

VAPNIK-CHERVONENKIS DENSITY IN SOME THEORIES WITHOUT THE INDEPENDENCE PROPERTY, I

MATTHIAS ASCHENBRENNER, ALF DOLICH, DEIRDRE HASKELL,
DUGALD MACPHERSON, AND SERGEI STARCHENKO

For Lou van den Dries, on his 60th birthday.

ABSTRACT. We recast the problem of calculating Vapnik-Chervonenkis (VC) density into one of counting types, and thereby calculate bounds (often optimal) on the VC density for some weakly o-minimal, weakly quasi-o-minimal, and P -minimal theories.

CONTENTS

1. Introduction	1
2. VC Density	7
3. The Model-Theoretic Context	15
4. Some VC Density Calculations	27
5. Theories with the VC d Property	35
6. Examples of VC d : Weakly O-minimal Theories and Variants	42
7. A Strengthening of VC d , and P -adic Examples	47
References	56

1. INTRODUCTION

The notion of VC dimension, which arose in probability theory in the work of Vapnik and Chervonenkis [98], was first drawn to the attention of model-theorists by Laskowski [55], who observed that a complete first-order theory does not have the independence property (as introduced by Shelah [86]) if and only if, in each model, each definable family of sets has finite VC dimension. With this observation, Laskowski easily gave several examples of classes of sets with finite VC dimension, by noting well-known examples of theories without the independence property. This line of thought was pursued by Karpinski and Macintyre [49], who calculated explicit bounds on the VC dimension of definable families of sets in some o-minimal structures (with an eye towards applications to neural networks), which were polynomial in the number of parameter variables. In a further paper [50], they observe that their arguments also lead to a linear bound on the VC *density* of definable families of sets in some o-minimal structures. They ask whether similar (linear) bounds hold for the p -adic numbers (whose theory also does not have the independence property). The bound in the o-minimal case in [50] was established independently, using a more combinatorial approach, by Wilkie (unpublished), and more recently, also by Johnson and Laskowski [47].

Date: September 2011.

In this paper we give a sufficient criterion (Theorem 5.7) on a first-order theory for the VC density of a definable family of sets to be bounded by a linear function in the number of parameter variables, and show that the criterion is satisfied by several theories of general interest, including the theory of the p -adics and all weakly o-minimal theories. In a sequel to this paper [6] we give different arguments to get similar bounds in a variety of other examples where our criterion does not apply. Before we state our main results, we introduce our setup and review some definitions and basic facts. We hope that the present paper (unlike its sequel [6]) can be read with only little technical knowledge of model theory beyond basic first order logic. The first few chapters of [42] or [63] or similar texts should provide sufficient background for a prospective reader.

1.1. VC dimension and VC density. Let X be an infinite set and \mathcal{S} be a non-empty collection of subsets of X . Given $A \subseteq X$, we say that a subset B of A is *cut out by \mathcal{S}* if $B = S \cap A$ for some $S \in \mathcal{S}$; we let $\mathcal{S} \cap A := \{S \cap A : S \in \mathcal{S}\}$ be the collection of subsets of A cut out by \mathcal{S} . We say that A is *shattered by \mathcal{S}* if every subset of A is cut out by some element of \mathcal{S} . The collection \mathcal{S} is said to be a *VC class* if there is a non-negative integer n such that no subset of X of size n can be shattered by \mathcal{S} . In this case, the *VC dimension of \mathcal{S}* is the largest $d \geq 0$ such that some set of size d is shattered by \mathcal{S} . We denote by $\pi_{\mathcal{S}}(n)$ the maximum, as A varies over subsets of X of size n , of the numbers of subsets of A that can be cut out by \mathcal{S} ; that is,

$$\pi_{\mathcal{S}}(n) := \max \left\{ |\mathcal{S} \cap A| : A \in \binom{X}{n} \right\}.$$

(Here and below, $\binom{X}{n}$ denotes the set of n -element subsets of X .) The function $n \mapsto \pi_{\mathcal{S}}(n)$ is called the *shatter function* of \mathcal{S} . Clearly $0 \leq \pi_{\mathcal{S}}(n) \leq 2^n$ for every n , and if \mathcal{S} is not a VC class, then $\pi_{\mathcal{S}}(n) = 2^n$ for every n . However, if \mathcal{S} is a VC class, of VC dimension d say, then by a fundamental observation of Sauer [83] (independently made in [87] and, implicitly, in [98]), the function $n \mapsto \pi_{\mathcal{S}}(n)$ is bounded above by a polynomial in n of degree d . (In fact, for $d, n \geq 1$ one has $\pi_{\mathcal{S}}(n) \leq (en/d)^d$, where e is the base of the natural logarithm.) Hence it makes sense to define the *VC density* of a VC class \mathcal{S} as the infimum of all reals $r \geq 0$ such that $\pi_{\mathcal{S}}(n)/n^r$ is bounded for all positive n . It turns out that in many cases, the VC density (rather than the VC dimension) is the decisive measure for the combinatorial complexity of a family of sets. For example, the VC density of \mathcal{S} governs the size of packings in \mathcal{S} with respect to the Hamming metric ([41], see also [64, Lemma 2.1]), and is intimately related to the notions of entropic dimension [7] and discrepancy [68]. We refer to the surveys [65, 33] for uses of VC density in combinatorics.

1.2. VC dimension and VC density of formulas. Let \mathcal{L} be a first-order language. In an \mathcal{L} -structure M , a natural way to generate a collection of subsets of M^m is to take the family of sets defined by a formula, as the parameters vary. Given a tuple $x = (x_1, \dots, x_m)$ of pairwise distinct variables we denote by $|x| := m$ the *length* of x . We often need to deal with \mathcal{L} -formulas whose free variables have been separated into *object* and *parameter* variables. We use the notation $\varphi(x; y)$ to indicate that the free variables of the \mathcal{L} -formula φ are contained among the components of the tuples $x = (x_1, \dots, x_m)$ and $y = (y_1, \dots, y_n)$ of pairwise distinct variables (which we also assume to be disjoint). Here the x_i are thought of as the object variables and the y_j as the parameter variables. We refer to $\varphi(x; y)$ as a *partitioned \mathcal{L} -formula*.

In the rest of this introduction we let \mathbf{M} be an infinite \mathcal{L} -structure. Let $\varphi(x; y)$ be a partitioned \mathcal{L} -formula, $m = |x|$, $n = |y|$, and denote by

$$\mathcal{S}_\varphi = \{\varphi^{\mathbf{M}}(M^m; b) : b \in M^n\}$$

the family of subsets of M^m defined by φ in \mathbf{M} using parameters ranging over M^n . We call \mathcal{S}_φ a definable family of sets (in \mathbf{M}). We say that φ defines a VC class in \mathbf{M} if \mathcal{S}_φ is a VC class; in this case the VC dimension of φ in \mathbf{M} is the VC dimension of the collection \mathcal{S}_φ of subsets of M^m , and similarly one defines the VC density of φ in \mathbf{M} . Since the shatter function $\pi_\varphi = \pi_{\mathcal{S}_\varphi}$ of \mathcal{S}_φ only depends on the elementary theory of \mathbf{M} (see Lemma 3.2 below), given a complete \mathcal{L} -theory T with no finite models, we may also speak of the shatter function of φ in T as well as VC dimension of φ in T and the VC density of φ in T .

1.3. NIP theories. A partitioned \mathcal{L} -formula $\varphi(x; y)$ as above is said to have the independence property for \mathbf{M} if for every $t \in \mathbb{N}$ there are $b_1, \dots, b_t \in M^n$ such that for every $S \subseteq \{1, \dots, t\}$ there is $a_S \in M^m$ such that for all $i \in \{1, \dots, t\}$, $\mathbf{M} \models \varphi(a_S; b_i) \iff i \in S$. The structure \mathbf{M} is said to have the independence property if some \mathcal{L} -formula has the independence property for \mathbf{M} , and not to have the independence property (or to be NIP or dependent) otherwise. By a classical result of Shelah [86] (with other proofs in [52, 55, 80]), for \mathbf{M} to be NIP it is actually sufficient that no formula $\varphi(x; y)$ with $|x| = 1$ has the independence property for \mathbf{M} . NIP is implied by (but not equivalent to) another prominent tameness condition on first-order structures called stability: An \mathcal{L} -formula $\varphi(x; y)$ is said to be unstable for \mathbf{M} if for every $t \in \mathbb{N}$ there are $a_1, \dots, a_t \in M^m$ and $b_1, \dots, b_t \in M^n$ such that $\mathbf{M} \models \varphi(a_i; b_j) \iff i \leq j$, for all $i, j \in \{1, \dots, t\}$. The \mathcal{L} -structure \mathbf{M} is called unstable if some \mathcal{L} -formula φ is unstable for \mathbf{M} ; and “stable” (for formulas and structures) is synonymous with “not unstable.”

Laskowski’s observation [55] is that an \mathcal{L} -formula defines a VC class in \mathbf{M} if and only if it does not have the independence property for \mathbf{M} . In fact, given a collection \mathcal{S} of subsets of a set X , define the dual shatter function of \mathcal{S} as the function $n \mapsto \pi_{\mathcal{S}}^*(n)$ whose value at n is the maximum number of equivalence classes defined by an n -element subfamily \mathcal{T} of \mathcal{S} , where two elements of X are said to be equivalent with respect to \mathcal{T} if they belong to the same sets of \mathcal{T} . Then a given partitioned \mathcal{L} -formula $\varphi(x; y)$ has the independence property precisely if $\pi_{\mathcal{S}_\varphi}^*(n) = 2^n$ for every n . The dual shatter function of \mathcal{S}_φ is really a shatter function in disguise: it agrees with the shatter function of \mathcal{S}_{φ^*} where $\varphi^*(y; x) := \varphi(x; y)$ is the dual of the partitioned formula φ . (See Section 3.)

A complete \mathcal{L} -theory T is said to have the independence property if some model of it does, and is said not to have the independence property (or to be NIP) otherwise. Thus a complete \mathcal{L} -theory T is NIP if and only if every \mathcal{L} -formula defines a VC class in every model of T . Many theories arising in mathematical practice turn out to be NIP: By [86], all stable theories (i.e., complete theories all of whose models are stable) are NIP; so, for example, algebraically closed (more generally, separably closed) fields, differentially closed fields, modules, or free groups furnish examples of NIP structures. Furthermore, o-minimal (or more generally, weakly o-minimal) theories are NIP [55, 61]. By [36] any ordered abelian group has NIP theory. Certain important theories of henselian valued fields are NIP, for example, the completions of the theory of algebraically closed valued fields and the theory of the field of p -adic numbers (and also their rigid analytic and p -adic subanalytic expansions, respectively). In fact, in the language of rings with a predicate for the valuation ring, an unramified henselian valued field of characteristic $(0, p)$ is NIP if and only if its residue field is NIP [12]. Similarly, henselian valued fields

of characteristic $(0, 0)$ and algebraically maximal Kaplansky fields of characteristic (p, p) are NIP iff their residue fields are NIP [13, 12].

On the other hand, each pseudofinite field (infinite model of the theory of all finite fields) is not NIP [29], since it defines the (Rado) random graph.

1.4. Uniform bounds on VC density. This paper is motivated by the following question: *Given a NIP theory T , can one find an upper bound, in terms of n only, on the VC densities (in T) of all \mathcal{L} -formulas $\varphi(x; y)$ with $|y| = n$?* The intuition behind this question is, of course, that the complexity of a family \mathcal{S}_φ of sets defined by a first-order formula $\varphi(x; y)$ in a NIP structure should be governed by the number n of freely choosable parameters. Note that the minimum possible bound is $|y| = n$: for if $\varphi(x; y)$, where x is a single variable, is the formula $x = y_1 \vee \cdots \vee x = y_n$, then the subsets of M cut out by \mathcal{S}_φ are exactly the non-empty subsets of M of cardinality at most n , so $\varphi(x; y)$ has VC density n (in any complete theory). We note here in passing that the VC density of a formula φ in a NIP theory may take fractional values, and that the shatter function of \mathcal{S}_φ , though not growing faster than polynomially, is not *asymptotic* to a real power function in general. See Section 4 below, where we explicitly compute the VC density of certain incidence structures (related to the Szémeredi-Trotter Theorem) and of the edge relation in Spencer-Shelah random graphs, and investigate the asymptotics of a shatter function in the infinitary hypercube.

In this paper we employ VC duality to translate the problem of bounding the VC density of a formula φ into the task of counting φ^* -types over finite parameter sets, which then can be treated by model-theoretic machinery. Viewing VC density as a bound on a number of types also illuminates the connection with a strengthening of the NIP concept, which is that of *dp-minimality*. (See Section 5.3 below for a definition.) Dolich, Goodrick and Lippel [24] have observed that, if, in a theory, the dual VC density of any \mathcal{L} -formula in a single object variable is less than 2, then the theory in question is dp-minimal. (No counterexample to the converse of this implication seems to be known.)

We now state our main results. First, an optimal bound on density is obtained for weakly o-minimal theories (see Theorem 6.1 below). Recall that a complete theory T in a language containing a binary relation symbol “ $<$ ” which expands the theory of linearly ordered sets is called *weakly o-minimal* if in every model of T , each partitioned \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$ defines a finite union of convex sets. (See [61] for more on this notion, which generalizes the probably more familiar concept of an o-minimal theory, cf. [25].)

Theorem 1.1. *Suppose \mathcal{L} contains a binary relation symbol “ $<$ ”, interpreted in \mathbf{M} as a linear ordering. If $T = \text{Th}(\mathbf{M})$ is weakly o-minimal, then every \mathcal{L} -formula $\varphi(x; y)$ has VC density at most $n = |y|$ in T (in fact, $\pi_\varphi(t) = O(t^n)$).*

This bound is the same as that obtained by Karpinski-Macintyre [49] for o-minimal expansions of the reals, or by Wilkie and by Johnson-Laskowski [47] for all o-minimal structures. The motivating example of a theory which is weakly o-minimal but not o-minimal is the theory of real closed valued fields, that is, real closed fields equipped with a predicate for a proper convex valuation ring. In fact, the methods of Karpinski and Macintyre can also be adapted to give the correct density bounds for this and certain other weakly o-minimal expansions of real closed fields [40]. Some interesting weakly o-minimal theories to which these methods do not readily adapt may be found in [5, 54].

Our approach to Theorem 1.1, via definable types, was partly inspired by the use of Puiseux series in [11, 81].

Let ACVF denote the theory of (non-trivially) valued algebraically closed fields, in the ring language expanded by a predicate for the valuation divisibility. This has completions $\text{ACVF}_{(0,0)}$ (for residue characteristic 0), $\text{ACVF}_{(0,p)}$ (field characteristic 0, residue characteristic p), and $\text{ACVF}_{(p,p)}$ (field characteristic p). Because $\text{ACVF}_{(0,0)}$ is interpretable in RCVF, our methods give (non-optimal) density bounds for $\text{ACVF}_{(0,0)}$ (Corollary 6.3). However, they give no information on density in the theories $\text{ACVF}_{(0,p)}$ and $\text{ACVF}_{(p,p)}$. The problems arise essentially because a definable set in 1-space in ACVF is a finite union of ‘Swiss cheeses’ but we have no way of choosing a particular Swiss cheese. This means that the definable types technique in our main tool (Theorem 5.7) breaks down. On the other hand, our methods do yield:

Theorem 1.2. *Suppose $M = \mathbb{Q}_p$ is the field of p -adic numbers, construed as a first-order structure in Macintyre’s language \mathcal{L}_p . Then the VC density of every \mathcal{L}_p -formula $\varphi(x; y)$ is at most $2|y| - 1$.*

The same result holds for the subanalytic expansions of \mathbb{Q}_p considered by Denef and van den Dries [22]. (Theorem 7.2 and Remark 7.9.) Key tools available here, but not in the case of ACVF, are cell decomposition and the existence of definable Skolem functions. We do not know whether the bound in Theorem 1.2 is optimal.

The investigation of the fine structure of type spaces over finite parameter sets in NIP theories is only just beginning, and the present paper can be seen as a first step in studying one particular measure (VC density) for their complexity. Applications of the results in this paper to transversals of definable families in NIP theories will appear in a separate manuscript, under preparation by the first- and last-named authors.

As remarked above, all stable theories are NIP, so it also makes sense to investigate VC density in stable theories. In a sequel of the present paper [6] we obtain bounds on VC density in certain finite U-rank theories (including all complete theories of finite Morley rank expansions of infinite groups).

We close off this introduction by pointing out that besides being of intrinsic interest, uniform bounds on VC density of first-order formulas (as obtained in this paper) often also help to explain why certain well-known effective bounds on the complexity of geometric arrangements, used in computational geometry, are polynomial in the number of objects involved. For example, the bound on the number of semialgebraically connected components of realizable sign conditions on polynomials over real closed fields from [11, 81] breaks up into a topological and a combinatorial part, where the polynomial nature of the latter may be seen as a consequence of Theorem 1.1:

Example. Let R be a real closed field, $\mathcal{P} = (P_1, \dots, P_s)$ be a tuple of polynomials from $R[X] = R[X_1, \dots, X_k]$, each of degree at most d . A sign condition for \mathcal{P} is an s -tuple $\sigma \in \{-1, 0, +1\}^s$, and we say that σ is realized in a subset V of R^k if

$$\sigma_V := \{a \in V : (\text{sign } P_1(a), \dots, \text{sign } P_s(a)) = \sigma\}$$

is non-empty. Theorem 1.1 in the semialgebraic case yields: *if V is an algebraic set defined by polynomials of degree at most d , then the number of sign conditions for \mathcal{P} realized in V is at most Cs^m , where $m = \dim(V)$ and the constant $C = C(d, k)$ only depends on d and k .*

To see this recall that by cell decomposition, V is a finite union of semialgebraic subsets of R^k each of which is semialgebraically homeomorphic to some R^n ; moreover, this decomposition (and the resulting homeomorphism) can be chosen uniformly in the parameters: Every zero set of polynomials from $R[X]$ of degree at most d is the zero set of M such polynomials, where $M = \binom{k+d}{d}$ is the dimension of the R -linear subspace of $R[X]$ consisting of the polynomials of degree at most d ; thus we may take a semialgebraic (in fact, algebraic) family $(V_b)_{b \in R^N}$, where $N = M^2$, whose fibers V_b are the algebraic subsets of R^k defined by polynomials of degree at most d . Then there are finitely many semialgebraic families $(V_b^{(i)})_{b \in R^N}$ of subsets of R^k and for each i there is a semialgebraic family $(F_b^{(i)})_{b \in R^N}$ of maps such that for each $b \in R^N$ we have $V_b = \bigcup_i V_b^{(i)}$, and $F_b^{(i)}$ is a homeomorphism $R^{m^{(i)}} \rightarrow V_b^{(i)}$, for some $m^{(i)}$.

Fix some i and write $m = m^{(i)}$. Let $\nu = (\nu_1, \dots, \nu_k)$ range over \mathbb{N}^k , with $|\nu| = \nu_1 + \dots + \nu_k$, and suppose $y = (y_\nu)_{|\nu| \leq d}$, so y has length M . Let $P(X; y)$ be the general polynomial in the indeterminates X of degree at most d with coefficient sequence y ; so every P_j is of the form $P_j = P(X; b_j)$ with $b_j \in R^M$. Suppose also $x = (x_1, \dots, x_m)$, and let z be a tuple of new variables of length N , let z' be a single new variable, and let $\varphi^{(i)}(x; y, z, z')$ be a formula in the language of ordered rings which expresses that $P(F_z^{(i)}(x); y)$ and z' have the same sign. So, e.g., for $a \in R^m$, $b \in R^N$ we have $R \models \varphi^{(i)}(a; b_j, b, 1)$ iff $P_j(F_b^{(i)}(a)) > 0$. In this way we see that the number of sign conditions for \mathcal{P} realized in $V_b^{(i)}$ is bounded by $\pi_{\varphi^{(i)}}^*(3s)$ and thus is $O(s^m)$ by Theorem 1.1, where the implicit constant only depends on $\varphi^{(i)}$ and hence on d and k . This yields the claim highlighted above. (Of course we have been very nonchalant with the constants. Indeed, [11] shows the more precise result that the sum of the number of semialgebraically connected components of the sets σ_V , where σ ranges over all sign conditions for \mathcal{P} realized in V , is bounded by $(O(d))^k \binom{s}{m}$.)

A simpler example is the number of non-empty sets definable by equalities and inequalities of a finite collection of polynomials over an algebraically closed field:

Example. Here we let $\nu = (\nu_1, \dots, \nu_m)$ range over \mathbb{N}^m , and suppose $y = (y_\nu)_{|\nu| \leq d}$. Let $\varphi(x; y)$ be the partitioned formula

$$\sum_{|\nu| \leq d} y_\nu x^\nu = 0$$

in the language \mathcal{L} of rings, and fix an algebraically closed field K . Then $\mathcal{S}_\varphi = \mathcal{S}_\varphi^K$ is the collection of all zero sets (in K^m) of polynomials in m indeterminates with coefficients in K having degree at most d . Hence $\pi_{\mathcal{S}_\varphi}^*(t)$ is the maximum number of non-empty Boolean combinations of t such hypersurfaces. In the sequel of our paper (see [6, Theorem 1.1]) we will show that the shatter function of any partitioned \mathcal{L} -formula with m parameter variables (such as φ^*) is $O(t^m)$ in $\text{Th}(K)$; hence $\pi_\varphi^*(t) = \pi_{\varphi^*}(t) = O(t^m)$. (In fact, [46] proves that $\pi_\varphi^*(t) \leq \sum_{k=0}^m \binom{t}{k} d^k$ for every t , and this bound is asymptotically optimal.)

1.5. Organization of the paper. In the preliminary Section 2 we set the scene by recalling the definitions and basic facts concerning VC dimension and VC density in a general combinatorial setting. In Section 3 we then move to the model-theoretic context; in particular we introduce the VC density function of a complete theory without finite models, and the (dual) VC density of a finite set of formulas. In Section 4 we give some interesting examples of formulas in NIP theories for which we can explicitly compute their VC density or determine the asymptotic behavior of their shatter function.

In Section 5 we introduce the VC d property (a refinement of Guingona’s notion of uniform definability of types over finite sets) and get our main tool for counting types (Theorem 5.7) in place, which is then employed, in Section 6, to prove Theorem 1.1 from above. A strengthening of the VC d property is defined and established for the p -adics in Section 7, thus proving Theorem 1.2. We refer to the introductions of each section for a more detailed description of their contents.

1.6. Notations and conventions. In this paper, d, k, m, n range over the set $\mathbb{N} := \{0, 1, 2, \dots\}$ of natural numbers. We set $[n] := \{1, \dots, n\}$. Given a set X , we write 2^X for the power set of X , and we let $\binom{X}{n}$ denote the set of n -element subsets of X and $\binom{X}{\leq n} := \binom{X}{0} \cup \binom{X}{1} \cup \dots \cup \binom{X}{n}$ the collection of subsets of X of cardinality at most n .

1.7. Acknowledgments. Part of the work on this paper was done while some of the authors were participating in the thematic program on O-minimal Structures and Real Analytic Geometry at the Fields Institute in Toronto (Spring 2009), and in the Durham Symposium on New Directions in the Model Theory of Fields (July 2009), organized by the London Mathematical Society and funded by EPSRC grant EP/F068751/1. The support of these institutions is gratefully acknowledged. Aschenbrenner was partly supported by NSF grant DMS-0556197. He would also express his gratitude to Gerhard Wöginger for suggesting the example in Section 4.4.1, and to Andreas Baudisch and Humboldt-Universität Berlin for their hospitality during Fall 2010. Haskell’s research was supported by NSERC grant 238875. Macpherson acknowledges support by EPSRC grant EP/F009712/1. Starchenko was partly supported by NSF grant DMS-0701364.

2. VC DENSITY

In this section we introduce various numerical parameters associated to abstract families of sets: VC dimension, VC density, and independence dimension, and we recall the well-known phenomenon of “VC duality” hinted at already in the introduction (which, in particular, allows us to relate VC dimension with independence dimension). An important role in later sections is played by a new parameter associated to a set system defined here, which we call *breadth*, and which is the focus of the last part of this section.

2.1. VC dimension and VC density. A *set system* is a pair (X, \mathcal{S}) consisting of a set X and a collection \mathcal{S} of subsets of X . We call X the *base set* of the set system (X, \mathcal{S}) , and we sometimes also speak of a *set system \mathcal{S} on X* . Given a set system (X, \mathcal{S}) and a set $A \subseteq X$, we let $\mathcal{S} \cap A := \{S \cap A : S \in \mathcal{S}\}$ and call $(A, \mathcal{S} \cap A)$ the *set system on A induced by \mathcal{S}* . Let now \mathcal{S} be a set system on an infinite set X . The function $\pi_{\mathcal{S}} : \mathbb{N} \rightarrow \mathbb{N}$ given by

$$\pi_{\mathcal{S}}(n) := \max \left\{ |\mathcal{S} \cap A| : A \in \binom{X}{n} \right\}$$

is called the *shatter function of \mathcal{S}* . We have $0 \leq \pi_{\mathcal{S}}(n) \leq 2^n$ and $\pi_{\mathcal{S}}(n) \leq \pi_{\mathcal{S}}(n+1)$ for all n . Note that if $Y \supseteq X$ then $\pi_{\mathcal{S}}$ does not change if \mathcal{S} is considered as a set system on Y . (This justifies our choice of notation for the shatter function, suppressing the base set X of our set system.)

