number of types - there are at most  $|2I|!$  many choices for the order of T-val,  $O(N)$  many choices for the subinterval for each  $p_i$ , and  $K$  many choices for the subinterval type for each  $p_i$ , giving a total of  $O(N^{|I|}) \cdot K^{|I|} \cdot |I|! = O(N^{|I|})$  many types. This implies  $\text{vc}^*(\Phi) \leq |I|$ . The biggest contribution to this bound are the choices among the  $O(N)$  many subintervals for each  $p_i$  with  $i \in I$ . Are all of those choices realized? Intuitively there are  $|x|$  many variables and  $|I|$  many equations, so once we choose an subinterval for  $|x|$  many  $p_i$ 's, the subinterval for the rest should be determined. This would give the required  $\text{vc}^*(\Phi) \leq |x|$  bound. The next section outlines this idea formally.

#### 4. MAIN PROOF

Alternative way to write  $p_i(c)$  is  $\vec{p}_i \cdot \vec{c}$ , where  $\vec{p}_i$  and  $\vec{c}$  are vectors in  $\mathbb{Q}_p^{|x|}$  (as  $p_i(x)$  is linear).

**Lemma 4.1.** *Suppose we have a finite collection of vectors  $\{\vec{p}_i\}_{i \in I}$  with each  $\vec{p}_i \in \mathbb{Q}_p^{|x|}$ . Suppose  $J \subset I$  and  $i \in I$  satisfy*

$$\vec{p}_i \in \text{span} \{\vec{p}_j\}_{j \in J},$$

*and we have  $\vec{c} \in \mathbb{Q}_p^{|x|}$ ,  $\alpha \in \mathbb{Z}$  with*

$$\text{val}(\vec{p}_j \cdot \vec{c}) > \alpha \text{ for all } j \in J$$

*Then*

$$\text{val}(\vec{p}_i \cdot \vec{c}) > \alpha - \gamma$$

*for some  $\gamma \in \mathbb{N}$ . Moreover  $\gamma$  can be chosen independently from  $J, j, \vec{c}, \alpha$  depending only on  $\{\vec{p}_i\}_{i \in I}$ .*

*Proof.* Fix  $i, J$  satisfying the conditions of the lemma. For some  $c_j \in \mathbb{Q}_p$  for  $j \in J$  we have

$$\vec{p}_i = \sum_{j \in J} c_j \vec{p}_j,$$

hence

$$\vec{p}_i \cdot \vec{c} = \sum_{j \in J} c_j \vec{p}_j \cdot \vec{c}.$$

We have

$$\text{val}(c_j \vec{p}_j \cdot \vec{c}) = \text{val}(c_j) + \text{val}(\vec{p}_j \cdot \vec{c}) > \text{val}(c_j) + \alpha.$$

Let  $\gamma = \max(0, -\max_{j \in J} \text{val}(c_j))$ . Then we have

$$\begin{aligned} \text{val}(c_j \vec{p}_j \cdot \vec{c}) &> \alpha - \gamma \text{ for all } j \in J \\ \text{val}\left(\sum_{j \in J} c_j \vec{p}_j \cdot \vec{c}\right) &> \alpha - \gamma \\ \text{val}(\vec{p}_i \cdot \vec{c}) &> \alpha - \gamma \end{aligned}$$

This shows that we can pick such  $\gamma$  for a given choice of  $i, J$ , but independent from  $\alpha, \vec{c}$ . To get a choice independent from  $i, J$ , go over all such eligible choices ( $i$  ranges over  $I$  and  $J$  ranges over subsets of  $I$ ), pick  $\gamma$  for each, and then take the maximum of those values.  $\square$

Fix  $\gamma$  according to Lemma 4.1 corresponding to  $\{\vec{p}_i\}_{i \in I}$  given by our collection of formulas  $\Phi$ . (The lemma above is a general result, but we only use it applied to the vectors given by  $\Phi$ .)

**Definition 4.2.** Suppose  $a \in \mathbb{Q}_p$  lies in a subinterval  $(B(t_L, \alpha_L), B(t_U, \alpha_U))$ . Define floor of  $a$  to be  $F(a) = \alpha_L$ .

**Definition 4.3.** Let  $f : \mathbb{Q}_p^{|x|} \rightarrow \mathbb{Q}_p^I$  with  $f(c) = (p_i(c))_{i \in I}$ . Define the segment space  $\text{Sg}$  to be the image of  $f$ .

Given a tuple  $(a_i)_{i \in I}$  in the segment space, look at the corresponding floors  $\{F(a_i)\}_{i \in I}$  and T-valuations  $\{\text{T-val}(a_i)\}_{i \in I}$ . Partition the segment space by the order types of  $\{F(a_i)\}_{i \in I}$  and  $\{\text{T-val}(a_i)\}_{i \in I}$  (as subsets of  $\mathbb{Z}$ ).