One says that $A \subseteq X$ is *shattered by \mathcal{S}* if $\mathcal{S} \cap A = 2^A$. If \mathcal{S} is non-empty, then we define the *VC dimension of \mathcal{S}* , denoted by $\text{VC}(\mathcal{S})$, as the supremum (in $\mathbb{N} \cup \{\infty\}$) of the sizes of all finite subsets of X shattered by \mathcal{S} ; so $\text{VC}(\mathcal{S}) = \infty$ means that arbitrarily large finite subsets of X can be shattered by \mathcal{S} . Equivalently,

$$\text{VC}(\mathcal{S}) = \sup \{n : \pi_{\mathcal{S}}(n) = 2^n\}. \quad (2.1)$$

One says that \mathcal{S} is a *VC class* if $\text{VC}(\mathcal{S}) < \infty$. Note that some sources (e.g., [55]) alternatively define the VC dimension of \mathcal{S} to be the minimum n such that no set of size n is shattered by \mathcal{S} (i.e., $\text{VC}(\mathcal{S}) + 1$, with $\text{VC}(\mathcal{S})$ as given by (2.1)).

We have the following fundamental fact about set systems:

Lemma 2.1 (Sauer-Shelah). *If \mathcal{S} has finite VC dimension d (so $\pi_{\mathcal{S}}(n) < 2^n$ for $n > d$), then*

$$\pi_{\mathcal{S}}(n) \leq \binom{n}{\leq d} := \binom{n}{0} + \cdots + \binom{n}{d} \quad \text{for every } n.$$

If $n \geq d$, then $\binom{n}{\leq d}$ is bounded above by $(en/d)^d$ (where e is the base of the natural logarithm). In particular, either $\pi_{\mathcal{S}}(n) = 2^n$ for every n (if \mathcal{S} is not a VC class), or $\pi_{\mathcal{S}}(n) = O(n^d)$. One may now define the *VC density* $\text{vc}(\mathcal{S})$ of \mathcal{S} as the infimum of all real numbers $r > 0$ such that $\pi_{\mathcal{S}}(n) = O(n^r)$, if there is such an r , and $\text{vc}(\mathcal{S}) := \infty$ otherwise. That is,

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

We also define $\text{VC}(\emptyset) := \text{vc}(\emptyset) := -1$. Then $\text{vc}(\mathcal{S}) \leq \text{VC}(\mathcal{S})$ by Lemma 2.1, and $\text{vc}(\mathcal{S}) < \infty$ iff $\text{VC}(\mathcal{S}) < \infty$. The VC density of \mathcal{S} is also known as the real density [7] or the VC exponent [17] of \mathcal{S} . It is related to the combinatorial dimension of \mathcal{S} introduced by Blei [8] and to compression schemes for \mathcal{S} [47].

Example. Suppose $\mathcal{S} = \binom{X}{\leq d}$. Then the inequality in the statement of Lemma 2.1 is an equality, and $\text{VC}(\mathcal{S}) = \text{vc}(\mathcal{S}) = d$.

Example. Suppose $X = \mathbb{R}^d$, and \mathcal{S} is the collection of all closed affine half-spaces in \mathbb{R}^d , i.e., sets of the form $\{x \in \mathbb{R}^d : \langle x, a \rangle \geq \beta\}$ where $a \in \mathbb{R}^d$, $\beta \in \mathbb{R}$, and $\langle \cdot, \cdot \rangle$ denotes the usual inner product on \mathbb{R}^d . Then $\text{VC}(\mathcal{S}) = d + 1$. (The proof of this fact is based on Radon's Theorem on convex sets; see [7, Corollaire 3.5].) Moreover, $\text{vc}(\mathcal{S}) = d$; in fact, $\pi_{\mathcal{S}}(n) = 2 \sum_{i=0}^d (-1)^{d-i} \binom{n}{\leq i}$ for every n ; see [30, Theorem 3.1].

Example. Suppose $X = \mathbb{R}$, $k \geq 1$, and let \mathcal{S} be the collection whose members are the unions of k disjoint (open) intervals in \mathbb{R} . Then $\text{VC}(\mathcal{S}) = \text{vc}(\mathcal{S}) = 2k$, in fact, $\pi_{\mathcal{S}}(n) = \binom{n}{\leq 2k}$ for each n . (See [28, Exercise 11, Chapter 4].)

In all three examples, $\pi_{\mathcal{S}}$ is actually given by a polynomial of degree $d = \text{vc}(\mathcal{S})$. It is worth pointing out that for a VC class \mathcal{S} , in general $\pi_{\mathcal{S}}$ is not even asymptotic to a real power function; see Section 4.4 below.

Clearly, VC and vc are increasing: if $\mathcal{S} \subseteq \mathcal{T} \subseteq 2^X$, then $\pi_{\mathcal{S}} \leq \pi_{\mathcal{T}}$ and so $\text{VC}(\mathcal{S}) \leq \text{VC}(\mathcal{T})$ and $\text{vc}(\mathcal{S}) \leq \text{vc}(\mathcal{T})$. If X' is an infinite subset of X then $\pi_{\mathcal{S} \cap X'} \leq \pi_{\mathcal{S}}$; more generally (see [7, Proposition 2.2]):

Lemma 2.2. *Let X' be an infinite set and $f: X' \rightarrow X$ be a map, and let $f^{-1}(\mathcal{S}) := \{f^{-1}(S) : S \in \mathcal{S}\}$. Then $\pi_{f^{-1}(\mathcal{S})} \leq \pi_{\mathcal{S}}$, with equality if f is surjective. In particular, $\text{VC}(f^{-1}(\mathcal{S})) \leq \text{VC}(\mathcal{S})$ and $\text{vc}(f^{-1}(\mathcal{S})) \leq \text{vc}(\mathcal{S})$, with equality if f is surjective.*

It is easy to verify that $\text{VC}(\mathcal{S}) = 0$ if and only if $|\mathcal{S}| = 1$, and $\text{vc}(\mathcal{S}) = 0$ if \mathcal{S} is finite; in fact, the converse of the latter implication also holds: if $\text{vc}(\mathcal{S}) < 1$, then \mathcal{S} is finite [7, Proposition 2.19] (and hence actually $\text{vc}(\mathcal{S}) = 0$). It is also easy to verify (cf. [7, Proposition 2.4]) that if $\mathcal{S}_1, \mathcal{S}_2$ are subsets of \mathcal{S} with $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$, then $\text{vc}(\mathcal{S}) = \max\{\text{vc}(\mathcal{S}_1), \text{vc}(\mathcal{S}_2)\}$. In particular, $\text{vc}(\mathcal{S})$ does not change if we alter finitely many sets from \mathcal{S} .

2.2. Independence dimension. Let X be a set. Given subsets A_1, \dots, A_n of X , we denote by $S(A_1, \dots, A_n)$ the set of atoms of the Boolean algebra of subsets of X generated by A_1, \dots, A_n (the “non-empty fields in the Venn diagram of A_1, \dots, A_n ”); that is, $S(A_1, \dots, A_n)$ is precisely the set of non-empty subsets of X of the form

$$\bigcap_{i \in I} A_i \cap \bigcap_{i \in [n] \setminus I} X \setminus A_i \quad \text{where } I \subseteq [n] = \{1, \dots, n\}.$$

Note that $S(A_1, \dots, A_n)$ does not depend on the particular order of the A_i , so sometimes we abuse notation and, e.g., write $S(A_i : i = 1, \dots, n)$ instead of $S(A_1, \dots, A_n)$. We have $0 \leq |S(A_1, \dots, A_n)| \leq 2^n$, and we say that the sequence A_1, \dots, A_n is *independent* (in X) if $|S(A_1, \dots, A_n)| = 2^n$, and call A_1, \dots, A_n *dependent* (in X) otherwise.

Suppose now that \mathcal{S} is a collection of subsets of X . We define $\pi_{\mathcal{S}}^* : \mathbb{N} \rightarrow \mathbb{N}$ by

$$\pi_{\mathcal{S}}^*(n) := \max \{|S(A_1, \dots, A_n)| : A_1, \dots, A_n \in \mathcal{S}\}.$$

Note that $0 \leq \pi_{\mathcal{S}}^*(n) \leq 2^n$ for each n . We say that \mathcal{S} is *independent* (in X) if $\pi_{\mathcal{S}}^*(n) = 2^n$ for every n , that is, if for every n there is an independent sequence of elements of \mathcal{S} of length n . Otherwise, we say that \mathcal{S} is *dependent* (in X). If \mathcal{S} is dependent, we define the *independence dimension* $\text{IN}(\mathcal{S})$ of \mathcal{S} as the largest n such that $\pi_{\mathcal{S}}^*(n) = 2^n$, and if \mathcal{S} is independent, we set $\text{IN}(\mathcal{S}) = \infty$. If \mathcal{S} is finite, then clearly $\text{IN}(\mathcal{S}) \leq |\mathcal{S}|$.

Example 2.3. $\text{IN}(\mathcal{S}) \leq 1$ iff for all $S, S' \in \mathcal{S}$ one of the following relations holds: $S \cap S' = \emptyset$, $S \subseteq S'$, $S' \subseteq S$, or $S \cup S' = X$.

The function $\pi_{\mathcal{S}}^*$ is called the *dual shatter function* of \mathcal{S} , since (for infinite \mathcal{S}) one has $\pi_{\mathcal{S}}^* = \pi_{\mathcal{S}^*}$ for a certain set system \mathcal{S}^* on $X^* = \mathcal{S}$, called the *dual* of \mathcal{S} (cf. [7, 2.7–2.11] or [66, Section 10.3]). For the same reason, the independence dimension of \mathcal{S} is sometimes also called the *dual VC dimension* of \mathcal{S} , denoted by $\text{VC}^*(\mathcal{S})$. The correspondence between \mathcal{S} and \mathcal{S}^* is explained in the following subsection.

2.3. VC duality. Let X and Y be infinite sets, and let $\Phi \subseteq X \times Y$. For $y \in Y$ we put

$$\Phi_y := \{x \in X : (x, y) \in \Phi\},$$

and we set

$$\mathcal{S}_{\Phi} := \{\Phi_y : y \in Y\} \subseteq 2^X.$$

We also write

$$\Phi^* \subseteq Y \times X := \{(y, x) \in Y \times X : (x, y) \in \Phi\}$$

for the dual of the binary relation Φ . In this way we obtain two set systems (X, \mathcal{S}_{Φ}) and $(Y, \mathcal{S}_{\Phi^*})$. To simplify notation, we denote the shatter function of \mathcal{S}_{Φ} by π_{Φ} , and its dual shatter function by π_{Φ}^* ; similarly for Φ^* in place of Φ . One verifies easily that given a finite set $A \subseteq X$, the assignment

$$A' \mapsto \bigcap_{x \in A'} \Phi_x^* \cap \bigcap_{x \in A \setminus A'} Y \setminus \Phi_x^*$$

defines a bijection

$$\mathcal{S}_{\Phi} \cap A \rightarrow S(\Phi_x^* : x \in A).$$

This implies:

Lemma 2.4. $\pi_{\Phi} = \pi_{\Phi^*}^*$.

We set $\text{VC}(\Phi) := \text{VC}(\mathcal{S}_\Phi)$, and similarly with IN and vc in place of VC . By the previous lemma, $\text{VC}(\Phi) = \text{IN}(\Phi^*)$, hence \mathcal{S}_Φ is a VC class iff \mathcal{S}_{Φ^*} is dependent. Reversing the role of Φ and Φ^* also yields $\pi_{\Phi^*} = \pi_\Phi^*$, hence $\text{VC}(\Phi^*) = \text{IN}(\Phi)$, and \mathcal{S}_{Φ^*} is a VC class iff \mathcal{S}_Φ is dependent. The following is also well-known (see, e.g., [7, 2.13 b]):

Lemma 2.5. $\text{VC}(\Phi) < 2^{1+\text{VC}(\Phi^*)}$. (In particular \mathcal{S}_Φ is a VC class iff \mathcal{S}_{Φ^*} is a VC class.)

Example 2.6. Suppose \mathcal{S}_Φ is finite (i.e., $\text{vc}(\Phi) = 0$). Then \mathcal{S}_{Φ^*} is also finite. (Take $y_1, \dots, y_N \in Y$, where $N = |\mathcal{S}_\Phi|$, such that $\mathcal{S}_\Phi = \{\Phi_{y_1}, \dots, \Phi_{y_N}\}$. Let $X_i = \Phi_{y_i}$ and $Y_i = \{y \in Y : \Phi_y = \Phi_{y_i}\}$ for $i \in [N]$; thus $\Phi = X_1 \times Y_1 \cup \dots \cup X_N \times Y_N$. Hence for each $x \in X$, Φ_x^* is a union of Y_1, \dots, Y_N , and so there are only finitely many choices for Φ_x^* . Thus \mathcal{S}_{Φ^*} is also finite, of size at most 2^N .)

Clearly every infinite set system \mathcal{S} on X is of the form $\mathcal{S} = \mathcal{S}_\Phi$ for some infinite set Y and some binary relation $\Phi \subseteq X \times Y$: just take $Y = \mathcal{S}$, $\Phi = \{(x, S) : x \in S, S \in \mathcal{S}\}$. The resulting set system \mathcal{S}_{Φ^*} on $Y = \mathcal{S}$ is called the dual \mathcal{S}^* of \mathcal{S} in [66, Section 10.3]. By the above $\text{VC}(\mathcal{S}^*) = \text{VC}^*(\mathcal{S})$. If \mathcal{S} is a dependent infinite set system on X , then by Lemmas 2.1 and 2.4, there is a real number $r \geq 0$ such that $\pi_{\mathcal{S}^*} = O(n^r)$, and the infimum of all such r is called the *dual VC density* of \mathcal{S} , denoted by $\text{vc}^*(\mathcal{S})$; note that $\text{vc}(\mathcal{S}^*) = \text{vc}^*(\mathcal{S})$ and $\text{vc}^*(\mathcal{S}) \leq \text{VC}^*(\mathcal{S})$.

Given $\Phi \subseteq X \times Y$ we write $\neg\Phi$ for the relative complement $(X \times Y) \setminus \Phi$ of Φ in $X \times Y$. We clearly have $\pi_{\neg\Phi}^* = \pi_\Phi^*$. It is also easy to show that given $\Phi, \Psi \subseteq X \times Y$ we have $\pi_{\Phi \cup \Psi}^* \leq \pi_\Phi^* \cdot \pi_\Psi^*$ and hence (using complementation) $\pi_{\Phi \cap \Psi}^* \leq \pi_\Phi^* \cdot \pi_\Psi^*$. By passing to duals and Lemma 2.4, this yields:

Lemma 2.7. Let $\Phi, \Psi \subseteq X \times Y$. Then

$$\text{vc}(\neg\Phi) = \text{vc}(\Phi), \quad \text{vc}(\Phi \cup \Psi) \leq \text{vc}(\Phi) + \text{vc}(\Psi), \quad \text{vc}(\Phi \cap \Psi) \leq \text{vc}(\Phi) + \text{vc}(\Psi).$$

VC dimension does not satisfy a similar subadditivity property for unions and intersections (cf. [27, Proposition 9.2.8]). In this way, VC density is better behaved than VC dimension.

An important class of relations Φ such that the associated set system \mathcal{S}_Φ is dependent are the stable ones. An *n-ladder* for Φ is a $2n$ -tuple $(a_1, \dots, a_n, b_1, \dots, b_n)$ where each $a_i \in X$ and each $b_j \in Y$, such that for all $i, j \in [n]$,

$$(a_i, b_j) \in \Phi \iff i \leq j.$$

If there is an n such that there is no n -ladder for Φ , then Φ is called *stable*, and Φ is said to be *unstable* otherwise. If Φ is stable then the largest n such that an n -ladder for Φ exists is called the *ladder dimension* of Φ ; if Φ is unstable then we say that the ladder dimension of Φ is infinite. Clearly if Φ is stable then \mathcal{S}_Φ is a VC class (with VC dimension bounded by the ladder dimension). It is well-known that Φ is stable iff Φ^* is stable (e.g. [88, Exercise II.2.8]), and that Boolean combinations of stable relations are stable.

2.4. Breadth. In many cases of interest complicated set systems are generated by simpler collections of subsets, and then the following lemma (essentially due to Dudley) can be used to show that the resulting set system is dependent. For this let X be a set and \mathcal{B} be a collection of subsets of X .

Lemma 2.8. *Let $N > 0$ and suppose \mathcal{S} is a set system on X such that each set in \mathcal{S} is a Boolean combination of at most N sets in \mathcal{B} . Then $\pi_{\mathcal{S}}^*(t) \leq \pi_{\mathcal{B}}^*(Nt)$ for each t . (In particular, if \mathcal{B} is dependent then so is \mathcal{S} .)*

Proof. Let $A_1, \dots, A_t \in \mathcal{S}$, and let each A_i be a Boolean combination of the sets $B_{i1}, \dots, B_{iN} \in \mathcal{B}$. Then the Boolean algebra of subsets of X generated by the sets A_i ($i \in [t]$) is contained in the Boolean algebra generated by the sets B_{ij} ($i \in [t], j \in [N]$), and every atom of the former Boolean algebra contains an atom of the latter. \square

Suppose there is a $d > 0$ such that every non-empty intersection $B_1 \cap \dots \cap B_n$ of $n > d$ sets from \mathcal{B} equals an intersection of a subset consisting of d of the B_i . We call the smallest such integer $d > 0$ the *breadth* of \mathcal{B} . This choice of terminology is motivated by lattice theory: Given a (meet-) semilattice (L, \wedge) , the smallest $d > 0$ (if it exists) such that any meet $b_1 \wedge \dots \wedge b_n$ of $n > d$ elements of L equals the meet of d of the b_i is called the *breadth* of L ; if there is no such d we say that L has infinite breadth. (See [16, Section II.5, Exercise 6, and Section IV.10].) So if \mathcal{B} is closed under (finite) intersection and only contains non-empty subsets of X , then the breadth of \mathcal{B} , viewed as a sub-semilattice of $(2^X, \cap)$, agrees with the breadth of \mathcal{B} as defined above. Every set system of finite breadth is dependent:

Lemma 2.9. $\text{breadth}(\mathcal{B}) \geq \text{IN}(\mathcal{B})$.

Proof. Suppose $d := \text{breadth}(\mathcal{B}) < n := \text{IN}(\mathcal{B})$. Let $B_1, \dots, B_n \in \mathcal{B}$ such that $|S(B_1, \dots, B_n)| = 2^n$. Choose $I \subseteq [n]$ with $|I| = d$ and $\bigcap_{i \in I} B_i = \bigcap_{i \in [n]} B_i$, and take $j \in [n] \setminus I$. Then $\bigcap_{i \in [n] \setminus \{j\}} B_i = \bigcap_{i \in [n]} B_i$ and hence $(X \setminus B_j) \cap \bigcap_{i \in [n] \setminus \{j\}} B_i = \emptyset$, contradicting $\text{IN}(\mathcal{B}) = n$. \square

The previous two lemmas in combination with Lemma 2.1 immediately yield the following useful fact (cf. [25, Chapter 5, Lemma 2.6]):

Corollary 2.10. *Suppose \mathcal{B} has breadth d , let $N > 0$, and let \mathcal{S} be a set system on X with the property that each set in \mathcal{S} is a Boolean combination of at most N sets in \mathcal{B} . Then*

$$\pi_{\mathcal{S}}^*(t) \leq \sum_{i=0}^d \binom{Nt}{i} \quad \text{for every } t.$$

In particular, $\pi_{\mathcal{S}}^(t) = O(t^d)$ and hence $\text{vc}^*(\mathcal{S}) \leq d$.*

Example 2.11. Let $<$ be a linear ordering on X . We first recall some terminology: A subset S of X is said to be *convex* (with respect to $<$) if for all $s, s' \in S$ and $x \in X$ the implication $s < x < s' \Rightarrow x \in S$ holds. So \emptyset and singleton subsets are convex, as are intervals in X . Here and in the rest of the paper, an *interval* in X is a subset of the form

$$(a, b) := \{x \in X : a < x < b\}$$

where a, b are elements of $X \cup \{\pm\infty\}$ with $a < b$. Other examples of a convex subset of X are its initial segments: a subset S of X is an *initial segment* of X if for all $s \in S$ and $x \in X$, the implication $x < s \Rightarrow x \in S$ holds. Now let \mathcal{S} be the family of unions of at most N convex subsets of X , for some given $N \in \mathbb{N}$, and let \mathcal{B} be the collection of all initial segments of X . Then \mathcal{B} has breadth 1, and every set in \mathcal{S} is a Boolean combination of at most $2N$ sets in \mathcal{B} . Thus $\pi_{\mathcal{S}}^*(t) = O(t)$ by Corollary 2.10.

Example 2.12. Let K be a field and $v: K \rightarrow \Gamma_\infty := \Gamma \cup \{\infty\}$ be a valuation on K . By an *open ball* in K we mean any subset of K of the form $\{x \in K : v(x - a) > \gamma\}$ where $a \in K, \gamma \in \Gamma_\infty$; similarly a set of the form $\{x \in K : v(x - a) \geq \gamma\}$ is called a *closed ball* in K . A *ball* in K is an open or a closed ball in K . Any two given balls in K are either disjoint, or one contains the other. Hence the collection \mathcal{B} of balls in a given valued field has breadth 1. Thus if \mathcal{S} is the family of all Boolean combinations of at most N balls in K , for some $N \in \mathbb{N}$, then $\pi_{\mathcal{S}}^*(t) = O(t)$.

The preceding examples can be subsumed under the following general example (inspired by [2]):

Example 2.13. A family \mathcal{B} of subsets of X is said to be *directed* if \mathcal{B} has breadth 1; i.e., for all $B, B' \in \mathcal{B}$ with $B \cap B' \neq \emptyset$ one has $B \subseteq B'$ or $B' \subseteq B$. If $\mathcal{B} \subseteq 2^X$ is directed and \mathcal{S} is the family of Boolean combinations of at most N sets in \mathcal{B} , for some $N \in \mathbb{N}$, then $\pi_{\mathcal{S}}^*(t) = O(t)$.

We also note:

Example 2.14. Let G be a group and let \mathcal{H} be a collection of subgroups of G with breadth d . Let $\mathcal{B} = \{gH : g \in G, H \in \mathcal{H}\}$ be the set of all (left) cosets of subgroups from \mathcal{H} . Then \mathcal{B} also has breadth d . This follows from the general fact that if H_1, \dots, H_n are subgroups of $G, g_1, \dots, g_n \in G$, then the intersection $\bigcap_{i \in [n]} g_i H_i$ is either empty or a coset of $\bigcap_{i \in [n]} H_i$. (So if \mathcal{S} is a family of Boolean combinations of at most N elements of \mathcal{B} , for some $N \in \mathbb{N}$, then $\pi_{\mathcal{S}}^*(t) = O(t^d)$.)

In connection with the previous example it is worth recording:

Lemma 2.15 (Poizat). *Let G be a group and let \mathcal{H} be a collection of subgroups of G . Then $\text{breadth}(\mathcal{H}) = \text{IN}(\mathcal{H})$.*

Proof. By Lemma 2.9 we already know that $\text{breadth}(\mathcal{H}) \geq \text{IN}(\mathcal{H})$. Suppose this inequality is strict. Then there are $H_1, \dots, H_{n+1} \in \mathcal{H}$, where $n = \text{IN}(\mathcal{H})$, such that $\bigcap_{i \in [n+1]} H_i \neq \bigcap_{i \in [n+1] \setminus \{j\}} H_i$ for each $j \in [n+1]$. So for each $j \in [n+1]$ we may take $g_j \in \left(\bigcap_{i \in [n+1] \setminus \{j\}} H_i \right) \setminus H_j$. Then for every subset I of $[n+1]$ the element $g_I := \prod_{i \in I} g_i$ (with $g_\emptyset = 1$) is in $\bigcap_{i \in [n+1] \setminus I} H_i \cap \bigcap_{i \in I} (G \setminus H_i)$. This contradicts $\text{IN}(\mathcal{H}) = n$. \square

Example. Let \mathcal{S} be the collection of all subgroups of $(\mathbb{Z}, +)$. Then \mathcal{S} has infinite breadth, hence infinite independence dimension by the previous lemma, and thus is not a VC class by Lemma 2.5. In particular, the collection of arithmetic progressions $a + b\mathbb{Z}$ ($a, b \in \mathbb{Z}$) in \mathbb{Z} is also not a VC class.

If our family \mathcal{B} has finite breadth d , then the Helly number of \mathcal{B} is at most d . The *Helly number* of \mathcal{B} is defined as the smallest $d > 0$ such that every finite subfamily $\{B_1, \dots, B_n\}$ of \mathcal{B} with $n > d$ which is d -consistent, is consistent, that is to say: if for every $I \in \binom{[n]}{d}$ we have $\bigcap_{i \in I} B_i \neq \emptyset$, then $\bigcap_{i \in [n]} B_i \neq \emptyset$. Note however that conversely, the breadth may be infinite yet the Helly number finite, even in the case of cosets: the collection of arithmetic progressions in \mathbb{Z} is independent, but has Helly number 2. Also, not every VC class has finite Helly number: the family whose members are the subsets of \mathbb{R} with two connected components, though a VC class (of VC dimension 4), has infinite Helly number. (For each n the elements $[0, i) \cup (i + 1, n]$, $i = 0, \dots, n - 1$ of this family form an $n - 1$ -consistent subfamily which is inconsistent.)