Work in a fixed partition  $\text{Sg}'$ . After relabeling we may assume that

$$F(a_1) \geq F(a_2) \geq \dots$$

Consider the (relabelled) sequence of vectors  $\vec{p}_1, \vec{p}_2, \dots, \vec{p}_I$ . There is a unique subset  $J \subset I$  such that all vectors with indices in  $J$  are linearly independent, and all vectors with indices outside of  $J$  are a linear combination of preceding vectors.

For any index  $i \in I$  we call it independent if  $i \in J$  and we call it dependent otherwise.

**Definition 4.4.**

- Denote  $\mathbb{Z}/p\mathbb{Z}^\gamma$  as  $\underline{\text{Ct}}$ . Note that  $|\text{Ct}| = p^\gamma$ .
- Let  $\underline{\text{Tp}}$  be the space of all subinterval types. By Lemma 3.11  $|\text{Tp}| \leq K$ .
- Let  $\underline{\text{Sub}}$  be the space of all subintervals. By Lemma 3.4  $|\text{Sub}| \leq 3|I|^2 \cdot N = O(N)$ .

**Definition 4.5.** Now, we define the following function

$$g_{\text{Sg}'} : \text{Sg}' \longrightarrow \text{Tp}^I \times \text{Sub}^J \times \text{Ct}^{I \setminus J}$$

Let  $a = (a_i)_{i \in I} \in \text{Sg}'$ . To define  $g_{\text{Sg}'}(a)$  we need to specify where it maps  $a$  in each individual component of the product.

For each  $a_i$  record its subinterval type, giving the first component  $\text{Tp}^I$ .

For  $a_j$  with  $j \in J$ , record the subinterval of  $a_j$ , giving the second component  $\text{Sub}^J$ .

For the third component  $\text{Ct}^{I \setminus J}$  do the following computation. Pick  $a_i$  with  $i$  dependent. Let  $j$  be the largest independent index with  $j < i$ . Record  $a_i \upharpoonright [F(a_j) - \gamma, F(a_j))$ .

Combine  $g_{\text{Sg}'}$  for all the partitions to get a function

$$g : \text{Sg} \longrightarrow \text{Tp}^I \times \text{Sub}^J \times \text{Ct}^{I \setminus J}.$$

**Lemma 4.6.** Suppose we have  $c, c' \in \mathbb{Q}_p^{|x|}$  such that  $f(c), f(c')$  are in the same partition and  $g(f(c)) = g(f(c'))$ . Then  $c, c'$  have the same  $\Phi$ -type over  $B$ .

*Proof.* Let  $a_i = \vec{p}_i \cdot \vec{c}$  and  $a'_i = \vec{p}_i \cdot \vec{c}'$  so that

$$f(c) = (p_i(c))_{i \in I} = (\vec{p}_i \cdot \vec{c})_{i \in I} = (a_i)_{i \in I}$$

$$f(c') = (p_i(c'))_{i \in I} = (\vec{p}_i \cdot \vec{c}')_{i \in I} = (a'_i)_{i \in I}$$

For each  $i$  we show that  $a_i, a'_i$  are in the same subinterval and have the same subinterval type, so the conclusion follows by Lemma 3.12 ( $f(c), f(c')$  are in the same partition ensuring the proper order of T-valuations for the 3rd condition of the lemma). Tp records the subinterval type of each element, so if  $g(\bar{a}) = g(\bar{a}')$  then  $a_i, a'_i$  have the same subinterval type for all  $i \in I$ . Thus it remains to show that  $a_i, a'_i$  lie in the same subinterval for all  $i \in I$ . Suppose  $i$  is an independent index. Then by construction, Sub records the subinterval for  $a_i, a'_i$ , so those have to belong to the same subinterval. Now suppose  $i$  is dependent. Pick the largest  $j < i$  such that  $j$  is independent. We have  $F(a_i) \leq F(a_j)$  and  $F(a'_i) \leq F(a'_j)$ . Moreover  $F(a_j) = F(a'_j)$  as  $a_j, a'_j$  lie in the same subinterval (using the earlier part of the argument as  $j$  is independent).