The following example is a prototype for finite-breadth families when we have a dimension function at our disposal:

Example 2.16. Define the *height* of \mathcal{B} to be the largest d (if it exists) such that there are $B_1, \dots, B_d \in \mathcal{B}$ with

$$B_1 \supsetneq B_1 \cap B_2 \supsetneq \dots \supsetneq B_1 \cap \dots \cap B_d \neq \emptyset.$$

So \mathcal{B} has height 0 iff \mathcal{B} does not contain a non-empty set, and \mathcal{B} has height 1 iff \mathcal{B} does contain a non-empty set, but any two distinct elements of \mathcal{B} are disjoint. Clearly if \mathcal{B} has height $d > 0$, then the breadth of \mathcal{B} is at most d . If \mathcal{B} has height $d > 1$ and in addition \mathcal{B} has a largest element (with respect to inclusion) then the breadth of \mathcal{B} is smaller than d : to see this let $B_1, \dots, B_d \in \mathcal{B}$ with $\bigcap_{i \in [d]} B_i \neq \emptyset$ be given; if B_1 is the largest element B of \mathcal{B} then clearly $\bigcap_{i \in [d]} B_i = \bigcap_{i \in [d] \setminus \{1\}} B_i$, and otherwise we have a chain

$$B \supsetneq B_1 \supseteq B_1 \cap B_2 \supseteq \dots \supseteq B_1 \cap \dots \cap B_d \neq \emptyset,$$

hence $\bigcap_{i \in [j]} B_i = \bigcap_{i \in [j+1]} B_i$ and so $\bigcap_{i \in [d]} B_i = \bigcap_{i \in [d] \setminus \{j+1\}} B_i$, for some $j \in [d-1]$.

The following observation (the proof of which we leave to the reader) allows us to produce new finite-breadth set systems from old ones:

Lemma 2.17. *Let $\mathcal{B}, \mathcal{B}'$ be set systems on X and X' , respectively, and consider the set system*

$$\mathcal{B} \boxtimes \mathcal{B}' := \{B \times B' : B \in \mathcal{B}, B' \in \mathcal{B}'\}$$

on $X \times X'$. Then

$$\text{breadth}(\mathcal{B} \boxtimes \mathcal{B}') \leq \text{breadth}(\mathcal{B}) + \text{breadth}(\mathcal{B}'),$$

and this inequality is an equality if both \mathcal{B} and \mathcal{B}' have breadth larger than 1 and contain a largest element (with respect to inclusion).

This lemma immediately yields:

Corollary 2.18. *Let $\mathcal{B}, \mathcal{B}'$ be set systems on X . Then the set system*

$$\mathcal{B} \cap \mathcal{B}' := \{B \cap B' : B \in \mathcal{B}, B' \in \mathcal{B}'\}$$

on X has breadth at most $\text{breadth}(\mathcal{B}) + \text{breadth}(\mathcal{B}')$.

Example. Suppose $<$ is a linear ordering of X and \mathcal{B} is the collection of convex subsets of X . Every element of \mathcal{B} can be expressed as an intersection of an initial segment of $(X, <)$ with a final segment of $(X, <)$ (i.e., an initial segment of the linearly ordered set $(X, >)$). Hence $\text{breadth}(\mathcal{B}) = 2$.

If \mathcal{B} is a sublattice of $(2^X, \cap, \cup)$ which does not contain \emptyset and X , then \mathcal{B} and the set system $\neg\mathcal{B} := \{X \setminus B : B \in \mathcal{B}\}$ have the same breadth; this is an immediate consequence of the following lemma:

Lemma 2.19. *Suppose \mathcal{B} is closed under (finite) intersections and unions, and \mathcal{B} does not contain the empty set. Then for each d the following are equivalent:*

- (1) *For all $B_1, \dots, B_{d+1} \in \mathcal{B}$ there is some $i \in [d+1]$ such that $\bigcap_{j \neq i} B_j \subseteq B_i$;*
- (2) *for all $B_1, \dots, B_{d+1} \in \mathcal{B}$ there is some $i \in [d+1]$ such that $B_i \subseteq \bigcup_{j \neq i} B_j$.*

Proof. To see (1) \Rightarrow (2) apply (1) to $B'_i = \bigcup_{j \neq i} B_j$ ($i \in [d+1]$) in place of the B_i , and for the converse implication apply (2) to $B''_i = \bigcap_{j \neq i} B_j$ ($i \in [d+1]$). \square

We finish our discussion of breadth by a surprising connection between breadth and stability. We will not use this observation later in the paper, but we include it here since it shows, under the assumption of stability, the ubiquity of set systems of *infinite* breadth. The breadth of a relation between two sets is by definition the breadth of the associated set system, cf. Section 2.3.

Proposition 2.20. *Let X, Y be infinite sets and $\Phi \subseteq X \times Y$ be a relation. If $\text{vc}(\Phi) > 0$ then Φ is unstable, or at least one of Φ or $\neg\Phi$ has infinite breadth.*

At the root of Proposition 2.20 is a theorem of Balogh and Bollobás [10], which we explain first. For this we need some additional terminology: Let (X, \mathcal{S}) and (X', \mathcal{S}') be set systems. We say that (X, \mathcal{S}) *contains* (X', \mathcal{S}') as a trace if there exists an injective map $f: X' \rightarrow X$ such that $f(\mathcal{S}') \subseteq \mathcal{S} \cap f(X')$. For example, if (X, \mathcal{S}) is a set system and $A \subseteq X$ then (X, \mathcal{S}) trivially contains $(A, \mathcal{S} \cap A)$. Also, if (X, \mathcal{S}) contains (X', \mathcal{S}') , and (X', \mathcal{S}') contains (X'', \mathcal{S}'') , then (X, \mathcal{S}) contains (X'', \mathcal{S}'') .

For $k \geq 2$ consider now the following set systems on $[k]$:

$$\begin{aligned} \mathcal{C}_k &= \{[i] : i \in [k]\} && \text{(the } k\text{-chain)} \\ \mathcal{S}_k &= \{\{i\} : i \in [k]\} && \text{(the } k\text{-star)} \\ \mathcal{T}_k &= \{[k] \setminus \{i\} : i \in [k]\} && \text{(the } k\text{-costar)}. \end{aligned}$$

Balogh and Bollobás [10, Theorem 1] showed that these set systems are unavoidable among sufficiently large set systems. More precisely: *for all integers $k, l, m \geq 2$ there is some $N = N(k, l, m)$ such that every set system \mathcal{S} on a finite base set with $|\mathcal{S}| \geq N$ contains the k -chain, the l -star, or the m -costar.* (Note that there is no condition on the size of the base set in this statement.)

Proof of Proposition 2.20. Let $\mathcal{S} = \mathcal{S}_\Phi$. We first observe, for $k \geq 2$:

- (1) \mathcal{S} contains \mathcal{C}_k iff there is a k -ladder for Φ ;
- (2) if $\text{breadth}(\Phi^*) \geq k$ then \mathcal{S} contains \mathcal{T}_k ; and
- (3) if \mathcal{S} contains \mathcal{T}_{k+1} then $\text{breadth}(\Phi^*) \geq k$.

Part (1) is obvious. For (2) note that $\text{breadth}(\Phi^*) \geq k$ iff there exist elements x_1, \dots, x_k of X such that $\bigcap_{j \in [k]} \Phi_{x_j}^* \neq \emptyset$ and for each $i \in [k]$,

$$(Y \setminus \Phi_{x_i}^*) \cap \bigcap_{j \in [k] \setminus \{i\}} \Phi_{x_j}^* \neq \emptyset,$$

and for such choice of x_i , setting $X' = \{x_1, \dots, x_k\}$ we have $X' \setminus \{x_i\} \in \mathcal{S} \cap X'$ for each i . Similarly, for (3), if $X' = \{x_1, \dots, x_{k+1}\} \in \binom{X}{k+1}$ such that $X' \setminus \{x_i\} \in \mathcal{S} \cap X'$ for each $i \in [k+1]$, then for each such i we have

$$(Y \setminus \Phi_{x_i}^*) \cap \bigcap_{j \in [k+1] \setminus \{i\}} \Phi_{x_j}^* \neq \emptyset;$$

in particular, taking $i = k+1$ we see that $\bigcap_{j \in [k]} \Phi_{x_j}^* \neq \emptyset$, and for each $i \in [k]$ we have $(Y \setminus \Phi_{x_i}^*) \cap \bigcap_{j \in [k] \setminus \{i\}} \Phi_{x_j}^* \neq \emptyset$, hence $\text{breadth}(\Phi^*) \geq k$. Also note that (2) and (3) are true with $\mathcal{T}_k, \mathcal{T}_{k+1}$ and Φ^* replaced by $\mathcal{S}_k, \mathcal{S}_{k+1}$ and $\neg\Phi^*$, respectively.

Suppose now that $\text{vc}(\Phi) > 0$, i.e., \mathcal{S} is infinite. Then $\mathcal{S}^* = \mathcal{S}_{\Phi^*}$ is also infinite (see Example 2.6). Then we have $\text{vc}(\mathcal{S}^*) \geq 1$, hence there are arbitrarily large n and $B \in \binom{Y}{n}$ such that $|\mathcal{S}^* \cap B| \geq n^{1/2}$. In particular, for each N there is a finite subset B_N of Y with $|\mathcal{S}^* \cap B_N| \geq N$. Now suppose Φ is stable; then Φ^* is also stable. Let $k_0 \geq 2$ be larger than the ladder dimension of Φ^* . Then if $k \geq 2$ and $N \geq N(k_0, k, k)$ then $\mathcal{S}^* \cap B_N$ (and

hence \mathcal{S}^*) contains the k -star or the k -costar. Thus by observation (3) above, at least one of Φ or $\neg\Phi$ has infinite breadth. \square

Of course, the converse of the implication in this proposition also holds: if $\text{vc}(\Phi) = 0$ then \mathcal{S}_Φ is finite, hence trivially Φ is stable, and both Φ and $\neg\Phi$ have finite breadth, since $\mathcal{S}_{\neg\Phi}$ is finite as well.

Example. Let $<$ be a linear ordering of X . Suppose \mathcal{B} is the collection of initial segments of $(X, <)$, as in Example 2.11. Then $\neg\mathcal{B} = \{X \setminus B : B \in \mathcal{B}\}$ consists of final segments of $(X, <)$. Hence \mathcal{B} and $\neg\mathcal{B}$ both have finite breadth (indeed, breadth 1). Proposition 2.20 shows that phenomena such as these are confined to unstable contexts (for infinite set systems).

Using Lemma 2.19, Proposition 2.20 also implies:

Corollary 2.21. *Suppose X and Y are infinite sets and $\Phi \subseteq X \times Y$ such that \mathcal{S}_Φ is an infinite sublattice of $(2^X, \cap, \cup)$ of finite breadth, with $\emptyset, X \notin \mathcal{S}_\Phi$. Then Φ is unstable.*

3. THE MODEL-THEORETIC CONTEXT

Throughout this section we fix a first-order language \mathcal{L} , and we let $\varphi(x; y)$ be a partitioned \mathcal{L} -formula (as defined in the introduction), with object variables $x = (x_1, \dots, x_m)$ and parameter variables $y = (y_1, \dots, y_n)$. The formula φ gives rise, in a given \mathcal{L} -structure, to a set system. The associated parameters introduced in the previous section (shatter function, VC density etc.) are elementary invariants of the structure in question. In Section 3.2 below we also introduce the VC density function vc^T of a complete first-order \mathcal{L} -theory T with no finite models: if $\text{vc}^T(n)$ is finite then $\text{vc}^T(n)$ is a uniform bound on the VC density of all partitioned formulas in T having n parameter variables. In Section 3.3 we illustrate this concept by computing $\text{vc}^T(1)$ for various T . In Section 3.4 we then extend the definition of dual VC density to finite sets of formulas; this is convenient for later sections, but, as we see in Section 3.5, does not add much extra generality. There is some indication that computing VC density is easier when only parameters coming from initial segments of indiscernible sequences are considered; although we will not pursue these issues in the rest of the paper, we think that the relationship between quantities like VC or vc and their “indiscernible” counterparts deserves further investigation; we explore some connections in the last subsection.

3.1. VC density of definable families. Given an \mathcal{L} -structure \mathbf{M} and a tuple $b \in M^n$, we denote the subset of M^m defined by the \mathcal{L} -formula $\varphi(x; b)$ with parameters b in \mathbf{M} by

$$\varphi^{\mathbf{M}}(M^m; b) := \{a \in M^m : \mathbf{M} \models \varphi(a; b)\}.$$

A subset of M^m is called *definable* (in \mathbf{M}) if it is of the form $\varphi^{\mathbf{M}}(M^m; b)$, for some φ and b . We also denote by

$$\mathcal{S}_\varphi^{\mathbf{M}} := \{\varphi^{\mathbf{M}}(M^m; b) : b \in M^n\}$$

the family of subsets of M^m defined by φ in \mathbf{M} , and we call $(M^m, \mathcal{S}_\varphi^{\mathbf{M}})$ the *set system associated with φ in \mathbf{M}* . More generally, to a given collection $\Phi(x) = \{\varphi_i(x; y_i)\}_{i \in I}$ of partitioned \mathcal{L} -formulas in the tuple of object variables x (and in various tuples of parameter variables y_i) we may associate the set system

$$\mathcal{S}_\Phi^{\mathbf{M}} := \{\varphi_i^{\mathbf{M}}(M^m; b) : i \in I, b \in M^{|y_i|}\}$$

on M^m defined by the instances of the formulas φ_i . If the \mathcal{L} -structure \mathbf{M} is understood from the context, we drop the superscript \mathbf{M} in our notation.

Suppose now \mathbf{M} is an infinite \mathcal{L} -structure. As usual, say that φ is invariant under an extension $\mathbf{M} \subseteq \mathbf{N}$ of \mathcal{L} -structures if $\mathbf{M} \models \varphi(a; b) \iff \mathbf{N} \models \varphi(a; b)$ for all $a \in M^m$ and $b \in M^n$. The following is obvious:

Lemma 3.1. *Suppose \mathbf{N} is an \mathcal{L} -structure with $\mathbf{M} \subseteq \mathbf{N}$ and φ is invariant under $\mathbf{M} \subseteq \mathbf{N}$. Then $\mathcal{S}_\varphi^{\mathbf{M}} \subseteq M^m \cap \mathcal{S}_\varphi^{\mathbf{N}}$, hence $\pi_\varphi^{\mathbf{M}} \leq \pi_\varphi^{\mathbf{N}}$ and therefore $\text{VC}(\mathcal{S}_\varphi^{\mathbf{M}}) \leq \text{VC}(\mathcal{S}_\varphi^{\mathbf{N}})$ and $\text{vc}(\mathcal{S}_\varphi^{\mathbf{M}}) \leq \text{vc}(\mathcal{S}_\varphi^{\mathbf{N}})$.*

For each $s, t \in \mathbb{N}$, consider the \mathcal{L} -sentence

$$\pi_\varphi^{s,t} := \forall x^{(1)} \dots \forall x^{(t)} \forall y^{(1)} \dots \forall y^{(s+1)} \left(\bigvee_{1 \leq i < j \leq t} \bigwedge_{k=1}^m x_k^{(i)} = x_k^{(j)} \vee \bigvee_{1 \leq k < l \leq s+1} \bigwedge_{1 \leq i \leq t} \varphi(x^{(i)}; y^{(k)}) \leftrightarrow \varphi(x^{(i)}; y^{(l)}) \right),$$

where $x^{(i)} = (x_1^{(i)}, \dots, x_m^{(i)})$ and $y^{(j)} = (y_1^{(j)}, \dots, y_n^{(j)})$ are tuples of new variables. Then, with $\pi_\varphi^{\mathbf{M}} := \pi_{\mathcal{S}_\varphi^{\mathbf{M}}}$ denoting the shatter function of $\mathcal{S}_\varphi^{\mathbf{M}}$, we obviously have:

Lemma 3.2. *For each $s, t \in \mathbb{N}$,*

$$\mathbf{M} \models \pi_\varphi^{s,t} \iff \pi_\varphi^{\mathbf{M}}(t) \leq s.$$

In particular, if \mathbf{N} is an \mathcal{L} -structure with $\mathbf{M} \equiv \mathbf{N}$, then $\pi_\varphi^{\mathbf{M}} = \pi_\varphi^{\mathbf{N}}$.

From now until the end of this section we fix a complete \mathcal{L} -theory T with only infinite models, and let \mathbf{M} range over models of T . By the previous lemma we may set

$$\pi_\varphi := \pi_\varphi^{\mathbf{M}}, \quad \text{VC}(\varphi) := \text{VC}(\mathcal{S}_\varphi^{\mathbf{M}}), \quad \text{vc}(\varphi) := \text{vc}(\mathcal{S}_\varphi^{\mathbf{M}}),$$

where \mathbf{M} is an arbitrarily chosen model of T . We call π_φ the *shatter function* of φ (in T), and we call $\text{VC}(\varphi)$ and $\text{vc}(\varphi)$ the *VC dimension* of φ and *VC density* of φ (in T), respectively. If we want to stress the dependence of π_φ on T we write π_φ^T , and similarly for VC and vc .

Note that the definition of π_φ only depends on the set system \mathcal{S}_φ and not on the particular representing formula φ . In particular, π_φ remains unchanged under \emptyset -definable reparameterizations:

Lemma 3.3. *Let $\gamma(z; y)$ be an \mathcal{L} -formula, where $z = (z_1, \dots, z_l)$, which defines the graph of a map $g: M^l \rightarrow M^n$. Let $\sigma(x; z) := \exists y(\gamma(z; y) \wedge \varphi(x; y))$, so*

$$\mathcal{S}_\sigma = \{\varphi^{\mathbf{M}}(M^m; g(c)) : c \in M^l\}.$$

Then $\pi_\sigma \leq \pi_\varphi$, with equality if g is surjective.

The *dual* of the partitioned \mathcal{L} -formula $\varphi(x; y)$ is $\varphi^*(y; x) := \varphi(x; y)$; that is, $\varphi^*(y; x)$ is syntactically the same \mathcal{L} -formula φ , only with the role of the object and parameter variables interchanged. We call $\text{VC}^*(\varphi) := \text{VC}(\varphi^*)$ and $\text{vc}^*(\varphi) := \text{vc}(\varphi^*)$ the *dual VC dimension* and *dual VC density* of φ , respectively. By Lemma 2.4 we have $\pi_\varphi^* = \pi_{\varphi^*}$ and hence $\text{VC}^*(\varphi) = \text{IN}(\varphi)$ and

$$\text{vc}^*(\varphi) = \inf \{r \in \mathbb{R}^{>0} : \pi_\varphi^*(n) = O(n^r)\}.$$

If any of the quantities $\text{VC}(\varphi)$, $\text{vc}(\varphi)$, $\text{VC}^*(\varphi)$, $\text{vc}^*(\varphi)$ is finite, then so are all the others, and in this case we say that φ is *dependent* or that φ *defines a VC class*. Note that

for every partitioned \mathcal{L} -formula $\varphi(x; y)$ we have $\text{vc}(\varphi) \geq 0$, with equality if \mathcal{S}_φ is finite. If \mathcal{S}_φ is infinite then $\text{vc}(\varphi) \geq 1$. (See the remarks following Lemma 2.2.)

Letting $\Phi := \varphi(M^m; M^n)$ and $X := M^m$, $Y := M^n$, in the notation introduced in the previous subsection we have $\mathcal{S}_\Phi = \mathcal{S}_\varphi$ and $\mathcal{S}_{\Phi^*} = \mathcal{S}_{\varphi^*}$. Hence Lemma 2.7 yields:

Corollary 3.4. *We have $\text{vc}(\neg\varphi) = \text{vc}(\varphi)$, and if $\psi(x; z)$ is another partitioned \mathcal{L} -formula, then $\text{vc}(\varphi \wedge \psi), \text{vc}(\varphi \vee \psi) \leq \text{vc}(\varphi) + \text{vc}(\psi)$.*

From Lemma 2.2 one also obtains the invariance of vc under inverse images of surjective \emptyset -definable maps:

Corollary 3.5. *Let $\delta(v; x)$ be an \mathcal{L} -formula, where $v = (v_1, \dots, v_k)$, which defines the graph of a map $f: M^k \rightarrow M^m$, and let $\rho(v; y) := \exists x(\delta \wedge \varphi)$, so $\mathcal{S}_\rho = f^{-1}(\mathcal{S}_\varphi)$. Then $\pi_\rho \leq \pi_\varphi$, with equality if f is surjective.*

The theory T is NIP iff every partitioned \mathcal{L} -formula defines a VC class. The theorem of Shelah [86] already mentioned in the introduction shows that in order for every partitioned \mathcal{L} -formula $\varphi(x; y)$ to define a VC class, it is enough that this holds for all such $\varphi(x; y)$ with a single parameter variable (i.e., $|y| = 1$). Hence if for each partitioned \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$ the set system \mathcal{S}_φ has finite breadth then T is NIP, by Lemma 2.9. The theory T is said to be stable if for every partitioned \mathcal{L} -formula $\varphi(x; y)$ the associated relation $\Phi = \varphi(M^m; M^n)$ is stable (in the sense of Section 2.3); if T is stable then for each $\varphi(x; y)$ with \mathcal{S}_φ infinite, at least one of \mathcal{S}_φ or $\mathcal{S}_{\neg\varphi}$ has infinite breadth, by Proposition 2.20. (Corollary 2.21 of the same proposition also yields that if T is stable then all finite-breadth sublattices \mathcal{S} of the lattice of all subsets of M^m which have the form $\mathcal{S} = \mathcal{S}_\varphi$ for some \mathcal{L} -formula $\varphi(x; y)$ with $|x| = m$ are finite.)

3.2. VC density of a theory. We define the *VC density of T* to be the function

$$\text{vc} = \text{vc}^T: \mathbb{N} \rightarrow \mathbb{R}^{\geq 0} \cup \{\infty\}$$

given by

$$\text{vc}(n) := \sup \{ \text{vc}(\varphi) : \varphi(x; y) \text{ is an } \mathcal{L}\text{-formula with } |y| = n \}.$$

Note that we could have also defined vc^T as

$$\text{vc}(m) = \sup \{ \text{vc}^*(\varphi) : \varphi(x; y) \text{ is an } \mathcal{L}\text{-formula with } |x| = m \}.$$

In the introduction we already observed that $\text{vc}(m) \geq m$ for every m . If \mathcal{L}' is an expansion of \mathcal{L} and $T' \supseteq T$ a complete \mathcal{L}' -theory, then $\text{vc}^T \leq \text{vc}^{T'}$, with equality if T' is an expansion of T by definitions. Moreover, vc does not change under expansions by constants:

Lemma 3.6. *Let $\mathcal{L}' = \mathcal{L} \cup \{c_i : i \in I\}$ where the c_i are new constant symbols, and let $T' \supseteq T$ be a complete \mathcal{L}' -theory. Then $\text{vc}^T = \text{vc}^{T'}$.*

Proof. Let $M' \models T'$ and $C := \{c_i^{M'} : i \in I\} \subseteq M'$. Let $\varphi(x; y, z)$ be an \mathcal{L} -formula with $|x| = m$, and let $c \in C^{|z|}$. Then $\pi_{\varphi(x; y, c)}^*(t) \leq \pi_{\varphi(x; y, z)}^*(t)$ for every t , hence $\text{vc}^*(\varphi(x; y, c)) \leq \text{vc}^*(\varphi(x; y, z)) \leq \text{vc}^T(m)$ and thus $\text{vc}^{T'}(m) \leq \text{vc}^T(m)$. \square

It is clear that $\text{vc}(n) \leq \text{vc}(n+1)$ for every n , by viewing a formula with n parameter variables as one with $n+1$ parameters; perhaps less obviously:

Lemma 3.7. *$\text{vc}(n) + 1 \leq \text{vc}(n+1)$ for every n .*

Proof. By the preceding lemma we may assume that \mathcal{L} contains a constant symbol 0. Let $\varphi(x; y)$ be a partitioned \mathcal{L} -formula with $|x| = m$, $|y| = n$. We construct a formula $\psi(x, x_{m+1}; y, y_{n+1})$ with $\pi_\varphi(t) \cdot t \leq \pi_\psi(2t)$ for every t (hence $\text{vc}(\varphi) + 1 \leq \text{vc}(\psi)$), which then shows the lemma. We set

$$\psi := (x_{m+1} = 0 \wedge \varphi(x; y)) \vee (x_{m+1} = y_{n+1}).$$

Then for $b \in M^n$, $c \in M$ we have

$$\psi(M^{m+1}; b, c) = (\varphi(M^m; b) \times \{0\}) \cup (M^m \times \{c\}).$$

Let $A \subseteq M^m$ with $|A| = t$ and $\pi_\varphi(t) = |A \cap \mathcal{S}_\varphi|$. Choose pairwise distinct elements $a_1, \dots, a_t \in M \setminus \{0\}$ and an arbitrary element $a' \in M^m$, and set

$$A' := (A \times \{0\}) \cup \{(a', a_1), \dots, (a', a_t)\}.$$

Then $|A'| = 2t$, and for $b \in M^n$ and $j = 1, \dots, t$ we have

$$A' \cap \psi(M^{m+1}; b, a_j) = ((A \cap \varphi(M^m; b)) \times \{0\}) \cup \{(a', a_j)\}.$$

Take $b_1, \dots, b_k \in M^n$, $k = \pi_\varphi(t)$, such that the sets $A \cap \varphi(M^m; b_i)$, $i = 1, \dots, k$, are pairwise distinct. Then the sets $A' \cap \psi(M^{m+1}; b_i, a_j)$ (where $i = 1, \dots, k$, $j = 1, \dots, t$) are also pairwise distinct. Hence $\pi_\psi(2t) \geq |A' \cap \mathcal{S}_\psi| \geq k \cdot t = \pi_\varphi(t) \cdot t$ as claimed. \square

In this paper we prove, for many (unstable) NIP theories T of interest, that $\text{vc}^T(m) < \infty$ for every m , and in fact, in these cases we establish that $\text{vc}^T(m)$ is bounded by a linear function of m . Note, however, that T NIP does not imply that $\text{vc}^T(m) < \infty$ for all m : it is easy to see that for every T (whether NIP or not) we have $\text{vc}^{T^{\text{eq}}}(1) = \infty$, whereas T is NIP iff T^{eq} is NIP. (We thank Martin Ziegler for pointing this out.)

By Laskowski's proof [55] of Shelah's theorem [86], the VC dimension $\text{VC}(\varphi)$ of an \mathcal{L} -formula $\varphi(x; y)$ is bounded in terms of the VC dimensions of certain \mathcal{L} -formulas with a single parameter variable (which, however, are astronomical, involving iterated Ramsey numbers). This together with the examples below raises the following question, the answer to which we don't know:

Question. If $\text{vc}^T(1) < \infty$, is $\text{vc}^T(m) < \infty$ for every m ?

Provided the answer to this question is positive, one may then also ask how $\text{vc}(m)$ depends on m and $\text{vc}(1)$; e.g.: *is there a function $\beta: \mathbb{N} \times \mathbb{R}^{\geq 0} \rightarrow \mathbb{R}^{\geq 0}$, independent of T , with the property that if $\text{vc}^T(1) < \infty$, then $\text{vc}^T(m) \leq \beta(m, \text{vc}^T(1))$ for every m ?* (In all examples which we considered where vc^T is known to be real-valued, vc^T grows at worst linearly.)

3.3. Computing $\text{vc}^T(1)$. In concrete cases it is often easy to see that $\text{vc}^T(1) = 1$:

Example 3.8. Suppose that \mathbf{M} is strongly minimal. The collection $\mathcal{B} = \binom{M}{1}$ of one-element subsets of M has breadth 1; so $\text{vc}^T(1) = 1$. (Corollary 2.10.)

Example 3.9. Suppose that \mathcal{L} contains a binary relation symbol " $<$ ", $\mathbf{M} = (M, <, \dots)$ is an expansion of a linearly ordered set $(M, <)$, and $T = \text{Th}(\mathbf{M})$ is weakly o-minimal. Then for every partitioned \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$ there exists an integer $N \geq 0$ such that for every $b \in M^m$, the set $\varphi^{\mathbf{M}}(M; b)$ is a finite union of at most N convex subsets of M . Hence $\text{vc}^T(1) = 1$ by Example 2.11.

Example 3.10. Suppose that \mathcal{L}_{div} is the expansion of the language $\{0, 1, +, -, \times\}$ of rings by a binary relation symbol “ $|$ ”. In a field K equipped with a valuation $v: K \rightarrow \Gamma \cup \{\infty\}$, we interpret $|$ by putting $a|b : \iff v(a) \leq v(b)$, for all $a, b \in K$. Suppose T is a complete theory of valued fields in an expansion of \mathcal{L}_{div} , and T is C -minimal, i.e., for every $\mathbf{K} \models T$, every definable subset of K is a finite Boolean combination of balls in K . Then for every partitioned \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$ there exists an integer $N \geq 0$ such that for every $b \in K^m$, the set $\varphi^{\mathbf{K}}(K; b)$ is a Boolean combination of at most N balls in K . Thus $\text{vc}^T(1) = 1$ by Example 2.12.

The definition of C -minimality used in the previous example agrees (for expansions of valued fields) with the one in [44]; this definition is slightly more restrictive than the original one, introduced in [38, 62]. Every completion of the \mathcal{L}_{div} -theory ACVF of non-trivially valued algebraically closed fields is C -minimal (essentially by A. Robinson’s quantifier elimination in ACVF; see [43]). Conversely, every valued field with C -minimal elementary theory is algebraically closed [38]. Moreover, the rigid analytic expansions of ACVF introduced by Lipshitz [57] are C -minimal [58].

Example 3.11. Let R be a ring and suppose $\mathcal{L} = \mathcal{L}_R$ is the language of R -modules. (In this paper, “ R -module” always means “left R -module.”) Suppose M is an R -module, construed as an \mathcal{L}_R -structure in the natural way. By the Baur-Monk Theorem, every \mathcal{L}_R -formula is equivalent in $T = \text{Th}(M)$ to a Boolean combination of positive primitive (p.p.) \mathcal{L}_R -formulas; given a p.p. \mathcal{L}_R -formula $\varphi(x; y)$ and $b \in M^{|y|}$, the set $\varphi(M^{|x|}; b)$ is a coset of $\varphi(M^{|x|}; 0)$. Suppose M is *p.p.-uniserial*, i.e., the subgroups of M definable by p.p. \mathcal{L}_R -formulas form a chain. By Example 2.14, if M is infinite, then we have $\text{vc}^T(1) = 1$. (In [6] this will be extended to $\text{vc}^T(m) = m$ for every m .) Examples for p.p.-uniserial abelian groups (viewed as \mathbb{Z} -modules) include $\mathbb{Q}^{(\alpha)}$, $\mathbb{Z}_{(p)}^{(\alpha)}$, $\mathbb{Z}(p^n)^{(\alpha)}$ and $\mathbb{Z}(p^\infty)^{(\alpha)}$, where p is a prime and α is a cardinal, possibly infinite. Here

$$\mathbb{Z}_{(p)} = \{a/b : a, b \in \mathbb{Z}, b \neq 0, p \nmid b\},$$

viewed as a subgroup of the additive group of \mathbb{Q} , $\mathbb{Z}(p^n)$ denotes the cyclic group $\mathbb{Z}/p^n\mathbb{Z}$ of order p^n , and $\mathbb{Z}(p^\infty)$ denotes the Prüfer p -group (the group of p^n th roots of unity, for varying n , written additively). Given an R -module M and an index set I , $M^{(I)}$ denotes, as usual, the R -submodule of the direct product M^I consisting of all sequences with cofinitely many zero entries.

Examples 3.8–3.11 may be generalized as follows:

Example 3.12. A family $\Phi(x) = \{\varphi_i(x; y_i)\}_{i \in I}$ of \mathcal{L} -formulas in the object variables x (and in various tuples of parameter variables y_i) is said to have *dual VC dimension* d if the set system $\mathcal{S} = \mathcal{S}_\Phi$ defined by the instances of the formulas φ_i has dual VC dimension d . If Φ has dual VC dimension at most 1, then we say that Φ is *VC-minimal*; cf. Example 2.3. We also say that Φ is *directed* if \mathcal{S} is directed in the sense of Example 2.13.

The \mathcal{L} -theory T is *VC-minimal* if there is a VC-minimal family of \mathcal{L} -formulas $\Phi(x)$ with $|x| = 1$ such that in every $\mathbf{M} \models T$ every definable (possibly with parameters) subset of M is a Boolean combination of finitely many sets in \mathcal{S}_Φ . (This definition was introduced in [2].) If T is a VC-minimal \mathcal{L} -theory, then for every \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$ there exists some $N \in \mathbb{N}$ such that in every $\mathbf{M} \models T$ every instance $\varphi(x; b)$ ($b \in M^{|y|}$) of φ defines a subset of M which is a Boolean combination of at most N sets in \mathcal{S}_Φ , by compactness.

One says that the VC-minimal theory T is *directed* if one can additionally choose $\Phi(x)$ to be directed; in that case we have $\text{vc}^T(1) = 1$ by Example 2.13. By [2, Proposition 6], if $\Phi(x)$ is VC-minimal and \mathcal{S}_Φ contains some \emptyset -definable set other than \emptyset or $M^{|x|}$, then there is a directed set $\Psi(x)$ of \mathcal{L} -formulas such that $\mathcal{S}_\Phi = \mathcal{S}_\Psi$ and $\mathcal{S}_{\neg\Phi} = \mathcal{S}_{\neg\Psi}$. By Lemma 3.6 this yields in fact $\text{vc}^T(1) = 1$ for every complete VC-minimal T (directed or not) without finite models.

Example 3.11 can also be generalized in a different direction:

Example 3.13. Suppose \mathcal{L} is a language expanding the language $\{1, \cdot\}$ of groups, and T is a complete \mathcal{L} -theory containing the theory of infinite groups. Suppose for every $\mathbf{G} \models T$, every definable subset of G is a Boolean combination of cosets of $\text{acl}^{\text{eq}}(\emptyset)$ -definable subgroups of G . (This condition holds, in particular, if T satisfies the model-theoretic condition known as 1-basedness, cf. [45].) By Example 2.14, if the collection of $\text{acl}^{\text{eq}}(\emptyset)$ -definable subgroups of G has breadth at most d (in particular, by Example 2.16, if it has height at most d), then we have $\text{vc}^T(1) \leq d$.

Here is a particular instantiation of the previous example:

Example 3.14. Let R be a ring, M an R -module, and $T = \text{Th}(M)$ in the language \mathcal{L}_R , as in Example 3.11. We have $M^{\aleph_0} \equiv M^{(\aleph_0)}$ (see, e.g., [42, Lemma A.1.6] or [82, Corollary 2.24]). Set $T^{\aleph_0} := \text{Th}(M^{\aleph_0}) = \text{Th}(M^{(\aleph_0)})$. It is well-known that $T = T^{\aleph_0}$ iff the class of models of T is closed under direct products, iff for all p.p. \mathcal{L}_R -formulas $\varphi(x)$, $\psi(x)$, either $\varphi(M^{|x|}) \subseteq \psi(M^{|x|})$ or the index

$$\text{Inv}(M, \varphi, \psi) := [\varphi(M^{|x|}) : (\varphi \wedge \psi)(M^{|x|})]$$

is infinite. (See, e.g., [42, Lemma A.1.7].) So if $T = T^{\aleph_0}$ and the Morley rank $\text{MR}(T)$ of T is finite then the length n of every sequence

$$M \supseteq \varphi_1(M) \supseteq \varphi_1(M) \cap \varphi_2(M) \supseteq \cdots \supseteq \varphi_1(M) \cap \cdots \cap \varphi_n(M),$$

where each $\varphi_i(x)$ is a p.p. \mathcal{L}_R -formula with $|x| = 1$, is bounded by $d = \text{MR}(T)$; so by Examples 2.14 and 2.16 we see that $\text{vc}^T(1) \leq d$. (Note that this bound is far from optimal: e.g., for $R = \mathbb{Z}$, $M = \mathbb{Z}(p^d)^{(\aleph_0)}$ we have $\text{MR}(T) = d$, yet $\text{vc}^T(1) = 1$ by Example 3.11.) In [6] we will extend this to $\text{vc}^T(m) \leq md$ for every m .

3.4. Dual VC density of sets of formulas. It is convenient to extend the definition of dual VC density to finite sets of formulas. Let $\Delta = \Delta(x; y)$ be a finite set of partitioned \mathcal{L} -formulas $\varphi = \varphi(x; y)$ with the object variables x and parameter variables y . We set $\neg\Delta := \{\neg\varphi : \varphi \in \Delta\}$, and for $B \subseteq M^{|y|}$ we let

$$\Delta(x; B) := \{\varphi(x; b) : \varphi \in \Delta, b \in B\}.$$

Given a finite set $B \subseteq M^{|y|}$, we call a consistent subset of $\Delta(x; B) \cup \neg\Delta(x; B)$ a $\Delta(x; B)$ -type. Note that our parameter sets are subsets of $M^{|y|}$, and not of M , as is more common in model theory. (This is simply a matter of convenience, in order to be compatible with VC duality.) Given a $\Delta(x; B)$ -type p we denote by $p^M \subseteq M^{|x|}$ its set of realizations in \mathbf{M} . Since we are only dealing with finite sets Δ and finite parameter sets $B \subseteq M^{|y|}$, all $\Delta(x; B)$ -types have realizations in \mathbf{M} itself (rather than in an elementary extension). Given another finite set $\Delta'(x; y')$ of partitioned \mathcal{L} -formulas and a finite $B' \subseteq M^{|y'|}$, we say that a $\Delta(x; B)$ -type p is equivalent to a $\Delta'(x; B')$ -type q if $p^M = q^M$.

Let now $B \subseteq M^{|y|}$ be finite. Given $a \in M^{|x|}$ we denote the $\Delta(x; B)$ -type of a by

$$\begin{aligned} \text{tp}^\Delta(a/B) := & \{ \varphi(x; b) : b \in B, \varphi \in \Delta, \mathbf{M} \models \varphi(a; b) \} \cup \\ & \{ \neg\varphi(x; b) : b \in B, \varphi \in \Delta, \mathbf{M} \not\models \varphi(a; b) \}. \end{aligned}$$

We write $S^\Delta(B)$ for the set of complete $\Delta(x; B)$ -types (in \mathbf{M}), that is, the set of (in \mathbf{M}) maximally consistent subsets of $\Delta(x; B) \cup \neg\Delta(x; B)$; equivalently,

$$S^\Delta(B) = \{ \text{tp}^\Delta(a/B) : a \in M^{|x|} \}.$$

If $\Delta = \{\varphi\}$ is a singleton, we also write $S^\varphi(B)$ instead of $S^\Delta(B)$. The elements of $S^\Delta(B)$ are syntactical objects (sets of formulas), but associating to a type $p \in S^\Delta(B)$ its set $p^{\mathbf{M}}$ of realizations in \mathbf{M} gives a bijection from $S^\Delta(B)$ onto the set

$$S(\varphi^{\mathbf{M}}(M^{|x|}; b) : b \in B, \varphi \in \Delta)$$

of atoms of the Boolean algebra generated by the subsets $\varphi^{\mathbf{M}}(M^{|x|}; b)$ of $M^{|x|}$. (See Section 2.2.) Hence for every partitioned \mathcal{L} -formula $\varphi(x; y)$ we have

$$\pi_\varphi^*(t) = \max \{ |S^\varphi(B)| : B \subseteq M^{|y|}, |B| = t \}.$$

In the general case, for every $t \in \mathbb{N}$ we also set

$$\pi_\Delta^*(t) := \max \{ |S^\Delta(B)| : B \subseteq M^{|y|}, |B| = t \},$$

so $0 \leq \pi_\Delta^*(t) \leq 2^{|\Delta|t}$. Similarly as in Lemma 3.2 one shows that if we pass from \mathbf{M} to an elementarily equivalent \mathcal{L} -structure then π_Δ^* does not change (justifying our notation, which suppresses \mathbf{M}).

Let $\Delta_0(x; y)$ be a finite set of partitioned \mathcal{L} -formulas with $\Delta_0 \subseteq \Delta$, and $B \subseteq M^{|y|}$ be finite. Then there is a natural restriction map $S^\Delta(B) \rightarrow S^{\Delta_0}(B)$, written as $p \mapsto p \upharpoonright \Delta_0$. This map is onto: given $p \in S^{\Delta_0}(B)$ let $a \in p^{\mathbf{M}}$ be arbitrary; then $q := \text{tp}^\Delta(a/B) \in S^\Delta(B)$ satisfies $q \upharpoonright \Delta_0 = p$. In particular, $|S^{\Delta_0}(B)| \leq |S^\Delta(B)|$. Note also that if $\Delta \neq \emptyset$, then the restriction maps $p \mapsto p \upharpoonright \varphi$, where $\varphi \in \Delta$, combine to an injective map $S^\Delta(B) \rightarrow \prod_{\varphi \in \Delta} S^\varphi(B)$; in particular, $|S^\Delta(B)| \leq \prod_{\varphi \in \Delta} |S^\varphi(B)|$. This shows:

Lemma 3.15. *If all $\varphi \in \Delta$ are dependent, then there exists a real number r with $0 \leq r \leq \sum_{\varphi \in \Delta} \text{vc}^*(\varphi)$ and*

$$|S^\Delta(B)| = O(|B|^r) \quad \text{for all finite } B \subseteq M^{|y|}. \quad (3.1)$$

We define the *dual VC density* of Δ as the infimum $\text{vc}^*(\Delta)$ of all real numbers $r \geq 0$ such that (3.1) holds; that is,

$$\text{vc}^*(\Delta) = \inf \{ r \geq 0 : \pi_\Delta^*(t) = O(t^r) \}.$$

We have

$$\max_{\varphi \in \Delta} \text{vc}^*(\varphi) \leq \text{vc}^*(\Delta) \leq \sum_{\varphi \in \Delta} \text{vc}^*(\varphi).$$

Clearly $\text{vc}^*(\Delta)$ agrees with $\text{vc}^*(\varphi)$ as defined previously if $\Delta = \{\varphi\}$ is a singleton. Moreover, $\text{vc}^*(\Delta) = 0$ iff $\text{vc}^*(\varphi) = 0$ for every $\varphi \in \Delta$, and if $\text{vc}^*(\Delta) < 1$ then $\text{vc}^*(\Delta) = 0$. (See the remarks following Lemma 2.2.) Note that in computing $\text{vc}^*(\Delta)$ there is no harm in assuming that Δ is *closed under negation*, i.e., with every $\varphi \in \Delta$ the set Δ also contains a formula equivalent (in \mathbf{M}) to $\neg\varphi$. (Passing from Δ to $\Delta \cup \neg\Delta$ does not change $S^\Delta(B)$.)

Example. Suppose $\Delta(x; y) = \{x_1 = y, \dots, x_m = y\}$ where $|x| = m$ and $|y| = 1$. Then for finite $B \subseteq M$ we have $|S^\Delta(B)| = (|B| + 1)^m$, hence $\text{vc}^*(\Delta) = \sum_{\varphi \in \Delta} \text{vc}^*(\varphi) = m$.

We finish this subsection with an easy result about interpretations (related to Lemma 3.3 and Corollary 3.5).

Lemma 3.16. *Let M' be an infinite structure in a language \mathcal{L}' and $\pi: X \rightarrow M'$ an interpretation of M' in M without parameters, where $X \subseteq M^r$ is \emptyset -definable. Then for any finite set $\Delta'(x; y)$ of \mathcal{L}' -formulas there exists a finite set $\Delta(\bar{x}; \bar{y})$ of \mathcal{L} -formulas such that $|\Delta| = |\Delta'|$, $|\bar{x}| = r|x|$, and $\pi_{\Delta'}^* \leq \pi_{\Delta}^*$.*

Proof. Let $m := |x|$ and $n := |y|$. Let $B' \subseteq (M')^n$ be finite. Choose $B \subseteq X^n$ with $|B| = |B'|$ such that each $b = (b_1, \dots, b_n) \in B'$ has the form $(\pi(\bar{b}_1), \dots, \pi(\bar{b}_n))$ for some $(\bar{b}_1, \dots, \bar{b}_n) \in B$. For each \mathcal{L}' -formula $\varphi(x; y)$ choose an \mathcal{L} -formula $\psi_{\varphi}(\bar{x}; \bar{y})$, where $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m)$, $\bar{y} = (\bar{y}_1, \dots, \bar{y}_n)$ and $|\bar{x}_1| = \dots = |\bar{x}_m| = |\bar{y}_1| = \dots = |\bar{y}_n| = r$, such that $\psi_{\varphi}(M^{(m+n)r}) \subseteq X^{m+n}$ and for any $\bar{a}_1, \dots, \bar{a}_m, \bar{b}_1, \dots, \bar{b}_n \in X$,

$$M \models \psi_{\varphi}(\bar{a}_1, \dots, \bar{a}_m; \bar{b}_1, \dots, \bar{b}_n) \iff M' \models \varphi(\pi(\bar{a}_1), \dots, \pi(\bar{a}_m); \pi(\bar{b}_1), \dots, \pi(\bar{b}_n)).$$

Let a finite set $\Delta'(x; y)$ of \mathcal{L}' -formulas be given. Set $\Delta := \{\psi_{\varphi} : \varphi \in \Delta'\}$. Then $S^{\Delta}(B) \subseteq X^m$, and $(\bar{a}_1, \dots, \bar{a}_m) \mapsto (\pi(\bar{a}_1), \dots, \pi(\bar{a}_m))$ yields a surjective map $S^{\Delta}(B) \rightarrow S^{\Delta'}(B')$, hence $|S^{\Delta'}(B')| \leq |S^{\Delta}(B)|$ as required. \square

By Lemmas 3.6 and 3.16:

Corollary 3.17. *Let M' be an infinite structure in a language \mathcal{L}' , interpretable in M (possibly with parameters) on a definable subset of M^r . Then, writing $T = \text{Th}(M)$ and $T' = \text{Th}(M')$, we have $\text{vc}^{T'}(m) \leq \text{vc}^T(rm)$ for every m .*

So for example if G is a group (considered as a structure in the usual first-order language of group theory) and H is a definable normal subgroup of G , then $\text{vc}^{\text{Th}(G/H)} \leq \text{vc}^{\text{Th}(G)}$ if H has infinite index in G , and $\text{vc}^{\text{Th}(H)} \leq \text{vc}^{\text{Th}(G)}$ if H is infinite.

3.5. Coding finite sets of formulas. We let \mathcal{L} , M and Δ be as in the previous subsection, and $T = \text{Th}(M)$. The following useful lemma, essentially due to Shelah [88, Lemma II.2.1], shows that counting $\Delta(x; B)$ -types where $|\Delta| > 1$ is not really more general than counting $\Delta(x; B)$ -types where Δ is a singleton:

Lemma 3.18. *Let $d = |\Delta|$ and $y' = (y_1, \dots, y_{2d}, z, z_1, \dots, z_{2d})$ with $|y| = |y_i| = |z_i| = |z|$ for every $i = 1, \dots, 2d$. There is an \mathcal{L} -formula $\psi_{\Delta}(x; y')$ with the following properties:*

- (1) *for every finite $B \subseteq M^{|y|}$ with $|B| \geq 2$ there is some $B' \subseteq M^{|y'|}$ with $|B'| = 2d|B|$ such that every $p \in S^{\Delta}(B)$ is equivalent to some $q \in S^{\psi_{\Delta}}(B')$;*
- (2) *for every finite $B' \subseteq M^{|y'|}$ there is some $B \subseteq M^{|y|}$ with $|B| \leq 2d|B'|$ such that every $q \in S^{\psi_{\Delta}}(B')$ is equivalent to some (possibly incomplete) $\Delta(x; B)$ -type p_0 .*

In particular, we have $\pi_{\Delta}^(t) \leq \pi_{\psi_{\Delta}}^*(2dt)$ for $t > 1$ and $\pi_{\psi_{\Delta}}^*(t) \leq \pi_{\Delta}^*(2dt)$ for $t \geq 0$. Thus $\text{vc}^*(\Delta) = \text{vc}^*(\psi_{\Delta}) \leq \text{vc}^T(m)$ where $m = |x|$.*

Proof. Write $\Delta = \{\varphi_1, \dots, \varphi_d\}$ and define ψ_{Δ} as follows:

$$\begin{aligned} \psi_{\Delta} = \bigwedge_{k=1}^d (z = z_k \rightarrow \varphi_k(x; y_k)) \wedge \bigwedge_{k=d+1}^{2d} (z = z_k \rightarrow \neg \varphi_{k-d}(x; y_k)) \wedge \\ \bigvee_{k=1}^{2d} z = z_k \wedge \bigwedge_{1 \leq k < l \leq 2d} \neg(z = z_k \wedge z = z_l). \end{aligned}$$

For (1), suppose $B \subseteq M^{|y|}$ is finite, and $b_0 \neq b_1$ are distinct elements of B . For $b \in B$ and $k \in [d]$ set

$$b_0^{(k)} := (b_0, \quad b_0, \quad \dots, \quad b, \quad \dots, \quad b_0, \quad b_1, \quad b_0, \quad \dots, \quad b_1, \quad \dots, \quad b_0)$$

$$y_1 \quad y_2 \quad \dots \quad y_{d+k} \quad \dots \quad y_{2d} \quad z \quad z_1 \quad \dots \quad z_{d+k} \quad \dots \quad z_{2d}$$

and

$$b_1^{(k)} := (b_0, \quad b_0, \quad \dots, \quad b, \quad \dots, \quad b_0, \quad b_1, \quad b_0, \quad \dots, \quad b_1, \quad \dots, \quad b_0)$$

$$y_1 \quad y_2 \quad \dots \quad y_k \quad \dots \quad y_{2d} \quad z \quad z_1 \quad \dots \quad z_k \quad \dots \quad z_{2d}$$

and put

$$B' := \{b_0^{(k)}, b_1^{(k)} : b \in B, k \in [d]\} \subseteq (M^{|y|})^{4d+1}.$$

Then $|B'| = 2d|B|$, and for every $b \in B$, $k \in [d]$ we have

$$\psi_\Delta(M^{|x|}; b_0^{(k)}) = \neg\varphi_k(M^{|x|}; b), \quad \psi_\Delta(M^{|x|}; b_1^{(k)}) = \varphi_k(M^{|x|}; b).$$

Given $p \in S^\Delta(B)$ we set

$$q := \{\neg\psi_\Delta(x; b_0^{(k)}), \psi_\Delta(x; b_1^{(k)}) : \varphi_k(x; b) \in p\} \cup$$

$$\{\psi_\Delta(x; b_0^{(k)}), \neg\psi_\Delta(x; b_1^{(k)}) : \varphi_k(x; b) \notin p\}.$$

Then clearly $q \in S^{\psi_\Delta}(B')$, and q is equivalent to p . The map $p \mapsto q : S^\Delta(B) \rightarrow S^{\psi_\Delta}(B')$ is injective, hence $|S^\Delta(B)| \leq |S^{\psi_\Delta}(B')| \leq \pi_{\psi_\Delta}^*(2d|B|)$.