**Claim 4.7.**  $\text{val}(a_i - a'_i) > F(a_j) - \gamma$

*Proof.* Let  $K$  be the set of the independent indices less than  $i$ . Note that by the definition for dependent indices we have  $\vec{p}_i \in \text{span}\{\vec{p}_k\}_{k \in K}$ . We also have

$$\text{val}(a_k - a'_k) > F(a_k) \text{ for all } k \in K$$

as  $a_k, a'_k$  lie in the same subinterval (using the earlier part of the argument as  $k$  is independent).

$$\text{val}(a_k - a'_k) > F(a_j) \text{ for all } k \in K \text{ by monotonicity of } F(a_k)$$

$$\text{val}(\vec{p}_k \cdot \vec{c} - \vec{p}_k \cdot \vec{c}') > F(a_j) \text{ for all } k \in K$$

$$\text{val}(\vec{p}_k \cdot (\vec{c} - \vec{c}')) > F(a_j) \text{ for all } k \in K$$

$K \subset I, i \in I, \vec{c} - \vec{c}' \in \mathbb{Q}_p^{|x|}, F(a_j) \in \mathbb{Z}$  satisfy the requirements of Lemma 4.1, so we apply it to conclude

$$\text{val}(\vec{p}_i \cdot (\vec{c} - \vec{c}')) > F(a_j) - \gamma$$

$$\text{val}(\vec{p}_i \cdot \vec{c} - \vec{p}_i \cdot \vec{c}') > F(a_j) - \gamma$$

$$\text{val}(a_i - a'_i) > F(a_j) - \gamma$$

as needed, finishing the proof of the claim.  $\square$

Additionally  $a_i, a'_i$  have the same image in Ct component, so we have

$$\text{val}(a_i - a'_i) > F(a_j)$$

We now would like to show that  $a_i, a'_i$  lie in the same subinterval. As  $F(a_i) \leq F(a_j)$ ,  $F(a'_i) \leq F(a'_j)$  and  $F(a_j) = F(a'_j)$  we have that  $\text{val}(a_i - a'_i) > F(a_i)$  and  $\text{val}(a_i - a'_i) > F(a'_i)$ . Suppose that  $a_i$  lies in the subinterval  $I(t, F(a_i), \alpha_U)$  and that  $a'_i$  lies in the subinterval  $I(t', F(a'_i), \alpha'_U)$ . Without loss of generality assume that  $F(a_i) \leq F(a'_i)$ . As  $\text{val}(a_i - a'_i) > F(a'_i)$ , this implies that

$$a_i \in B(a'_i, F(a'_i))$$

$$a_i \in B(t', F(a'_i))$$

$$B(t, F(a_i)) \cap B(t', F(a'_i)) \neq \emptyset$$

$$B(t, F(a_i)) \subset B(t', F(a'_i))$$

For the subintervals to be disjoint we need  $I(t, F(a_i), \alpha_U) \cap B(t', F(a'_i)) = \emptyset$ . But  $\text{val}(t' - a_i) > F(a'_i)$  implying that  $a_i \in I(t, F(a_i), \alpha_U) \cap B(t', F(a'_i))$  giving a contradiction. Therefore the subintervals coincide finishing the proof.  $\square$

**Corollary 4.8.**  $\Phi(x, y)$  has dual vc-density  $\leq |x|$ .

*Proof.* Suppose we have  $c, c' \in \mathbb{Q}_p^{|x|}$  such that  $f(c), f(c')$  are in the same partition and  $g(f(c)) = g(f(c'))$ . Then by the previous lemma  $c, c'$  have the same  $\Phi$ -type.

Thus the number of possible  $\Phi$ -types is bounded by the size of the range of  $g$  times the number of possible partitions

$$(\text{number of partitions}) \cdot |\text{Tp}|^{|I|} \cdot |\text{Sub}|^{|J|} \cdot |\text{Ct}|^{|I-J|}$$

There are at most  $(|2I|!)^2$  many partitions of  $\text{Sg}$ , so in the product above, the only component dependent on  $B$  is

$$|\text{Sub}|^{|J|} \leq (N \cdot 3|I|^2)^{|J|} = O(N^{|J|})$$

Every  $p_i$  is an element of a  $|x|$ -dimensional vector space, so there can be at most  $|x|$  many independent vectors. Thus we have  $|J| \leq |x|$  and the bound follows.  $\square$

**Corollary 4.9** (Theorem 2.6).  $(\mathbb{Q}_p, \mathcal{L}_{aff})$  has  $\text{vc}(n) = n$ .

*Proof.* Previous lemma implies that  $\text{vc}^*(\phi) \leq \text{vc}^*(\Phi) \leq |x|$ . As choice of  $\phi$  was arbitrary, this implies that  $\text{vc}$ -density of any formula is bounded by the arity of  $x$ .  $\square$

This proof relies heavily on the linearity of functions  $a_1, a_2, c$  in the cell decomposition result (see Definition 2.4). Linearity is used to separate  $x$  and  $y$  variables as well as for Lemma 4.1 to reduce the number of independent factors from  $|I|$  to  $|x|$ . The paper [2] has cell decomposition results for more expressive reducts of  $\mathbb{Q}_p$ , including, for example, restricted multiplication. While our results don't apply to it directly, it is this author's hope that similar techniques can be used to compute  $\text{vc}(n)$  function for those 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] E. Leenknegt. *Reducts of  $p$ -adically closed fields*, Archive for Mathematical logic, 53(3):285-306, 2014

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