For (2) note that if $b_1, \dots, b_{2d}, c, c_1, \dots, c_{2d} \in M^{|y|}$ then the formula

$$\psi_\Delta(x; b_1, \dots, b_{2d}, c, c_1, \dots, c_{2d})$$

defines $\varphi_k(M^{|x|}; b_k)$, $\neg\varphi_k(M^{|x|}; b_{k+d})$, or \emptyset (since the c_i 's are not necessarily distinct). Let $B' \subseteq M^{|y'|}$ be finite, and $q \in S^\Delta(B')$. Set

$$B := \{b \in M^{|y'|} : b = b_i \text{ for some } (b_1, \dots, b_{2d}, c, c_1, \dots, c_{2d}) \in B'\}$$

and let p_0 be the set of formulas which have the form $\varphi_k(x; b)$ where $k \in [d]$, $b = b_k$ for some $\psi_\Delta(x; b_1, \dots, b_{2d}, c, c_1, \dots, c_{2d}) \in q$ with $c = c_k$, or the form $\neg\varphi_k(x; b)$ with $k \in [d]$, $b = b_{d+k}$ for some $\psi_\Delta(x; b_1, \dots, b_{2d}, c, c_1, \dots, c_{2d}) \in q$ with $c = c_{k+d}$. Then $|B| \leq 2d|B'|$, and p_0 is a $\Delta(x; B)$ -type equivalent to q . For each q choose an extension p of p_0 to a complete $\Delta(x; B)$ -type. Then the map $q \mapsto p : S^{\psi_\Delta}(B') \rightarrow S^\Delta(B)$ is injective, so $|S^{\psi_\Delta}(B')| \leq |S^\Delta(B)| \leq \pi_\Delta^*(2d|B'|)$. \square

In the rest of this subsection we give some applications of this lemma. We first note:

Corollary 3.19. *Let Φ be a set of \mathcal{L} -formulas with the tuple of object variables x and varying parameter variables such that every \mathcal{L} -formula $\varphi(x; y)$ is equivalent in T to a Boolean combination of formulas in Φ . Then*

$$\text{vc}^T(m) = \sup \{ \text{vc}^*(\Delta) : \Delta \subseteq \Phi \text{ finite} \} \quad \text{where } m = |x|.$$

Proof. The inequality “ \leq ” is a consequence of the hypothesis: for each \mathcal{L} -formula $\varphi(x; y)$ there is a finite subset $\Delta = \Delta(x; y)$ of Φ such that $|S^\varphi(B)| \leq |S^\Delta(B)|$ for each finite $B \subseteq M^{|y|}$. The reverse inequality follows from the previous lemma. \square

Let $\mathbf{M}^* \succ \mathbf{M}$ be a monster model of T . Consider the expansion \mathcal{L}^{Sh} of \mathcal{L} by a new predicate symbol $R_{\psi, c}(x)$ for every \mathcal{L} -formula $\psi(x; z)$ and every $c \in (M^*)^{|z|}$. The *Shelah expansion* of \mathbf{M} is the expansion of \mathbf{M} to an \mathcal{L}^{Sh} -structure \mathbf{M}^{Sh} where each predicate symbol $R_{\psi, c}(x)$ as before is interpreted by $M^{|x|} \cap \psi^{\mathbf{M}^*}((M^*)^{|x|}; c)$. Shelah showed [89] (with another proof given in [20]) that if T is NIP then $T^{\text{Sh}} = \text{Th}(\mathbf{M}^{\text{Sh}})$ admits

quantifier elimination and is also NIP. This provides an interesting way of constructing new NIP theories from old ones. The previous lemma and its Corollary 3.19 allows us to prove that T and T^{Sh} share the same VC density function:

Corollary 3.20. $\text{vc}^{T^{\text{Sh}}} = \text{vc}^T$.

Proof. Fix some m and assume $|x| = m$. The inequality $\text{vc}^{T^{\text{Sh}}}(m) \geq \text{vc}^T(m)$ being obvious, we only need to show that $\text{vc}^{T^{\text{Sh}}}(m) \leq \text{vc}^T(m)$. Let $\Delta = \Delta(x; y)$ be a finite set of atomic \mathcal{L}^{Sh} -formulas; by Corollary 3.19 and Shelah's theorem mentioned above, it suffices to show that $\text{vc}^*(\Delta) \leq \text{vc}^T(m)$. Take a finite set $\Psi = \Psi(x; y, z)$ of partitioned \mathcal{L} -formulas and some $c \in M^{|z|}$ such that $\Delta = \{R_{\psi, c}(x; y) : \psi \in \Psi\}$. Let $B \subseteq M^{|y|}$ be finite, $B^* := B \times \{c\}$, and let $p \in S^\Delta(B)$. Let a be an arbitrary realization of p (in \mathbf{M}^{Sh}), and define $p^* := \text{tp}^\Psi(a/B^*)$ (in \mathbf{M}^*). Then for $\psi \in \Psi$ and $b \in B$ we have

$$\begin{aligned} \psi(x; b, c) \in p^* &\iff \mathbf{M}^* \models \psi(a; b, c) \\ &\iff \mathbf{M}^{\text{Sh}} \models R_{\psi, c}(a; b) \\ &\iff R_{\psi, c}(x; b) \in p. \end{aligned}$$

In particular, the map $p \mapsto p^* : S^\Delta(B) \rightarrow S^\Psi(B^*)$ is injective, so $\text{vc}^*(\Delta) \leq \text{vc}^*(\Psi) \leq \text{vc}^T(m)$ by Lemma 3.18. \square

It is well-known (see, e.g., [100, Theorem 4.7]) that the direct product of two NIP structures is again NIP. As a consequence of the last lemma we can also now estimate the VC density of a direct product in terms of the VC densities of its factors. We refer to [42, Section 9.1] for the definition of the product of two \mathcal{L} -structures, and to [42, Corollary 9.6.4] for the Feferman-Vaught Theorem used in the proof below.

Lemma 3.21. *Let \mathbf{M}' be another infinite \mathcal{L} -structure, $T' = \text{Th}(\mathbf{M}')$, and let $T^\times = \text{Th}(\mathbf{M} \times \mathbf{M}')$ be the \mathcal{L} -theory of the direct product of \mathbf{M} and \mathbf{M}' . Then*

$$\text{vc}^{T^\times} \leq \text{vc}^T + \text{vc}^{T'}.$$

Proof. Given n -tuples $a = (a_1, \dots, a_n) \in M^n$ and $a' = (a'_1, \dots, a'_n) \in (M')^n$ we denote by $a \times a'$ the n -tuple $((a_1, a'_1), \dots, (a_n, a'_n))$ of elements of $M \times M'$; every element of $(M \times M')^n$ has the form $a \times a'$ for some $a \in M^n$, $a' \in (M')^n$.

Let $\varphi(x; y)$ be an \mathcal{L} -formula. By the Feferman-Vaught Theorem there exist finitely many pairs of \mathcal{L} -formulas $(\theta_i(x; y), \theta'_i(x; y))$, $i \in [n] = \{1, \dots, n\}$, such that for all $a \in M^{|x|}$, $a' \in (M')^{|x|}$ and $b \in M^{|y|}$, $b' \in (M')^{|y|}$,

$$\mathbf{M} \times \mathbf{M}' \models \varphi(a \times a'; b \times b') \iff \text{for some } i \in [n], \mathbf{M} \models \theta_i(a; b) \text{ and } \mathbf{M}' \models \theta'_i(a'; b').$$

Set $\Theta := \{\theta_1, \dots, \theta_n\}$, $\Theta' := \{\theta'_1, \dots, \theta'_n\}$. Let C be a finite set of tuples from $(M \times M')^{|y|}$. Take $B \subseteq M^{|y|}$, $B' \subseteq (M')^{|y|}$ with $|B|, |B'| \leq |C|$ such that each $c \in C$ is of the form $c = b \times b'$ for a unique pair $(b, b') \in B \times B'$. For every $p \in S^\varphi(C)$ choose a realization $a_p \times a'_p \in (M \times M')^{|x|}$ of p in $\mathbf{M} \times \mathbf{M}'$, and put

$$q := \text{tp}^\Theta(a_p/B), \quad q' := \text{tp}^{\Theta'}(a'_p/B').$$

Then for all $(b, b') \in B \times B'$ we have

$$\begin{aligned} \varphi(x; b \times b') \in p &\iff \mathbf{M} \times \mathbf{M}' \models \varphi(a_p \times a'_p; b \times b') \\ &\iff \mathbf{M} \models \theta_i(a_p; b) \text{ and } \mathbf{M}' \models \theta'_i(a'_p; b'), \text{ for some } i \in [n] \\ &\iff \theta_i(x; b) \in q \text{ and } \theta'_i(x; b') \in q', \text{ for some } i \in [n]. \end{aligned}$$

Hence the map $p \mapsto (q, q')$ is an injection $S^\varphi(C) \rightarrow S^\Theta(B) \times S^{\Theta'}(B')$. In particular we obtain $\pi_\varphi^*(t) \leq \pi_\Theta^*(t) \cdot \pi_{\Theta'}^*(t)$ for every t and hence $\text{vc}^*(\varphi) \leq \text{vc}^*(\Theta) + \text{vc}^*(\Theta')$; here π_φ^* is computed in $\mathbf{M} \times \mathbf{M}'$ and $\pi_\Theta^*, \pi_{\Theta'}^*$ in \mathbf{M} and \mathbf{M}' , respectively, and similarly for vc^* . By Lemma 3.18 therefore $\text{vc}^{T^\times}(m) \leq \text{vc}^T(m) + \text{vc}^{T'}(m)$ where $m = |x|$. \square

Remark. In a similar way one shows that if \mathbf{M}' is a finite \mathcal{L} -structure and $T^\times = \text{Th}(\mathbf{M} \times \mathbf{M}')$, then $\text{vc}^{T^\times} \leq \text{vc}^T$.

We finish this subsection by noting a further restriction on the growth of vc (cf. also Lemma 3.7):

Lemma 3.22. $d \text{vc}(m) \leq \text{vc}(dm)$ for all $d, m > 0$.

Proof. Let $\Delta(x; y)$ be a finite set of \mathcal{L} -formulas with $|x| = m$. Let x_1, \dots, x_d be new m -tuples of variables and set

$$\Delta'(x_1, \dots, x_d; y) := \{\varphi(x_i; y) : \varphi(x; y) \in \Delta, i = 1, \dots, d\}.$$

Let $B \subseteq M^{|y|}$, $|B| = t \in \mathbb{N}$, such that $r := \pi_\Delta^*(t) = |S^\Delta(B)|$. Let $a_1, \dots, a_r \in M^m$ be realizations of the types in $S^\Delta(B)$. For each $\mathbf{i} = (i_1, \dots, i_d) \in [r]^d$ let $\mathbf{a}_i := (a_{i_1}, \dots, a_{i_d}) \in (M^m)^d = M^{dm}$. Then the \mathbf{a}_i realize pairwise distinct $\Delta'(x_1, \dots, x_d; B)$ -types. This yields $(\pi_\Delta^*(t))^d = |S^\Delta(B)|^d \leq |S^{\Delta'}(B)| \leq \pi_{\Delta'}^*(t)$. Since t was arbitrary, we obtain $d \text{vc}^*(\Delta) \leq \text{vc}^*(\Delta')$. Hence $d \text{vc}(m) \leq \text{vc}(dm)$ by Lemma 3.18. \square

3.6. VC density and indiscernible sequences. In this subsection we assume that \mathbf{M} is sufficiently saturated. Recall that $\pi_\varphi(t)$ is the maximum size of $\mathcal{S}_\varphi \cap A$ as A ranges over t -element subsets of M^m , and $\pi_\varphi^*(t)$ is the maximum size of $S^\varphi(B)$ as B ranges over all t -element subsets of M^n ; here, as above $m = |x|$, $n = |y|$. These definitions may naturally be relativized to parameters coming from indiscernible sequences. More precisely:

Definition 3.23. For every t let $\pi_{\varphi, \text{ind}}(t)$ be the maximum of $|\mathcal{S}_\varphi \cap A|$ as A ranges over all sets of the form $A = \{a_0, \dots, a_{t-1}\}$ for some indiscernible sequence $(a_i)_{i \in \mathbb{N}}$ in M^m , and let $\pi_{\varphi, \text{ind}}^*(t)$ be the maximum of $|S^\varphi(B)|$ where $B = \{b_0, \dots, b_{t-1}\}$ for some indiscernible sequence $(b_i)_{i \in \mathbb{N}}$ in M^n . We call $\pi_{\varphi, \text{ind}}$ the *indiscernible shatter function* of φ and $\pi_{\varphi, \text{ind}}^*$ the *dual indiscernible shatter function* of φ .

The indiscernible shatter functions give rise to corresponding notions of *indiscernible VC dimension* $\text{VC}_{\text{ind}}(\varphi)$ and *indiscernible VC density* $\text{vc}_{\text{ind}}(\varphi)$ of φ (and their duals $\text{VC}_{\text{ind}}^*(\varphi)$ and $\text{vc}_{\text{ind}}^*(\varphi)$) in a natural way; for example, $\text{vc}_{\text{ind}}^*(\varphi)$ is the infimum of all $r > 0$ having the property that there is some $C > 0$ such that for all t and indiscernible sequences $(b_i)_{i \in \mathbb{N}}$ we have $|S^\varphi(B)| \leq Ct^r$, where $B = \{b_0, \dots, b_{t-1}\}$; if there is no such r then $\text{vc}_{\text{ind}}^*(\varphi) = \infty$.

As in the classical case (cf. Lemma 2.4) we see that $\pi_{\varphi, \text{ind}}^* = \pi_{\varphi^*, \text{ind}}$ and hence $\text{VC}_{\text{ind}}(\varphi^*) = \text{VC}_{\text{ind}}^*(\varphi)$ and $\text{vc}_{\text{ind}}(\varphi^*) = \text{vc}_{\text{ind}}^*(\varphi)$. Directly from the definition we have $\pi_{\varphi, \text{ind}} \leq \pi_\varphi$ and hence $\text{VC}_{\text{ind}}(\varphi) \leq \text{VC}(\varphi)$ and $\text{vc}_{\text{ind}}(\varphi) \leq \text{vc}(\varphi)$. In particular $\text{VC}_{\text{ind}}(\varphi)$ and $\text{vc}_{\text{ind}}(\varphi)$ are finite if φ defines a VC class. Conversely, if $\text{VC}_{\text{ind}}(\varphi)$ is finite, then so is $\text{VC}(\varphi)$. (This follows by saturation of \mathbf{M} and extraction of an indiscernible sequence; see proof of Proposition 4 in [3].) Hence if one of the quantities $\text{VC}(\varphi)$, $\text{vc}(\varphi)$, $\text{VC}_{\text{ind}}(\varphi)$, or $\text{vc}_{\text{ind}}(\varphi)$ is finite, then so are all the others.

Another numerical parameter associated to φ and defined via indiscernible sequences is the *alternation number* $\text{alt}(\varphi)$ of φ (in \mathbf{M}). This is the largest d (if it exists) such

that for some indiscernible sequence $(a_i)_{i \in \mathbb{N}}$ in M^m and some $b \in M^n$ we have

$$a_i \in \varphi(M^m; b) \iff a_{i+1} \notin \varphi(M^m; b) \quad \text{for all } i < d - 1.$$

If there is no such d we set $\text{alt}(\varphi) = \infty$. It is well-known (and essentially due to Poizat) that $\text{alt}(\varphi) \leq 2 \text{VC}_{\text{ind}}(\varphi) + 1$ (see, e.g., [3, Proposition 3]) and that if $\text{alt}(\varphi)$ is finite then φ defines a VC class [3, Proposition 4]. Moreover:

Lemma 3.24. $\text{vc}_{\text{ind}}(\varphi) \leq \text{alt}(\varphi) - 1$.

Proof. Since this is trivial if φ has infinite alternation number, we assume that $d := \text{alt}(\varphi) < \infty$. Let $(a_i)_{i \in \mathbb{N}}$ be an indiscernible sequence in M^m and $A = \{a_0, \dots, a_{t-1}\}$. Then for each $b \in M^n$, there are less than d indices $i < t - 1$ such that $\varphi(a_i; b)$ and $\varphi(a_{i+1}; b)$ have different truth value in \mathbf{M} , and the set $A \cap \varphi(M^m; b)$ is uniquely determined by knowledge of these indices i . Thus $|A \cap \mathcal{S}_\varphi| \leq 2 \sum_{i=0}^{d-1} \binom{t}{i} = O(t^{d-1})$ and hence $\text{vc}_{\text{ind}}(\varphi) \leq d - 1$ as required. \square

Example. Suppose $\mathcal{S}_\varphi \subseteq \binom{M^m}{d}$ where $d > 0$. Then $\text{alt}(\varphi) \leq 2d + 1$ and $\text{vc}_{\text{ind}}(\varphi) \leq \text{vc}(\varphi) \leq d$, and all these inequalities are equalities if $\mathcal{S}_\varphi = \binom{M^m}{d}$.

The previous example shows that the inequality in Lemma 3.24, in general, is strict. The inequality $\text{VC}_{\text{ind}}(\varphi) \leq \text{VC}(\varphi)$ may be strict if there are no non-trivial indiscernible sequences:

Example. Suppose $\mathcal{L} = \{A, S, P\}$ where A and S are unary relation symbols and P is a binary relation symbol, and suppose \mathbf{M} is an \mathcal{L} -structure, with the interpretations of A , S and P in \mathbf{M} denoted by the same symbols, such that

- (1) $|A| = d$ and $|S| = 2^d$;
- (2) for $s \in S$, $P(x, s)$ defines a subset of A so that when s runs through S we obtain all subsets of A ;
- (3) for $s \notin S$, $P(x, s)$ defines the empty set.

Then $\text{VC}(P) = d$ and $\text{VC}_{\text{ind}}(P) = 1$ (as well as $\text{vc}(P) = \text{vc}_{\text{ind}}(P) = 0$).

The inequality $\text{vc}_{\text{ind}}(\varphi) \leq \text{vc}(\varphi)$ may also be strict, as Lemma 4.8 in the next section shows. We do not know the answer to the following question:

Question. Is $\text{vc}_{\text{ind}}(\varphi)$ always integral-valued?

(After a first version of this manuscript had been completed, Guingona and Hill [35] showed that this question indeed has a positive answer.)

We finish this section with a connection between vc_{ind}^* and the Helly number. We already remarked (see Section 2.4) that if $\mathbf{M} = (M, <)$ is a dense linearly ordered set and $\varphi(x; y_1, z_1, y_2, z_2) = (y_1 < x < z_1 \vee y_2 < x < z_2)$ then the set system \mathcal{S}_φ has infinite Helly number: that is, for each d there is a finite subfamily of \mathcal{S}_φ which is d -consistent yet inconsistent. In contrast to this, we have:

Lemma 3.25. *Put $d = \lfloor \text{vc}_{\text{ind}}^*(\varphi) \rfloor + 1$. Then for every indiscernible sequence $(b_i)_{i \in \mathbb{N}}$ in $M^{|y|}$ the set system $\mathcal{S} = \{\varphi(M^m; b_i) : i \in \mathbb{N}\}$ has Helly number at most d .*

Proof. Suppose for a contradiction that $(b_i)_{i \in \mathbb{N}}$ is an indiscernible sequence such that $\mathcal{S} = \{\varphi(M^m; b_i) : i \in \mathbb{N}\}$ has Helly number larger than d . Then some finite subfamily \mathcal{S}_0 of \mathcal{S} is d -consistent but not consistent. By indiscernibility of (b_i) , every finite subfamily of \mathcal{S} of size at least $|\mathcal{S}_0|$ has this property. In particular, we can take $D \in \mathbb{N}$ maximal such that the set $\{\varphi(M^m; b_i) : i < D\}$ is consistent. Obviously $D \geq d$. Since (b_i) is

indiscernible, we obtain that for any $I_0 \in \binom{\mathbb{N}}{D}$ the set $\{\varphi(M^m; b_i) : i \in I_0\}$ is consistent, but for any $D' > D$ and any $I_1 \in \binom{\mathbb{N}}{D'}$ the set $\{\varphi(M^m; b_i) : i \in I_1\}$ is inconsistent. Let $t > D$ be arbitrary, and set $B_t = \{b_i : i < t\}$. For $I \in \binom{t}{D}$ let $q_I(x)$ be the unique φ -type over B_t with $\varphi(x; b_i) \in q_I$ for $i \in I$ and $\neg\varphi(x; b_i) \in q_I$ for $i \notin I$. Since $|I| = D$ every q_I is consistent. Thus $|S^\varphi(B_t)| \geq \binom{t}{D} = \Theta(t^D)$. Since $D \geq d$, this contradicts $\text{vc}_{\text{ind}}^*(\varphi) < d$. \square

Remark. Note that in the context of the previous lemma, we cannot achieve the stronger conclusion that \mathcal{S} has breadth at most d : for the formula $\varphi(x; y) = x \neq y$ and any indiscernible sequence (b_i) , the set system \mathcal{S} always has infinite breadth.

By Lemma 3.25 and extraction of an indiscernible sequence (using that \mathbf{M} is assumed to be sufficiently saturated) we obtain a consequence which does not mention indiscernibles:

Corollary 3.26. *Suppose the set system \mathcal{S}_φ is d -consistent, where $d = \lfloor \text{vc}^*(\varphi) \rfloor + 1$. Then there is an infinite subset of \mathcal{S}_φ which is consistent.*

This is a weak version of a theorem of Matoušek [67], according to which, if \mathcal{S}_φ is d -consistent, where $d > \text{vc}^*(\varphi)$, then one may write $\mathcal{S}_\varphi = \mathcal{S}_1 \cup \dots \cup \mathcal{S}_N$ (for some $N \in \mathbb{N}$) where each \mathcal{S}_i is consistent.

4. SOME VC DENSITY CALCULATIONS

In this section we give an example of a formula in the language of rings which, in every infinite field, defines a set system with fractional VC density, depending on the characteristic of the field. The construction of this formula (which is inspired by an example by Assouad [7], who in turn credits Frankl) proceeds in two steps: we first associate to a given partitioned formula φ a bigraph (= bipartite graph with a fixed ordering of the bipartition of the vertex set), and then we realize the set of edges of this bigraph as a definable family \mathcal{S}_φ . For our example we choose φ so as to encode point-line incidences in the affine plane; the calculation of $\text{vc}(\widehat{\varphi})$ in characteristic zero uses an analogue of the Szémeredi-Trotter Theorem due to Tóth. We also discuss the question whether VC density in NIP theories can take irrational values, and give examples of formulas in NIP theories whose shatter function is not asymptotic to a real power function.

Throughout this section \mathcal{L} is a first-order language and \mathbf{M} is an \mathcal{L} -structure.

4.1. Associating a bigraph to a partitioned formula. We follow [59] and make a distinction between bipartite graphs and bigraphs. A *bipartite graph* is a graph (V, E) whose set V of vertices can be partitioned into two classes such that all edges connect vertices in different classes. By a *bigraph* we mean a triple $G = (X, Y, \Phi)$ where X and Y are (not necessarily disjoint) sets and $\Phi \subseteq X \times Y$. Thus a bipartite graph can be viewed as a bigraph if we fix a partition and specify which bipartition class is first and second. Conversely, if $G = (X, Y, \Phi)$ is a bigraph then we obtain a bipartite graph $(V(G), E(G))$ (the bipartite graph associated to G) by letting $V(G)$ be the disjoint union of the sets X and Y , and $E(G) = \Phi$; by abuse of language we call $V(G)$ the set of *vertices* of G and $E(G)$ the set of *edges* of G . We also say that G is a bigraph *on* $V = V(G)$. (What we call a bigraph $G = (X, Y, \Phi)$ is sometimes called an incidence structure, and $(V(G), E(G))$ is called its Levi graph or incidence graph.) A bigraph is said to be finite if its set of vertices is finite. It is easy to see that a finite bigraph G can have at most $\frac{1}{4}|V(G)|^2$ edges.

A bigraph $G' = (X', Y', \Phi')$ is a *sub-bigraph* of $G = (X, Y, \Phi)$ if $X \subseteq X'$, $Y \subseteq Y'$, and $\Phi' \subseteq \Phi$. We say that a bigraph G contains a given bigraph G' (as a sub-bigraph) if G' is isomorphic to a sub-bigraph of G . Given a bigraph $G = (X, Y, \Phi)$ and a subset V of its vertex set $V(G)$, we denote by

$$G \upharpoonright V := (X \cap V, Y \cap V, \Phi \cap (V \times V))$$

the sub-bigraph of G *induced on* V . The *complement* of a bigraph $G = (X, Y, \Phi)$ is the bigraph $\neg G := (X, Y, \neg \Phi)$, and its *dual* is $G^* := (Y, X, \Phi^*)$ where $\neg \Phi$ and Φ^* are as in Section 2.3.

Let $\varphi(x; y)$ be a partitioned \mathcal{L} -formula, where $|x| = m$, $|y| = n$. We may associate a bigraph $G_\varphi = (X, Y, \Phi)$ to φ and \mathbf{M} , where $X = M^m$, $Y = M^n$, and

$$\Phi = \varphi(M^m; M^n) = \{(a, b) \in M^m \times M^n : \mathbf{M} \models \varphi(a; b)\}.$$

Note that $G_{\neg \varphi} = \neg G_\varphi$ and $G_{\varphi^*} = (G_\varphi)^*$. If we want to stress the dependence of G_φ on \mathbf{M} , then we write $G_\varphi^{\mathbf{M}}$ instead of G_φ . If φ is invariant under the extension $\mathbf{M} \subseteq \mathbf{N}$ of \mathcal{L} -structures, then $G_\varphi^{\mathbf{N}} \upharpoonright V = G_\varphi^{\mathbf{M}}$ where $V = V(G_\varphi^{\mathbf{M}})$.

From now on until the end of this subsection we assume that M is infinite and $m = n$. The collection

$$E(G_\varphi) = \{(a, b) : (a, b) \in \varphi(M^m; M^m)\} \subseteq M^m \times M^m$$

of edges of G_φ then maps naturally onto the definable family

$$\mathcal{S}_{\widehat{\varphi}} = \{\{a, b\} : (a, b) \in \varphi(M^m; M^m)\} \subseteq \binom{M^m}{\leq 2}$$

of subsets of M^m by a map whose fibers have at most 2 elements; here $\widehat{\varphi}(v; x, y)$ is the partitioned \mathcal{L} -formula with object variables $v = (v_1, \dots, v_m)$ and parameter variables (x, y) given by

$$\widehat{\varphi}(v; x, y) := \varphi(x; y) \wedge (v = x \vee v = y).$$

Note that $\text{VC}(\widehat{\varphi}) \leq 2$. Also, $\mathcal{S}_{\widehat{\varphi}^*} = \mathcal{S}_{\widehat{\varphi}}$ and hence $\widehat{\varphi}^*$ and $\widehat{\varphi}$ have the same VC dimension and VC density. A bound on the number of subsets of a given finite set which are cut out by $\mathcal{S}_{\widehat{\varphi}}$ may be computed as follows:

Lemma 4.1. *Let $A \subseteq M^m$ be finite. Then*

$$|A_0| + \frac{1}{2}|E(G_\varphi \upharpoonright V)| \leq |A \cap \mathcal{S}_{\widehat{\varphi}}| \leq 1 + |A_0| + |E(G_\varphi \upharpoonright V)|$$

where

- (1) A_0 is the set of all $a \in A$ such that $\mathbf{M} \models \varphi(a; b)$ or $\mathbf{M} \models \varphi(b; a)$ for some $b \in M^m$, but there is no $b \in A$ with $\mathbf{M} \models \varphi(a; b)$ or $\mathbf{M} \models \varphi(b; a)$, and
- (2) $V \subseteq V(G_\varphi)$ is the disjoint union of A considered as a subset of X and A considered as a subset of Y .

Proof. Each set $S \in A \cap \mathcal{S}_{\widehat{\varphi}}$ is of one of the following types: $S = \emptyset$; $S = \{a\}$ where $a \in A_0$; or $S = \{a, b\}$ where $a, b \in A$ with $\mathbf{M} \models \varphi(a; b)$ or $\mathbf{M} \models \varphi(b; a)$. Each set of the last two types actually occurs in $A \cap \mathcal{S}_{\widehat{\varphi}}$, whereas $S = \emptyset$ only occurs iff there is some edge (a, b) of G_φ with $a, b \notin A$. \square

Hence if we set

$$\Pi_\varphi(t) := \max\{|E(G_\varphi \upharpoonright V)| : V \subseteq V(G_\varphi), |V| = t\} \in \mathbb{N},$$

then the lemma shows that

$$\frac{1}{2}\Pi_\varphi(t) \leq \pi_{\widehat{\varphi}}(t) \leq 1 + t + \Pi_\varphi(2t) \quad \text{for every } t. \quad (4.1)$$

This observation opens up a road to computing (upper or lower) bounds on the VC density of the formula $\widehat{\varphi}$: *find a bound on the number of edges of the subgraph of G_φ induced on finite subsets of its vertex set, in terms of the number of vertices.* In the following we give some applications of this approach.

For positive integers r and s we denote by $K_{r,s} := ([r], [s], [r] \times [s])$ the complete bigraph with the vertex set $[r] \cup [s]$. The following is a fundamental fact about finite bigraphs:

Theorem 4.2 (Kővári, Sós and Turán [51]). *Let $r \leq s$ be positive integers. There exists a real number $C = C(r, s)$ such that every finite bigraph G which does not contain $K_{r,s}$ as a sub-bigraph has at most $C |V(G)|^{2-1/r}$ edges.*

(In fact, a more precise bound is also available, in terms of the sizes of the vertex sets X and Y , but we won't need this.)

Corollary 4.3. *Let $r \leq s$ be positive integers. There is a real number $C_1 = C_1(r, s)$ with the following property: if $\varphi(x; y)$ is an \mathcal{L} -formula such that G_φ does not contain $K_{r,s}$ as a subgraph, then $\pi_{\widehat{\varphi}}(t) \leq C_1 t^{2-1/r}$ for every t ; in particular, $\text{vc}(\widehat{\varphi}) \leq 2 - \frac{1}{r}$.*

Proof. If $V \subseteq V(G_\varphi)$ is finite, and the bigraph $G_\varphi \upharpoonright V$ does not contain $K_{r,s}$, then $|E(G_\varphi \upharpoonright V)| \leq C |V|^{2-1/r}$ by Theorem 4.2, where $C = C(r, s) > 0$ is as in that theorem. Thus $\pi_{\widehat{\varphi}}(t) \leq 1 + t + \Pi_\varphi(2t) \leq 2(1 + 2^{1-1/r}C) t^{2-1/r}$ by (4.1). \square

Given integers $r, s \geq 1$, the bigraph G_φ contains $K_{r,s}$ if and only if there are pairwise distinct $a_1, \dots, a_r \in M^m$ and pairwise distinct $b_1, \dots, b_s \in M^m$ such that $\mathbf{M} \models \varphi(a_i; b_j)$ for all $i \in [r], j \in [s]$. It is interesting to note that if G_φ does not contain $K_{r,s}$ as a sub-bigraph, for some $r, s \geq 1$, then the bigraph $G_{\neg\varphi}$ associated to $\neg\varphi$ does contain $K_{t,t}$, for every $t \geq 1$: by an analogue of Ramsey's Theorem for bigraphs due to Erdős and Rado [32], for every t there exists an n such that for all bigraphs G with $|V(G)| \geq n$, one of $G, \neg G$ contains $K_{t,t}$ as a sub-bigraph. Hence in this case the VC density of the formula $\widehat{\neg\varphi}$ associated to $\neg\varphi$ equals 2, by (4.1).

4.2. Point-line incidences. Let K be an infinite field, construed as a first-order structure in the language of rings as usual. The partitioned formula

$$\varphi(x_1, x_2; y_1, y_2) := x_2 = y_1 x_1 + y_2$$

gives rise to the bigraph $G_\varphi = (X, Y, \Phi)$ where $X = Y = K^2$ and

$$\Phi = \{((\eta, \xi), (a, b)) \in K^2 \times K^2 : \eta = a\xi + b\}.$$

We may think of $V(G_\varphi) = X \cup Y$ as the disjoint union of the set X of points $p = (\eta, \xi) \in K^2$ in the affine plane $\mathbb{A}^2(K)$ over K and the set Y of non-vertical lines ℓ in $\mathbb{A}^2(K)$; thus $E(G_\varphi)$ is the set of point-line incidences (p, ℓ) where $p \in \mathbb{A}^2(K)$ and $\ell \subseteq \mathbb{A}^2(K)$ is a non-vertical line containing p . The bigraph G_φ does not contain $K_{2,2}$ as a subgraph. (Two distinct points in $\mathbb{A}^2(K)$ lie on a unique line.) Hence by Corollary 4.3:

Corollary 4.4. *There is a real number $C_1 > 0$ (independent of K) such that $\pi_{\widehat{\varphi}}(t) \leq C_1 t^{3/2}$ for every t ; in particular, $\text{vc}(\widehat{\varphi}) \leq \frac{3}{2}$.*

Note that this bound is better than what we get from the general estimate $\text{vc} \leq \text{VC}$, since $\text{VC}(\widehat{\varphi}) = 2$. Also, if $K = \mathbb{R}$, then for our original formula φ we have $\pi_\varphi(t) = 1 + t + \binom{t}{2}$ for every t . In particular $\text{VC}(\varphi) = \text{vc}(\varphi) = 2$.

A lower bound on $\text{vc}(\widehat{\varphi})$ is given by:

Lemma 4.5. *Suppose K has characteristic 0. Then $\text{vc}(\widehat{\varphi}) \geq \frac{4}{3}$.*

Proof. This is due to Erdős, with the following simpler argument by Elekes [31]: let k be a positive integer, $t = 4k^3$, and consider the subsets

$$\begin{aligned} P &:= \{(\eta, \xi) : \eta = 0, 1, \dots, k-1, \xi = 0, 1, \dots, 4k^2 - 1\} \\ L &:= \{(a, b) : a = 0, 1, \dots, 2k-1, b = 0, 1, \dots, 2k^2 - 1\} \end{aligned}$$

of \mathbb{Z}^2 , and set $V := P \cup L \subseteq V(G_\varphi)$. Then for each $i = 0, 1, \dots, k-1$, each line $\eta = a\xi + b$ with $(a, b) \in L$ contains a point $(\eta, \xi) \in P$ with $\xi = i$, so

$$|E(G_\varphi \upharpoonright V)| \geq k \cdot |L| = 4k^4 = \frac{1}{4^{1/3}} t^{4/3} = \frac{1}{4} |V|^{4/3}$$

and hence $\text{vc}(\widehat{\varphi}) \geq \frac{4}{3}$ by (4.1). \square

The precise value of $\text{vc}(\widehat{\varphi})$ depends on the characteristic of K :

Proposition 4.6.

- (1) *Suppose K has characteristic 0. Then $\text{vc}(\widehat{\varphi}) = \frac{4}{3}$.*
- (2) *Suppose K has positive characteristic. Then $\text{vc}(\widehat{\varphi}) = \frac{3}{2}$.*

In the proof of this proposition we use the following generalization of a famous theorem of Szémerédi and Trotter [96] (although a weaker version of this theorem from [93], with a somewhat simpler proof, would also suffice for our purposes):

Theorem 4.7 (Tóth [97]). *There exists a real number C such that for all $m, n > 0$ there are at most $C(m^{2/3}n^{2/3} + m + n)$ incidences among m points and n lines in the affine plane over \mathbb{C} .*

Proof of Proposition 4.6. The lower bound $\text{vc}(\widehat{\varphi}) \geq \frac{4}{3}$ in (1) was shown in the previous lemma. From Theorem 4.7 and Lemma 4.1 we obtain $\text{vc}^{\mathbb{C}}(\widehat{\varphi}) \leq \frac{4}{3}$. If K is any field of characteristic 0 with algebraic closure K^{alg} , then $\pi_{\widehat{\varphi}}^K \leq \pi_{\widehat{\varphi}}^{K^{\text{alg}}} = \pi_{\widehat{\varphi}}^{\mathbb{C}}$ by Lemmas 3.1 and 3.2, showing part (1) of Proposition 4.6.

The upper bound $\text{vc}(\widehat{\varphi}) \leq \frac{3}{2}$ in (2) is a consequence of Corollary 4.4. For the lower bound we use the following observation: if F is a finite subfield of K , say $|F| = q$, then $|V(G_\varphi^F)| = 2q^2$ and $|E(G_\varphi^F)| = q^3$, hence

$$|E(G_\varphi^K \upharpoonright V)| = |E(G_\varphi^F)| = \frac{1}{\sqrt{8}} |V|^{3/2} \quad \text{where } V = V(G_\varphi^F).$$

Together with (4.1) this yields the inequality $\text{vc}(\widehat{\varphi}) \geq \frac{3}{2}$ in (2). \square

Proposition 4.6 shows in particular that there is no hope for a “Łos Theorem” for VC density: if \mathbf{M} is a non-principal ultraproduct of a family $(\mathbf{M}_i)_{i \in I}$ of infinite \mathcal{L} -structures, then one may have $\text{vc}^{\mathbf{M}}(\varphi) \neq \text{vc}^{\mathbf{M}_i}(\varphi)$ for all $i \in I$.

It is interesting to contrast Proposition 4.6 with the outcome of only considering parameters from an indiscernible sequence:

Lemma 4.8. *The formula $\widehat{\varphi}$ has alternation number 2, hence $\text{vc}_{\text{ind}}(\widehat{\varphi}) = 1$.*

Proof. It suffices to show $\text{alt}(\widehat{\varphi}) = 2$, since then Lemma 3.24 yields $\text{vc}_{\text{ind}}(\widehat{\varphi}) = 1$. Suppose for a contradiction that $(a_i)_{i \in \mathbb{N}}$ is an indiscernible sequence in K^2 and $b = (p, \ell) \in K^2 \times K^2$ witnessing that $\text{alt}(\widehat{\varphi}) \geq 3$. We think of the elements of K^2 both as points in the affine space $\mathbb{A}^2(K)$ over K and as non-vertical lines in $\mathbb{A}^2(K)$, and let i, j range over $\{0, 1, 2, 3\}$. The a_i are pairwise distinct, $p \in \ell$, and $a_i = p, a_j = \ell$ for some $i \neq j$; hence $a_i \in a_j$ for some $i \neq j$. If $a_i \in a_j$ where $i < j$, then $a_i \in a_j$ for all $i < j$ (by indiscernibility) and hence $a_0, a_1 \in a_2 \cap a_3$, and this forces $a_0 = a_1$ or $a_2 = a_3$, in both cases a contradiction. Similarly the assumption that $a_i \in a_j$ with $i > j$ leads to a contradiction. \square

Many other results in the combinatorial literature lead to non-trivial (upper and lower) bounds on $\text{vc}(\widehat{\varphi})$ if φ encodes the incidence of points on various geometric objects; see [66, Chapter 4] or [74]. For example, let $\mathbf{R} = (\mathbb{R}, 0, 1, +, -, \times, <)$ be the ordered field of real numbers. Let

$$\varphi(x_1, x_2; y_1, y_2) := (x_1 - y_1)^2 + (x_2 - y_2)^2 = 1,$$

so $\mathcal{S}_\varphi^{\mathbf{R}}$ is the collection of circles with radius 1 in the plane, and $E(G_\varphi^{\mathbf{R}})$ is the set of incidences between points in \mathbb{R}^2 and circles of radius 1. Then an analogue of the Szémeredi-Trotter Theorem [95] (or a more general result due to Pach and Sharir [73] on families of simple plane curves) and (4.1) yields $\text{vc}^{\mathbf{R}}(\widehat{\varphi}) \leq \frac{4}{3}$. (However, it is unknown whether this bound is sharp, cf. [74, Section 2].)

4.3. Irrational VC density. In [8] it is shown that for every real number $r \geq 1$ there exists a set system $\mathcal{S} \subseteq \binom{\mathbb{N}}{[r]}$ with $\text{vc}(\mathcal{S}) = r$. We do not know the answer to the following question (though we suspect the answer to be negative):

Question. Is the VC density of a formula in a NIP theory always rational?

Let $\mathcal{L}_{\text{gr}} = \{E\}$ be the language with a single binary relation symbol E . The \mathcal{L}_{gr} -structures are nothing but the (directed) graphs (with E interpreted as the edge relation). Given a graph G we denote by $V(G)$ its set of vertices and by $E(G)$ its set of edges. Spencer and Shelah [91] established a 0-1-law for \mathcal{L}_{gr} -sentences about random (symmetric, loopless) graphs with n vertices and edge probability $n^{-\alpha}$, where α is an irrational number between 0 and 1. We denote the resulting complete \mathcal{L}_{gr} -theory by T_α . It was shown by Baldwin and Shelah [9] that T_α is stable. (This can also be checked by simply verifying that T_α is superflat in the sense of [77]; cf. Section 4.4 below.) In particular, T_α is NIP, so it makes sense to investigate VC density of formulas in T_α . We consider the following \mathcal{L}_{gr} -formula:

$$\varphi(x; y) = E(x, y) \vee x = y.$$

For any graph G and vertex v of G , the formula $\varphi(x; v)$ defines the (closed) neighborhood of v , i.e., the set consisting of v together with all vertices adjacent to it. It is tempting to guess that $\text{vc}(\varphi) = 1/\alpha$. (This would give rise to a negative answer of the question posed above.) However, it turns out that $\text{vc}(\varphi)$ is an integer:

Lemma 4.9. $\text{vc}(\varphi) = \lfloor 1/\alpha \rfloor$.

Before we give the proof, we recall some basic facts about the theory T_α ; our main reference is [94]. We let G be a model of T_α .

A *rooted graph* is a pair (R, H) where H is a finite graph and R a proper subset of its set of vertices; the elements of R will be called *roots*. We consider each finite non-empty graph H as a rooted graph by identifying it with (\emptyset, H) . Given a rooted graph (R, H) , a

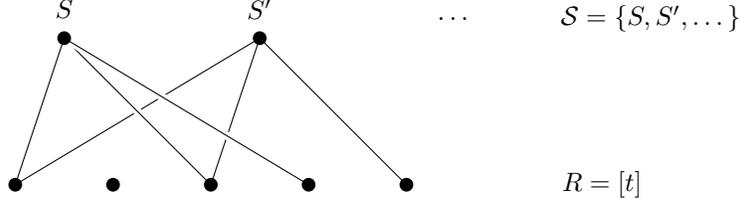


FIGURE 4.1. The rooted graph associated to a set system

rooted graph (R, H') , where H' is subgraph of H whose vertex set properly contains R , is called a rooted subgraph of (R, H) .

A *weak embedding* of a rooted graph (R, H) into G is an injective map $\iota: V(H) \rightarrow V(G)$ such that for all roots v and non-roots w of (R, H) , v and w are adjacent in H iff $\iota(v)$ and $\iota(w)$ are adjacent in G ; such a weak embedding is called an *embedding* if also any two non-roots v and w of (R, H) are adjacent in H iff $\iota(v)$ and $\iota(w)$ are adjacent in G . Note that there is no requirement about edges between roots. (This terminology does not appear in [94] which talks about “ (R, H) -extensions” instead.)

Let (R, H) be a rooted graph. The *average degree* of (R, H) is $\text{adeg}(R, H) := 2e/v$ where $v = v(R, H) > 0$ is the number of vertices of H which are not roots and $e = e(R, H)$ is the number of edges of H which do not have both ends in R . The *maximum average degree* $\text{mdeg}(R, H)$ of (R, H) is defined as the maximum of $\text{adeg}(R, H')$ where (R, H') is a rooted subgraph of (R, H) . If $\text{adeg}(R, H) > 2/\alpha$ then (R, H) is called *dense*, and *sparse* otherwise (i.e., if $\text{adeg}(R, H) < 2/\alpha$). If $\text{mdeg}(R, H) < 2/\alpha$ then (R, H) is called *safe*, and *unsafe* otherwise.

Now if H is dense then G does not contain a copy of H , whereas if H is safe then G contains a copy (indeed, an induced copy) of H . More generally, if (R, H) is unsafe then there is no weak embedding of (R, H) into G [94, p. 69], and if (R, H) is safe then every injective map $R \rightarrow V(G)$ extends to an embedding of (R, H) into G [94, Theorem 5.2.1].

Let now \mathcal{S} be a non-empty set system on $[t] = \{1, \dots, t\}$, where $t > 0$. We associate a rooted graph $(R, H) = (R_{\mathcal{S}}, H_{\mathcal{S}})$ to \mathcal{S} as follows: the set of vertices of H is the disjoint union of $[t]$ and \mathcal{S} , the set of roots is $R = [t]$, there are no edges between two roots and no edges between two non-roots, and a root $i \in [t]$ and a non-root $S \in \mathcal{S}$ are related by an edge iff $i \in S$. (Cf. Figure 4.1.) Note that this rooted graph has average degree $\text{adeg}(R, H) = \frac{2}{|\mathcal{S}|} \sum_{S \in \mathcal{S}} |S|$ and maximum average degree

$$\text{mdeg}(R, H) = \max_{\emptyset \neq \mathcal{S}' \subseteq \mathcal{S}} \frac{2}{|\mathcal{S}'|} \sum_{S \in \mathcal{S}'} |S|.$$

So if $\mathcal{S} \subseteq \binom{[t]}{k}$ where $k \in \{0, \dots, t\}$ then $\text{adeg}(R, H) = \text{mdeg}(R, H) = 2k$; hence if in addition $k < 1/\alpha$ then (R, H) is safe (so there exists an embedding of (R, H) into G) whereas if $k > 1/\alpha$ then (R, H) is dense (and so there is no weak embedding of (R, H) into G).

Proof of Lemma 4.9. Applying the remarks above to $\mathcal{S} = \binom{[t]}{[1/\alpha]}$, where $t > 0$, we see that, as $[1/\alpha] < 1/\alpha$, there are pairwise distinct vertices a_1, \dots, a_t and b_S ($S \in \mathcal{S}$) of G such that a_i and b_S are adjacent iff $i \in S$, i.e., writing $A = \{a_1, \dots, a_t\}$ we have

$A \cap \varphi(G; b_S) = \{a_i : i \in S\}$ and hence $|A \cap \mathcal{S}_\varphi| \geq |S|$. So $\pi_\varphi(t) \geq \binom{t}{\lfloor 1/\alpha \rfloor}$ for each t , therefore $\text{vc}(\varphi) \geq \lfloor 1/\alpha \rfloor$.

To show the reverse inequality suppose for a contradiction that $\text{vc}(\varphi) > \lfloor 1/\alpha \rfloor$. Let ρ be a real number with $\text{vc}(\varphi) > \rho > \lfloor 1/\alpha \rfloor$. Note that for every set A of vertices of G the set system $A \cap \mathcal{S}_\varphi$ is the union of the set system

$$\{\{b\} \cup \{a \in A : (a, b) \in E(G)\} : b \in A\} \quad (4.2)$$

consisting of at most $|A|$ sets, and

$$\{\{a \in A : (a, b) \in E(G)\} : b \in V(G) \setminus A\}. \quad (4.3)$$

Since $\text{vc}(\varphi) > \rho$, for every $C \geq 1$ there are arbitrarily large $t > 0$ and $A \subseteq V(G)$ with $|A| = t$ such that $|A \cap \mathcal{S}_\varphi| \geq 2Ct^\rho$. The set system (4.2) has at most t elements; hence (4.3) contains at least $2Ct^\rho - t$ sets and so, since $\rho > \lfloor 1/\alpha \rfloor \geq 1$, contains at least Ct^ρ sets. Identifying A with $[t]$, the set system (4.3) on A thus gives rise to a set system \mathcal{S} on $[t]$ with $|\mathcal{S}| \geq Ct^\rho$ whose associated rooted graph $(R_\mathcal{S}, H_\mathcal{S})$ weakly embeds into G .

On the other hand, let \mathcal{S} be any set system on $[t]$ such that $(R_\mathcal{S}, H_\mathcal{S})$ weakly embeds into G , and for $k = 0, \dots, t$ consider the set system $\mathcal{S}_k := \mathcal{S} \cap \binom{[t]}{k}$ on $[t]$. If $\mathcal{S}_k \neq \emptyset$ then $(R_{\mathcal{S}_k}, H_{\mathcal{S}_k})$ is a rooted subgraph of $(R_\mathcal{S}, H_\mathcal{S})$ and hence

$$2k = \text{adeg}(R_{\mathcal{S}_k}, H_{\mathcal{S}_k}) \leq \text{mdeg}(R_\mathcal{S}, H_\mathcal{S}) < 2/\alpha.$$

Therefore \mathcal{S} does not contain a k -element subset of $[t]$ with $k > 1/\alpha$, i.e., $\mathcal{S} \subseteq \binom{[t]}{\leq \lfloor 1/\alpha \rfloor}$. Hence $|\mathcal{S}| \leq Ct^{\lfloor 1/\alpha \rfloor}$ where $C = C_\alpha$ is a constant only depending on α . This contradicts the previous paragraph. \square

We remark that a similar analysis shows that the simpler formula $E(x, y)$ also has VC density $\lfloor 1/\alpha \rfloor$ in T_α . We chose φ as above because it allows us to compare Lemma 4.9 with the main result of [4], where the precise value of the VC dimension of φ (as it depends on α) is computed. In particular, [4, Corollary 8] shows that if $0 < \alpha < \frac{93}{650}$ then

$$\lfloor 1/\alpha \rfloor + 3 \leq \text{VC}(\varphi) \leq \lfloor 1/\alpha + 3(\alpha + 1) \rfloor.$$

4.4. Shatter functions not growing like a power. We finish this section with two examples of VC classes definable in NIP theories whose shatter function is not asymptotic to a real power function.

4.4.1. The hypercube. Let Q be the ‘‘infinitary hypercube’’, i.e., the (symmetric, loopless) graph whose vertex set is the set of all sequences $s = (s_n)$ in $\{0, 1\}^{\mathbb{N}}$ with finite support, with two sequences related by an edge iff they differ in only one component. (Alternatively, Q can be represented as the set of all finite sets of natural numbers, with an edge between them iff their symmetric difference is a singleton.) Note that Q is the increasing union $Q = \bigcup_{d>0} Q_d$ of its induced subgraphs Q_d with vertex set

$$V(Q_d) = \{s = (s_n) \in Q : s_n = 0 \text{ for } n \geq d\},$$

which we may identify with $\{0, 1\}^d$ in the natural way. (So Q_d is the d -dimensional hypercube.) We construe Q as an \mathcal{L}_{gr} -structure. In [77] a condition sufficient for a (symmetric) graph to be stable is introduced, called superflatness: a graph G is superflat if for every m there is some n such that no subdivision of the complete graph K_n on n vertices, obtained by placing at most m additional vertices on each edge, embeds into G . Note that the graph Q is not superflat: in fact, for every d there is an embedding of a subdivision of K_{d+1} , obtained by placing at most one additional vertex on each edge, into Q_d , cf. [37]. However, we do have:

Proposition 4.10. *Q is ω -stable.*

Towards a proof of this proposition, we first introduce some notation and terminology: Given a language \mathcal{L} let $\mathcal{L}(P) = \mathcal{L} \cup \{P\}$ where P is a new predicate symbol, and given an \mathcal{L} -structure \mathbf{M} and a subset A of its domain let (\mathbf{M}, A) be the expansion of \mathbf{M} to an $\mathcal{L}(P)$ -structure obtained by interpreting P by A . The *induced* structure A_{ind} on A is the structure whose language consists of an m -ary relation symbol R_φ for every \mathcal{L} -formula $\varphi(x)$, where $m = |x|$, interpreted in A_{ind} by $\varphi^{\mathbf{M}}(M^m) \cap A^m$. An $\mathcal{L}(P)$ -formula $\varphi(x)$, where $x = (x_1, \dots, x_m)$, is said to be *bounded* if it has the following form (slightly abusing syntax):

$$\varphi(x) = \diamond_1 y_1 \in P \cdots \diamond_n y_n \in P \psi(x_1, \dots, x_m, y_1, \dots, y_n),$$

where $\diamond_i \in \{\forall, \exists\}$ and ψ is an \mathcal{L} -formula. Casanovas and Ziegler have shown:

Theorem 4.11. *Let \mathbf{M} be a structure in a language \mathcal{L} and let $A \subseteq M$.*

- (1) *Suppose \mathbf{M} is strongly minimal. Then in (\mathbf{M}, A) , every $\mathcal{L}(P)$ -formula is equivalent to a bounded formula.*
- (2) *Suppose that in (\mathbf{M}, A) , every $\mathcal{L}(P)$ -formula is equivalent to a bounded formula, and let $\lambda \geq |\mathcal{L}|$ be a cardinal. If both \mathbf{M} and A_{ind} are λ -stable then (\mathbf{M}, A) is λ -stable.*

(See [18, Corollary 5.4 and Proposition 3.1]; part (1) had actually first been shown by Pillay [76].)

Consider now, slightly more general than necessary, an arbitrary field K , and let $\mathcal{L}_K = \{0, +, (\lambda \cdot)_{\lambda \in K}\}$ be the language of K -vector spaces. Let M be an infinite-dimensional K -vector space. Then M , construed as an \mathcal{L}_K -structure, has quantifier elimination and is ω -stable. Let A be a set of linearly independent elements of M . Then the induced structure on A is trivial: every subset of A^m definable in A_{ind} is definable in the empty language. Hence by the theorem above, the $\mathcal{L}_K(P)$ -structure (M, A) is ω -stable.

For the proof of Proposition 4.10, it now suffices to note that the \mathcal{L}_{gr} -structure Q is definable in (M, A) , for suitable choice of K , M and A : Take $K = \mathbb{F}_2$ and let $M = \bigoplus_n \mathbb{F}_2 a_n$ be a countably infinite \mathbb{F}_2 -vector space with distinguished basis $A = \{a_n : n \geq 0\}$. Then Q is definable in (M, A) : identifying $V(Q)$ with M in the natural way, we have, for all vertices s, t of Q : $(s, t) \in E(Q)$ iff $s - t \in A$. Since (M, A) is ω -stable, so is Q . \square

Now consider the \mathcal{L}_{gr} -formula $\varphi(x; y) := E(x, y)$. Then

$$\mathcal{S}_\varphi = \{\{s, s'\} : s, s' \in Q, (s, s') \in E(Q)\}$$

is the collection of undirected edges of Q . In the following A denotes a subset of Q (unlike in the proof of Proposition 4.10). Note that if $A \subseteq Q_d$ then

$$A \cap \mathcal{S}_\varphi = \binom{A}{1} \cup E[A] \quad \text{where } E[A] := \{\{a, a'\} : a, a' \in A, (a, a') \in E(Q)\}.$$

Since Q_d has 2^d vertices and $\frac{1}{2}d2^d$ undirected edges, we thus see that $\pi_\varphi(t) \geq t + \frac{1}{2}t \log t$ for infinitely many t ; in fact:

Proposition 4.12. $\pi_\varphi(t) = \frac{1}{2}t \log t (1 + o(1))$ as $t \rightarrow \infty$.

Proof. Set

$$E_d(t) := \max \{ |E[A]| : A \subseteq Q_d, |A| = t \} \quad \text{for } d > 0 \text{ and } t \leq 2^d.$$

Then $\pi_{\widehat{\varphi}}(t) = t + \max_{d \geq \lceil \log t \rceil} E_d(t)$. It is known (see, e.g., [1]) that there is some function g with $g(t) = \frac{1}{2}t \log t (1 + o(1))$ as $t \rightarrow \infty$ such that $E_d(t) = g(t)$ for all d and $t \leq 2^d$. This yields the claim. \square

4.4.2. *An example in $\mathbf{R} = (\mathbb{R}, 0, 1, +, -, \times, <)$.* For this we use another one of the rare examples (besides the Szémeredi-Trotter Theorem) where tight bounds on the number of incidences are known:

Theorem 4.13 (Pach and Sharir [72]). *Let α be a real number with $0 < \alpha < \pi$. The maximum number of times that α occurs as an angle among the ordered triples of t points in the plane is $O(t^2 \log t)$. Furthermore, suppose $\tan(\alpha) \in \mathbb{Q}\sqrt{d}$ where $d \in \mathbb{N}$ is not a square. Then there exists a constant $C = C_\alpha > 0$ and, for every $t > 3$, a t -element set $S_t \subseteq \mathbb{R}^2$ with the property that at least $Ct^2 \log t$ ordered triples of points from S_t determine the angle α .*

Let $x = (x_1, x_2)$, $y = (y_1, y_2)$, $z = (z_1, z_2)$ and consider the formula

$$\varphi(x, y, z) := x \neq y \wedge x \neq z \wedge 2\langle y - x, z - x \rangle = \|y - x\| \|z - x\|$$

in the language of the ordered field of real numbers \mathbf{R} , where $\langle \cdot, \cdot \rangle$ denotes the usual inner product on \mathbb{R}^2 and $\| \cdot \|$ the associated norm. Then for $a, b, c \in \mathbb{R}^2$ we have $\mathbf{R} \models \varphi(a, b, c)$ iff the vectors $b - a$ and $c - a$ are non-zero and the angle $\angle(b, a, c)$ between them is $\frac{\pi}{3}$. Let $\widehat{\varphi}(v; x, y, z)$ be the partitioned formula with object variables $v = (v_1, v_2)$ and parameter variables (x, y, z) given by

$$\widehat{\varphi}(v; x, y, z) := \varphi(x, y, z) \wedge (v = x \vee v = y \vee v = z),$$

so $\mathcal{S}_{\widehat{\varphi}}$ consists of all $\{a, b, c\} \in \binom{\mathbb{R}^2}{3}$ with $\mathbf{R} \models \varphi(a, b, c)$. We now have:

Corollary 4.14. *There exist constants $C_1, C_2 > 0$ such that*

$$C_1 t^2 \log t < \pi_{\widehat{\varphi}}(t) < C_2 t^2 \log t \quad \text{for every } t > 0.$$

That is, $\pi_{\widehat{\varphi}}(t) = \Theta(t^2 \log t)$ as $t \rightarrow \infty$.

Proof. Let $A \subseteq \mathbb{R}^2$ be finite. Then

$$A \cap \mathcal{S}_{\widehat{\varphi}} = \binom{A}{\leq 2} \cup \left\{ \{a, b, c\} \in \binom{A}{3} : \mathbf{R} \models \varphi(a, b, c) \right\}.$$

Since $\tan(\pi/3) = \sqrt{3}$, the second part of Theorem 4.13 applies to $\alpha = \frac{\pi}{3}$. The upper bound in Corollary 4.14 now follows from the first assertion in Theorem 4.13, and the lower bound from the second assertion in the same theorem. \square

5. THEORIES WITH THE VC d PROPERTY

After defining the VC d property we prove that if a theory has this property then the dual VC density of any finite set of partitioned formulas in the tuple of object variables x is at most $d|x|$. In Section 6 we will then show that various theories have the VC d property. In the following \mathbf{M} is a structure in a language \mathcal{L} , $\Delta(x; y)$ is a finite non-empty set of partitioned \mathcal{L} -formulas, and $m = |x|$.

5.1. Uniform definability of types over finite sets. Given a $\Delta(x; B)$ -type $q \in S^\Delta(B)$, where $B \subseteq M^{|y|}$, a family $\mathcal{F} = (\varphi_\#)_{\varphi \in \Delta}$ of $\mathcal{L}(M)$ -formulas $\varphi_\#(y)$ is said to *define* q if for all $\varphi \in \Delta$ and $b \in B$ we have

$$\varphi(x; b) \in q \quad \iff \quad \mathbf{M} \models \varphi_\#(b).$$

We also say that \mathcal{F} is a *definition* of q . For a family $\mathcal{F} = \mathcal{F}(y; v)$ of partitioned \mathcal{L} -formulas $\psi(y; v)$, we denote by $\mathcal{F}(y; c)$ the family $(\psi(y; c))_{\psi \in \mathcal{F}}$ of $\mathcal{L}(M)$ -formulas obtained by substituting a given tuple $c \in M^{|v|}$ for the tuple of variables v . The following generalizes a definition due to Guingona [34]:

Definition 5.1. We say that Δ has *uniform definability of types over finite sets* (abbreviated as *UDTFS*) in \mathbf{M} if there are finitely many families

$$\mathcal{F}_i = (\varphi_i(y; y_1, \dots, y_d))_{\varphi \in \Delta} \quad (i \in I)$$

of \mathcal{L} -formulas (where $|y_j| = |y|$ for every $j = 1, \dots, d$) such that for every finite set $B \subseteq M^{|y|}$ and $q \in S^\Delta(B)$ there are $b_1, \dots, b_d \in B$ and some $i \in I$ such that $\mathcal{F}_i(y; b_1, \dots, b_d)$ defines q . We call the family $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ a *uniform definition of $\Delta(x; B)$ -types over finite sets* in \mathbf{M} with d parameters. If $\Delta = \{\varphi\}$ is a singleton, we also speak of φ having UDTFS.

The following observation shows in particular that every finite set $\Delta(x; y)$ of partitioned \mathcal{L} -formulas which is directed (see Example 3.12) has UDTFS in T with a single parameter:

Lemma 5.2. *Let $\Delta(x; y)$ be a finite set of partitioned \mathcal{L} -formulas, and suppose the set system $\mathcal{S}_\Delta = \{\varphi(M^{|x|}; b) : b \in M^{|y|}\}$ has breadth d . Then Δ has UDTFS with d parameters.*

Proof. To see this set $\mathcal{F}_0(y) := (\exists z(z \neq z))_{\varphi \in \Delta}$ and, for each d -tuple $\psi = (\psi_1, \dots, \psi_d) \in \Delta^d$, define

$$\mathcal{F}_\psi(y; y_1, \dots, y_d) := (\varphi_\psi(y; y_1, \dots, y_d))_{\varphi \in \Delta}$$

where

$$\varphi_\psi(y; y_1, \dots, y_d) := \forall x(\psi_1(x; y_1) \wedge \dots \wedge \psi_d(x; y_d) \rightarrow \varphi(x; y)).$$

Then $\mathcal{F} = (\mathcal{F}_\psi)_{\psi \in \{0\} \cup \Delta^d}$ is a uniform definition of $\Delta(x; B)$ -types over finite sets. For suppose $q \in S^\Delta(B)$ where $B \subseteq M^{|y|}$ is finite. If $\varphi(x; b) \notin q$ for all $\varphi \in \Delta$, $b \in B$, then $\mathcal{F}_0(y)$ defines q . Otherwise, by assumption we can pick $\psi_1(x; b_1), \dots, \psi_d(x; b_d) \in q$ such that

$$\bigcap_{\varphi(x; b) \in q} \varphi(M^{|x|}; b) = \psi_1(M^{|x|}; b_1) \cap \dots \cap \psi_d(M^{|x|}; b_d).$$

Then $\mathcal{F}_\psi(y; b_1, \dots, b_d)$, where $\psi = (\psi_1, \dots, \psi_d)$, defines q . \square

A uniform definition of $\Delta(x; B)$ -types over finite sets in \mathbf{M} remains a uniform definition of $\Delta(x; B)$ -types over finite sets in any elementarily equivalent structure, and so it makes sense to speak of uniform definability of types over finite sets in a complete theory. If we do not care about the number of parameters, we can always do with a single defining scheme (at least for non-trivial parameter sets):

Lemma 5.3. *Let $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ be a uniform definition of $\Delta(x; B)$ -types over finite sets with d parameters as above, and let $n = |I|$. Then there exists a family*

$$\mathcal{F}_\# = (\varphi_\#(y; y_1, \dots, y_d, v, w_1, \dots, w_n))_{\varphi \in \Delta}$$

such that for every finite set $B \subseteq M^{|y|}$ with $|B| \geq 2$ and every $q \in S^\Delta(B)$ there are $b_1, \dots, b_d, c, c_1, \dots, c_n \in B$ such that $\mathcal{F}_\#(u; b_1, \dots, b_d, c, c_1, \dots, c_n)$ defines q .

Proof. This is a simple coding trick due to Shelah (proof of Theorem II.2.12 (1) in [88], cf. also [34, Lemma 2.5]). For every $\varphi \in \Delta$ define

$$\varphi_\#(y; y_1, \dots, y_d, v, w_1, \dots, w_n) := \bigwedge_{i=1}^n (v = w_i \rightarrow \varphi_i(y; y_1, \dots, y_d))$$

and let $\mathcal{F}_\# = (\varphi_\#)_{\varphi \in \Delta}$. Let $B \subseteq M^{|y|}$, $|B| \geq 2$, and $q \in S^\Delta(B)$. By hypothesis, there are $b_1, \dots, b_d \in B$ and $i \in I$ such that $\mathcal{F}_i(y; b_1, \dots, b_d)$ defines q . Pick $c, c' \in B$ with $c \neq c'$, and put $c_i := c$, $c_j := c'$ for $j \neq i$. Then $\mathcal{F}_\#(y; b_1, \dots, b_d, c, c_1, \dots, c_n)$ defines q . \square

Similarly, the proof of Lemma 3.18, (1) shows that if the \mathcal{L} -formula ψ_Δ which we associated there to the finite set of \mathcal{L} -formulas Δ admits a uniform definition of $\psi_\Delta(x; B')$ -types over finite sets with d parameters, then Δ itself admits a uniform definition of $\Delta(x; B)$ -types over finite parameter sets B (with at least 2 elements) having $d + 2$ parameters.

On the other hand, if we have tight control over the number d of parameters in our defining schemes, then we can bound the sizes of the $\Delta(x; B)$ -type spaces over finite sets by polynomial functions (in the size of the parameter set) of degree d : more precisely, if Δ allows a uniform definition $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ of $\Delta(x; B)$ -types over finite sets in \mathbf{M} with d parameters, then $|S^\Delta(B)| \leq |I| |B|^d$ for every finite $B \subseteq M^{|y|}$, hence $\pi_\Delta^*(t) \leq |I| t^d$ for each t , and so $\text{vc}^*(\Delta) \leq d$.

5.2. The VC d property. We say that \mathbf{M} has the VC d property if any $\Delta(x; y)$ with $|x| = 1$ has a uniform definition of $\Delta(x; B)$ -types over finite sets with d parameters. Clearly, if \mathbf{M} has the VC d property, then so does every elementarily equivalent \mathcal{L} -structure. We say that a theory T has the VC d property if every model of T has the VC d property.

We point out that the VC0 property only holds in a very special situation; recall that a structure is called rigid if it has no automorphisms besides the identity.

Lemma 5.4. *The following are equivalent:*

- (1) \mathbf{M} has the VC0 property;
- (2) every model of $\text{Th}(\mathbf{M})$ is rigid;
- (3) \mathbf{M} is finite and rigid.

Proof. Suppose \mathbf{M} has the VC0 property; to see (2), it suffices to show that \mathbf{M} is rigid. Let $\varphi(x; y)$ be the \mathcal{L} -formula $x = y$, and let $\mathcal{F} = (\varphi_i(y))_{i \in I}$ be a uniform definition of $\varphi(x; B)$ -types over finite sets. Suppose $\sigma \in \text{Aut}(\mathbf{M})$ and $b \in M$ satisfy $b \neq \sigma(b)$. Let $p = \text{tp}^\varphi(b/B)$ where $B = \{b, \sigma(b)\}$, and choose $i \in I$ such that φ_i defines p . Then

$$b = b \iff \mathbf{M} \models \varphi_i(b) \iff \mathbf{M} \models \varphi_i(\sigma(b)) \iff b = \sigma(b),$$

a contradiction. This shows (1) \Rightarrow (2), and (2) \Rightarrow (3) is obvious. Suppose now that \mathbf{M} is finite, and let $\Delta(x; y)$, where $|x| = 1$, be a finite set of partitioned \mathcal{L} -formulas. It is easy to see that then for each $p \in S^\Delta(M^{|y|})$ there is a family $\mathcal{F}_p = \{\varphi_p(y) : \varphi \in \Delta\}$ of \mathcal{L} -formulas such that for all $b \in M^{|y|}$ we have $\mathbf{M} \models \varphi_p(b)$ iff there is some automorphism σ of \mathbf{M} such that $\varphi(x; b) \in \sigma(p)$. Hence if in addition \mathbf{M} is rigid then $\mathcal{F} = (\mathcal{F}_p)_{p \in S^\Delta(M^{|y|})}$ is a uniform definition of $\Delta(x; B)$ -types over finite sets in \mathbf{M} . This shows (3) \Rightarrow (1). \square

The following result and its Corollary 5.6 are useful if we have some kind of quantifier elimination result at hand:

Lemma 5.5. *Suppose $\Delta = \Delta(x; y)$ and $\Psi = \Psi(x; y)$ are finite sets of partitioned \mathcal{L} -formulas such that every formula in Δ is equivalent in T to a Boolean combination of formulas in Ψ , and Ψ has UDTFs in T with d parameters. Then Δ has UDTFs in T with d parameters.*

Proof. We may assume that each $\varphi \in \Delta$ has the form

$$\varphi = \bigwedge_{r \in R_\varphi} \bigvee_{s \in S_{r,\varphi}} \square_{r,s,\varphi} \psi_{r,s,\varphi}$$

where $R_\varphi, S_{r,\varphi}$ are finite index sets, $\square_{r,s,\varphi}$ is \neg or no condition, and $\psi_{r,s,\varphi} \in \Psi$. Suppose $\mathcal{G} = (\mathcal{G}_i)_{i \in I}$ is a uniform definition of $\Psi(x; B)$ -types over finite sets in T , where

$$\mathcal{G}_i = (\psi_i(y; y_1, \dots, y_d))_{\psi_i \in \Psi} \quad \text{for each } i \in I.$$

Let $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ where

$$\mathcal{F}_i = (\varphi_i(y; y_1, \dots, y_d))_{\varphi_i \in \Delta} \quad \text{where } \varphi_i = \bigwedge_{r \in R_\varphi} \bigvee_{s \in S_{r,\varphi}} \square_{r,s,\varphi} (\psi_{r,s,\varphi})_i.$$

Let $q \in S^\Delta(B)$ where $B \subseteq M^{|y|}$ is finite. Let $a \in M^{|x|}$ realize q , and put $p := \text{tp}^\Psi(a/B)$. Take $i \in I$ and $b_1, \dots, b_d \in B$ such that $\mathcal{G}_i(y; b_1, \dots, b_d)$ defines p . One now easily verifies that $\mathcal{F}_i(y; b_1, \dots, b_d)$ defines q . \square

Corollary 5.6. *Suppose Φ is a family of partitioned \mathcal{L} -formulas in the single object variable x such that*

- (1) *every partitioned \mathcal{L} -formula in the object variable x is equivalent in T to a Boolean combination of formulas from Φ , and*
- (2) *every finite set of \mathcal{L} -formulas from Φ has UDTFs in T with d parameters.*

Then T has the VC d property.

The following theorem is at the root of the proof of Theorem 1.1 from the introduction; it shows that having UDTFs with a constant number of parameters for all sets of formulas in a single object variable entails UDTFs with a linearly bounded number of parameters for sets of formulas in an arbitrary number of object variables:

Theorem 5.7. *Suppose that \mathbf{M} has the VC d property. Then every $\Delta(x; y)$ has a uniform definition of $\Delta(x; B)$ -types over finite sets in \mathbf{M} with $d|x|$ parameters.*

Before we embark on the proof, we introduce some notation: for a sequence $a \in M^m$ and a set $B \subseteq M^n$ we write $aB := \{(a, b) : b \in B\} \subseteq M^{m+n}$ and $Ba := \{(b, a) : b \in B\} \subseteq M^{n+m}$.

Proof. We proceed by induction on $m = |x|$. The base case $m = 1$ holds by hypothesis. For the inductive step write $x = (x_0, x')$ where $x' = (x_1, \dots, x_m)$, and let $\Delta(x; y)$ be given. Let

$$\Delta_0(x_0; x', y) = \{\varphi(x_0; x', y) : \varphi(x; y) \in \Delta\}.$$

By the VC d property applied to Δ_0 , we take finitely many families

$$\mathcal{F}_i = (\varphi_i(x', y; y_1, \dots, y_d))_{\varphi_i \in \Delta} \quad (i \in I)$$

of \mathcal{L} -formulas with the following property: for any $a' \in M^m$, any finite set $B \subseteq M^{|y|}$ and any $q \in S^{\Delta_0}(a'B)$, there are $b_1, \dots, b_d \in B$ and $i \in I$ such that $\mathcal{F}_i(a', y; b_1, \dots, b_d)$ defines q , i.e., for all $\varphi \in \Delta$, $b \in B$:

$$\varphi(x_0; a', b) \in q \iff \mathbf{M} \models \varphi_i(a', b; b_1, \dots, b_d). \quad (5.1)$$

In the rest of this proof let φ range over Δ and i over I . For each i , let

$$\Delta_i(x'; y, y_1, \dots, y_d) = \{\varphi_i(x'; y, y_1, \dots, y_d) : \varphi(x; y) \in \Delta\}$$

and apply the inductive hypothesis to each Δ_i . Thus for each i there are finite families

$$\mathcal{F}_{ij} := (\varphi_{ij}(y, y_1, \dots, y_d; v_1, \dots, v_n))_{\varphi \in \Delta} \quad (j \in J_i)$$

of \mathcal{L} -formulas, where $n = md$, such that for all finite subsets $B \subseteq M^{|y|}$, all $\bar{b} = (b_1, \dots, b_d) \in (M^{|y|})^d$, and every $p \in S^{\Delta_i}(B\bar{b})$, there exists some $j \in J_i$ and $c_1, \dots, c_n \in B$ such that for each φ and each $b \in B$ we have

$$\varphi_i(x'; b, \bar{b}) \in p \iff \mathbf{M} \models \varphi_{ij}(b, \bar{b}; c_1, \dots, c_n). \quad (5.2)$$

Partition the variable tuple of the \mathcal{L} -formulas φ_{ij} as $(y; y_1, \dots, y_d, v_1, \dots, v_n)$, and set $\mathcal{F} := (\mathcal{F}_{ij})_{i \in I, j \in J_i}$. We claim that \mathcal{F} is a uniform definition of $\Delta(x; B)$ -types over finite sets in \mathbf{M} ; since \mathcal{F} has $d + n = d(m + 1)$ parameters, this will then finish the inductive step. To see this, let a finite $B \subseteq M^{|y|}$ and some $a = (a_0, a') \in M^{1+m}$ be given. We let b range over B . We need to show that there are some $i \in I$, $j \in J_i$ and $\bar{b} \in B^d$, $c_1, \dots, c_n \in B$ such that for all φ and b we have

$$\mathbf{M} \models \varphi(a; b) \iff \mathbf{M} \models \varphi_{ij}(b; \bar{b}, c_1, \dots, c_n). \quad (5.3)$$

Let $q = \text{tp}^{\Delta_0}(a_0/a'B)$ be the type in $S^{\Delta_0}(a'B)$ realized by a_0 . Take i and $\bar{b} = (b_1, \dots, b_d) \in B^d$ such that (5.1) holds for all φ and all b . Set $p = \text{tp}^{\Delta_i}(a'/B\bar{b})$. Then we may take $j \in J_i$ and $c_1, \dots, c_n \in B$ such that for each φ and each b , the equivalence (5.2) holds. This yields (5.3), for all φ and b . \square

In the next corollary we assume that M is infinite (so we can meaningfully talk about VC density). We already remarked that if Δ admits a uniform definition of $\Delta(x; B)$ -types with d parameters, then $\text{vc}^*(\Delta) \leq d$; in particular, if \mathbf{M} has the VC d property then this conclusion holds for all $\Delta(x; y)$ with $|x| = 1$. The previous theorem generalizes this observation to the upper bound $\text{vc}^*(\Delta) \leq d|x|$ for all Δ . Hence:

Corollary 5.8. *If $T = \text{Th}(\mathbf{M})$ has the VC d property then $m \leq \text{vc}^T(m) \leq d \cdot m$ for every m . In particular, if T has the VC 1 property then $\text{vc}^T(m) = m$ for every m .*

In a multi-sorted setting, the VC d property can be naturally localized. Suppose \mathbf{M} is a multi-sorted structure, and let S be one of the sorts of \mathbf{M} . We say that \mathbf{M} has the VC d property in the sort S if any $\Delta(x; y)$ with x a single variable of sort S has a uniform definition of $\Delta(x; B)$ -types over finite sets with d parameters. With this definition, the following analogue of Theorem 5.7 holds (with the same proof):

Corollary 5.9. *If \mathbf{M} has the VC d property in the sort S , then every $\Delta(x; y)$ with each variable x_i of sort S has a uniform definition of $\Delta(x; B)$ -types over finite sets with $d|x|$ parameters.*

So for example, if \mathbf{M} is an infinite single-sorted structure and some expansion of \mathbf{M}^{eq} has the VC d property in the home sort (where the uniform defining formulae are in the expanded language), then $\text{vc}^T(m) \leq dm$ for each m .

5.3. Relationship to other notions. We now want to put the VC d property into perspective and compare it to two other strengthenings of the NIP concept, namely, *uniform definability of types over finite sets* and *dp-minimality*. The following notion was introduced and studied in [34]:

Definition 5.10. The structure \mathbf{M} is said to have *uniform definability of types over finite sets (UDTFS)* if every partitioned \mathcal{L} -formula has UDTFS in \mathbf{M} . Clearly UDTFS is an invariant of the elementary theory of \mathbf{M} , and so we say that an \mathcal{L} -theory T has *uniform definability of types over finite sets (UDTFS)* if every model of T does.

By Theorem 5.7, if \mathbf{M} has the VC d property, for some d , then \mathbf{M} has UDTFS. More generally, [34, Lemma 2.6] shows that if every partitioned \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$ has UDTFS in \mathbf{M} , then \mathbf{M} has UDTFS. The proof of this lemma as given in [34] in fact shows that if every partitioned \mathcal{L} -formula with a single object variable admits a uniform definition of $\varphi(x; B)$ -types with d parameters, then every partitioned \mathcal{L} -formula in the object variables x admits a uniform definition of $\varphi(x; B)$ -types with at most $(d + 1)^{|x|} - 1$ parameters. (In contrast, our bound in Theorem 5.7 is linear in $|x|$.)

Every stable formula has UDTFS; in particular, every stable theory has UDTFS. In fact, Laskowski [56] has shown that if T is stable then every partitioned \mathcal{L} -formula $\varphi(x; y)$ has UDTFS in T with $R^m(x = x, \varphi, 2)$ parameters. (See [88, Definition II.1.1] for the definition of the rank $R^m(-, -, 2)$. Stability of T is equivalent to $R^m(x = x, \varphi, 2) < \omega$ for all $\varphi(x; y)$, cf. [88, Theorem II.2.2].)

An *ICT pattern* in \mathbf{M} consists of a pair $\alpha(x; y), \beta(x; y)$ of partitioned \mathcal{L} -formulas, where $|x| = 1$, and sequences $(a_i)_{i \in \mathbb{N}}, (b_j)_{j \in \mathbb{N}}$ in $M^{|y|}$ such that for all i and j the set of $\mathcal{L}(M)$ -formulas

$$\{\alpha(x; a_i), \beta(x; b_j)\} \cup \{\neg\alpha(x; a_k) : k \neq i\} \cup \{\neg\beta(x; b_l) : l \neq j\}$$

is consistent (with \mathbf{M}). This notion and the following definition originate in [90]:

Definition 5.11. An \mathcal{L} -theory T is said to be *dp-minimal* if in no model of T there is an ICT pattern, and \mathbf{M} is *dp-minimal* if $\text{Th}(\mathbf{M})$ is dp-minimal (equivalently, if there is no ICT pattern in an elementary extension of \mathbf{M}).

The following proposition (which shows that in the definition of dp-minimality we could have restricted ourselves to ICT patterns given by identical formulas α, β) and its Corollary 5.13 are due to Dolich, Goodrick and Lippel [24]; for convenience of the reader we indicate their proofs:

Proposition 5.12. *Suppose \mathbf{M} is a monster model of the complete \mathcal{L} -theory T . Then T is dp-minimal iff there are no \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$ and sequences $(c_i)_{i \in \mathbb{N}}, (d_j)_{j \in \mathbb{N}}$ in $M^{|y|}$ such that for all i, j ,*

$$\{\varphi(x; c_i), \varphi(x; d_j)\} \cup \{\neg\varphi(x; c_k) : k \neq i\} \cup \{\neg\varphi(x; d_l) : l \neq j\}$$

is consistent.

Proof. Suppose $\alpha(x; y), \beta(x; y)$ (where $|x| = 1$) and the sequences $(a_i)_{i \in \mathbb{N}}, (b_j)_{j \in \mathbb{N}}$ are an ICT pattern in \mathbf{M} . Let $\varphi(x; y, z) := \alpha(x; y) \vee \beta(x; z)$, and for every i, j let $c_i := (a_{2i}, b_{2i})$ and $d_j := (a_{2j+1}, b_{2j+1})$. Let $a \in M$ realize the type

$$\{\alpha(x; a_{2i}), \beta(x; b_{2j+1})\} \cup \{\neg\alpha(x; a_k) : k \neq 2i\} \cup \{\neg\beta(x; b_l) : l \neq 2j + 1\}.$$

Then a satisfies $\varphi(x; c_i)$ and $\varphi(x; d_j)$. If $k \neq i$ then a satisfies $\neg\alpha(x; a_{2k})$ (since $2k \neq 2i$) and $\neg\beta(x; b_{2k})$ (since $2k \neq 2j + 1$) and hence also $\neg\varphi(x; c_k)$. Similarly we see that a satisfies $\neg\varphi(x; d_l)$ for $l \neq j$. \square

Let us tentatively say that a complete \mathcal{L} -theory T is *vc-minimal* if $\text{vc}^*(\varphi) < 2$ for every \mathcal{L} -formula $\varphi(x; y)$ with $|x| = 1$. So if $\text{vc}^T(1) < 2$ (in particular, if T is VC-minimal), then T is vc-minimal. An example of a theory T which is not VC-minimal yet satisfies $\text{vc}^T(1) = 1$ (and thus is vc-minimal) was given in [24, Proposition 3.7].

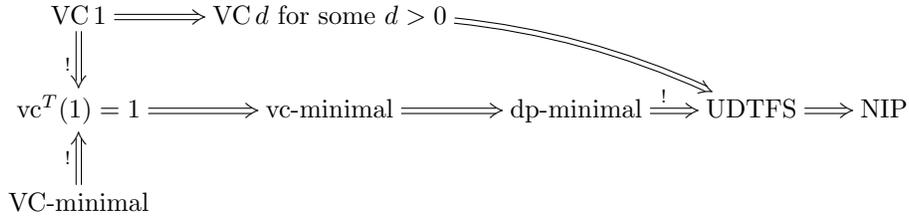
Corollary 5.13. *Every vc-minimal theory is dp-minimal.*

Proof. Suppose \mathbf{M} is a monster model of $T = \text{Th}(\mathbf{M})$, and T is not dp-minimal. Take φ and $(c_i), (d_j)$ with the properties in the previous proposition. For every n let $B_n := \{c_i : i < n\} \cup \{d_j : j < n\}$. Then $|S^\varphi(B_n)| \geq n^2 \geq \frac{1}{4}|B_n|^2$, hence $\text{vc}^*(\varphi) \geq 2$. \square

In particular, each of the theories in Examples 3.8–3.10 (being VC-minimal) is dp-minimal. Other proofs of the dp-minimality of weakly o-minimal theories can be found in [2, 24]. See also [48] for a generalization of Corollary 5.13 to a bound on “dp-rank” in terms of VC density.

A characterization of dp-minimal theories among stable theories was given in [71]. The main result of [34] is that every dp-minimal theory T has UDTFS. In particular, by Corollary 5.13, every vc-minimal T has UDTFS. (Actually, [34, Theorem 3.14] gives a more precise result: if $\varphi(x; y)$ is an \mathcal{L} -formula such that $\pi_\varphi(t) \leq \binom{t+1}{2}$ for some $t > 0$, then φ has UDTFS.)

We summarize the implications between the properties of a theory T discussed above in the following diagram:



Here the arrows marked with an exclamation mark are known not to be reversible. (For an example showing that $\text{vc}(1) = 1 \not\Rightarrow \text{VC } 1$ see [6, Example 3.15].) We do not know which of the other arrows are reversible; whether the converse of the implication $\text{UDTFS} \Rightarrow \text{NIP}$ holds was first asked by Laskowski [34, Open Question 4.1].

Recall from Corollary 3.20 that the Shelah expansion \mathbf{M}^{Sh} of \mathbf{M} has the same VC density function as \mathbf{M} . In [71] it is observed that the Shelah expansion of a dp-minimal structure is again dp-minimal. We finish this section by showing that the analogous statement also holds for the VC d property; in fact, we have more generally:

Proposition 5.14. *Suppose every finite set of partitioned \mathcal{L} -formulas in m object variables has UDTFS in T with d parameters. Then every finite set of partitioned \mathcal{L}^{Sh} -formulas in m object variables has UDTFS in T^{Sh} with d parameters.*

Proof. As in the definition of the Shelah expansion (cf. Section 3.5) let \mathbf{M}^* be a very saturated elementary extension of \mathbf{M} . Let $\Delta = \Delta(x; y)$ be a finite set of partitioned \mathcal{L}^{Sh} -formulas where $m = |x|$; below φ ranges over Δ . We need to show that Δ has UDTFS in T^{Sh} with d parameters. For this, by Lemma 5.5 and since T^{Sh} admits quantifier elimination, we may assume that each of the \mathcal{L}^{Sh} -formulas in Δ is atomic; that is, there exist \mathcal{L} -formulas $\psi_\varphi(x, y; z)$, one for each φ , and a tuple $c \in (M^*)^{|z|}$ such that each φ

has the form $\varphi(x; y) = R_{\psi_\varphi, c}(x; y)$. Let now $\Psi(x; y, z) := \{\psi_\varphi(x; y, z) : \varphi \in \Delta\}$ and take a uniform definition $\mathcal{G} = (\mathcal{G})_{i \in I}$ of $\Psi(x; B^*)$ -types over finite sets in T^{Sh} , where

$$\mathcal{G}_i = ((\psi_\varphi)_i(y, z; (y_1, z_1), \dots, (y_d, z_d)))_{\varphi \in \Delta} \quad \text{for each } i \in I.$$

For each $i \in I$ set

$$\mathcal{F}_i := (R_{\varphi_i, c}(y; y_1, \dots, y_d))_{\varphi \in \Delta}$$

$$\text{where } \varphi_i(y, y_1, \dots, y_d; z) := (\psi_\varphi)_i(y, z, (y_1, z), \dots, (y_d, z)).$$

We claim that $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ is a uniform definition of $\Delta(x; B)$ -types over finite sets in T^{Sh} . To see this let $p \in S^\Delta(B)$ where $B \subseteq M^{|y|}$ is finite, and let $a \in M^m$ be a realization of p in M^{Sh} . Put $B^* := B \times \{c\} \subseteq M^{|y|} \times (M^*)^{|z|}$ and let $p^* := \text{tp}^\Psi(a/B^*)$ (in M^*). Take $b_1, \dots, b_d \in B$ and $i \in I$ such that $\mathcal{G}_i(y, z; (b_1, c), \dots, (b_d, c))$ defines p^* ; then for every φ and $b \in B$ we have

$$\begin{aligned} M^{\text{Sh}} \models \varphi(a; b) &\iff M^* \models \psi_\varphi(a; b, c) \\ &\iff M^* \models (\psi_\varphi)_i(b, c; (b_1, c), \dots, (b_d, c)) \\ &\iff M^{\text{Sh}} \models R_{\varphi_i, c}(b; b_1, \dots, b_d). \end{aligned}$$

That is, $\mathcal{F}_i(y; b_1, \dots, b_d)$ defines p (in M^{Sh}). \square

6. EXAMPLES OF VC d : WEAKLY O-MINIMAL THEORIES AND VARIANTS

In this section we apply Theorem 5.7 from the preceding section to give a proof of Theorem 1.1 on VC density in weakly o-minimal theories from the introduction. We also observe that a similar technique allows us to treat all (weakly) quasi-o-minimal theories.

Throughout this section \mathcal{L} is a language containing a binary relation symbol “ $<$ ” and T is a theory extending the theory of infinite linear orderings.

6.1. Weakly o-minimal theories. We begin by introducing some terminology concerning ordered sets. Let $(X, <)$ be a linearly ordered set, and let S be a subset of X which is a union of finitely many non-empty convex subsets of X . We refer to the convex sets in the unique minimal such presentation of S as its (*convex*) *components*. Suppose S has N convex components, where $N > 0$. These components are ordered by $<$, so for $i = 1, \dots, N$ we can refer to the i th component of S ; for $i > N$ we declare the i th component of S to be equal to the N th.

Recall that T is called *weakly o-minimal* if for any $M \models T$, any definable subset of M is a finite union of convex subsets of M .

Theorem 6.1. *Assume that T is weakly o-minimal. Then T has the VC1 property, and hence any finite set $\Delta(x; y)$ of \mathcal{L} -formulas has dual VC density at most $|x|$.*

Proof. Let $M \models T$. Fix a finite non-empty set of \mathcal{L} -formulas $\Delta(x; y)$ with $|x| = 1$. We let φ range over Δ and b over $M^{|y|}$. By the weak o-minimality of T , there is an integer $N > 0$ such that for any φ and any b , $\varphi(M; b)$ has at most N components. For any φ and $i \in [N]$ there is an \mathcal{L} -formula $\varphi^i(x; y)$ such that for every b with $\varphi(M; b) \neq \emptyset$, the i th component of $\varphi(M; b)$ equals $\varphi^i(M; b)$, and such that for every b with $\varphi(M; b) = \emptyset$ we have $\varphi^i(M; b) = \emptyset$. So

$$\varphi(M; b) = \varphi^1(M; b) \cup \dots \cup \varphi^N(M; b) \quad \text{for every } b. \quad (6.1)$$

Set

$$\begin{aligned}\varphi_{\leq}^i(x; y) &:= \exists x_0(\varphi^i(x_0; y) \wedge x \leq x_0), \\ \varphi_{<}^i(x; y) &:= \forall x_0(\varphi^i(x_0; y) \rightarrow x < x_0).\end{aligned}$$

Then clearly

$$\varphi^i(M; b) = \varphi_{\leq}^i(M; b) \cap (M \setminus \varphi_{<}^i(M; b)) \quad \text{for all } \varphi, b \text{ and } i \in [N]. \quad (6.2)$$

Now set

$$\Psi(x; y) := \{\varphi_{\square}^i(x; y) : \varphi \in \Delta, i \in [N], \square \in \{\leq, <\}\}.$$

For each $\psi \in \Psi$ and each b , the set $\psi(M; b)$ is an initial segment of M ; hence \mathcal{S}_{Ψ} is directed, so Ψ has UDTFs with a single parameter, by Lemma 5.2. Moreover, by (6.1) and (6.2), every φ is equivalent to a Boolean combination of $2N$ formulas from Ψ . Hence Δ also has UDTFs with a single parameter, by Lemma 5.5. Thus \mathbf{M} has the VC1 property. By Corollary 5.8 therefore $\text{vc}^*(\Delta) \leq |x|$ for every finite set $\Delta(x; y)$ of \mathcal{L} -formulas. \square

Remark. In the previous theorem we assume that the theory T is weakly o-minimal (i.e., all models of T are weakly o-minimal). Recall that a weakly o-minimal structure need not have weakly o-minimal theory. We do not know whether the conclusion of Theorem 6.1 holds if T is merely assumed to have *some* weakly o-minimal model.

Put $\mathcal{L}_{\text{div}, <} := \mathcal{L}_{\text{div}} \cup \{<\}$, where $\mathcal{L}_{\text{div}} = \{0, 1, +, -, \times, |\}$ is the language of rings expanded by a divisibility predicate (see Example 3.10 above) and “ $<$ ” is a binary relation symbol. Let RCVF denote the theory of real closed fields equipped with a proper convex valuation ring, parsed in the language $\mathcal{L}_{\text{div}, <}$. The following corollary is now immediate, as by [23], RCVF is weakly o-minimal.

Corollary 6.2. *Let $K \models \text{RCVF}$. Then any finite set $\Delta(x; y)$ of $\mathcal{L}_{\text{div}, <}$ -formulas has dual VC density at most $|x|$ in K .*

This result in turn yields a VC density bound for algebraically closed valued fields of residue characteristic 0 (which is non-optimal by Example 3.10):

Corollary 6.3. *Let $\text{ACVF}_{(0,0)}$ be the theory of non-trivially valued algebraically closed fields of residue characteristic 0, in the language \mathcal{L}_{div} . Let $\Delta(x; y)$ be a finite set of \mathcal{L}_{div} -formulas. Then Δ has dual VC density at most $2|x|$ in $\text{ACVF}_{(0,0)}$.*

Proof. The theory $\text{ACVF}_{(0,0)}$, which is complete, is interpretable in RCVF: if K is a model of RCVF, then its algebraic closure K^{alg} is a degree 2 extension of K : $K^{\text{alg}} = K(i)$, where $i^2 = -1$. So K^{alg} can be identified with K^2 , and the valuation v of K can be definably extended to one of K^{alg} by setting $v(a + bi) = \frac{1}{2}v(a^2 + b^2)$ for $a, b \in K$. Thus, by Lemma 3.16 and Corollary 6.2, $\Delta(x; y)$ has dual VC density at most $2|x|$. \square

We have no results on VC density for ACVF in characteristics other than $(0, 0)$.

6.2. Quasi-o-minimal theories. We now turn to quasi-o-minimal theories: T is said to be *quasi-o-minimal* if for any $\mathbf{M} \models T$, any definable subset of M is a finite Boolean combination of singletons, intervals in M , and \emptyset -definable sets. (See [14].)

Theorem 6.4. *Assume that T is quasi-o-minimal. Then T has the VC1 property, and hence $\text{vc}^T(n) = n$ for each n .*

Proof. Let $M \models T$. Fix a finite set $\Delta(x; y)$ of \mathcal{L} -formulas with $|x| = 1$; we let φ range over Δ and b over $M^{|y|}$. There is some positive integer N and \emptyset -definable subsets D_1, \dots, D_N of M so that for any φ and any choice of parameters b , the set $\varphi(M; b)$ of realizations of $\varphi(x; b)$ is a Boolean combination of the D_i and at most N singletons and intervals in M [14, Theorem 3].

Let \mathcal{D} be the collection of sets of the form $\tilde{D}_1 \cap \dots \cap \tilde{D}_N$, where \tilde{D}_i is either D_i or its complement in M (so \mathcal{D} is a partition of M into at most 2^N sets). We let D range over \mathcal{D} . For every D , φ , and b , the set $D \cap \varphi(M; b)$ is then a finite union of at most N convex subsets of the ordered set D . For every $i \in [N]$ and every D let $\varphi^{i,D}(x; u)$ be an \mathcal{L} -formula such that for every b , if the set $D \cap \varphi(M; b)$ is non-empty, then the i th convex component of $D \cap \varphi(M; b)$ (viewed as a subset of the ordered set D) is given by $\varphi^{i,D}(M; b)$, and if $D \cap \varphi(M; b) = \emptyset$ then $\varphi^{i,D}(M; b) = \emptyset$. Hence for each φ and b we have

$$\varphi(M; b) = \bigcup_{D \in \mathcal{D}, i \in [N]} \varphi^{i,D}(M; b).$$

Now let (slightly abusing syntax)

$$\begin{aligned} \varphi_{\leq}^{i,D}(x; y) &:= x \in D \wedge \exists x_0 (x_0 \in D \wedge \varphi^i(x_0; y) \wedge x \leq x_0), \\ \varphi_{<}^{i,D}(x; y) &:= x \in D \wedge \forall x_0 (x_0 \in D \wedge \varphi^i(x_0; y) \rightarrow x < x_0). \end{aligned}$$

Then

$$\varphi^{i,D}(M; b) = \varphi_{\leq}^{i,D}(M; b) \cap (M \setminus \varphi_{<}^{i,D}(M; b)) \quad \text{for all } \varphi, b, D \text{ and } i \in [N].$$

Each set $\varphi_{\square}^{i,D}(x; b)$, where $\square \in \{\leq, <\}$, is an initial segment of D , and any two distinct elements of \mathcal{D} are disjoint. Thus the set system \mathcal{S}_{Ψ} , where

$$\Psi(x; y) = \{\varphi_{\square}^{i,D}(x; y) : \varphi \in \Delta, i \in [N], \square \in \{\leq, <\}, D \in \mathcal{D}\},$$

is directed. As in the proof of Theorem 6.1 it now follows that Δ has UDTFs with a single parameter. \square

Corollary 6.5. *The following structures all have the VC1 property:*

- (1) $(\mathbb{R}, <, \mathbb{Q})$ (i.e., the ordered set of reals with a predicate for the rationals);
- (2) $(\mathbb{Z}^n, <, +)$ where $<$ is the lexicographic ordering on \mathbb{Z}^n ;
- (3) $(\mathbb{Z}^n \times \mathbb{Q}, <, +)$ where $<$ is the lexicographic ordering on $\mathbb{Z}^n \times \mathbb{Q}$.

Proof. Each of the examples has quasi-o-minimal theory: For (1) this was noted in [14, Section 1], and for (2) and (3) this is proved (based on a quantifier-elimination result from [99]) in [15, Theorem 15]. The corollary now follows from Theorem 6.4. \square

Remark. The ordered abelian groups in (2) and (3) of the previous corollary are typical for quasi-o-minimal groups. Here and below, “quasi-o-minimal group” means “quasi-o-minimal expansion of an ordered group.” (A quasi-o-minimal group is necessarily abelian [14, Theorem 11].) An expansion \mathbf{G} of an ordered group is called *coset-minimal* if every subset of G definable in \mathbf{G} is a finite union of cosets of definable subgroups, intersected with intervals. A theory expanding the theory of ordered groups is said to be coset-minimal if all its models are. (See [15].) Now by [79, Theorem 5.3], the theory of \mathbf{G} is coset-minimal iff the theory of the expansion of \mathbf{G} by constant symbols for the elements of G is quasi-o-minimal, and in this case \mathbf{G} is an expansion of an ordered group elementarily equivalent to either $(\mathbb{Z}^n, <, +)$ or $(\mathbb{Z}^n \times \mathbb{Q}, <, +)$, for some n .

Part (2) of the previous corollary shows in particular that Presburger Arithmetic, i.e., the theory of the ordered group $(\mathbb{Z}, <, +)$ of integers, has the VC1 property, and hence is dp-minimal. By Theorem 10 in [14], $(\mathbb{Z}, <, +)$ has no proper quasi-o-minimal expansions. One can strengthen this statement:

Proposition 6.6. *No proper expansion of $(\mathbb{Z}, <, +)$ is dp-minimal.*

The proof is the same as in [14], replacing the use of [14, Theorem 7] by a result from [92]; we state the latter employing some convenient terminology from [14]: Let $(X, <)$ be a linearly ordered set. We say that two subsets S, T of X are *eventually equal* (in symbols: $S \approx T$) if there is some $a \in X$ such that $S \cap (a, +\infty) = T \cap (a, +\infty)$. Clearly \approx is an equivalence relation on subsets of X . We say that a family of subsets of X is *eventually finite* if it is partitioned into finitely many classes by \approx .

Lemma 6.7 (Simon [92, Lemma 2.9]). *Suppose T is dp-minimal. Let $\mathbf{M} \models T$ and let $\varphi(x; y)$ be a partitioned $\mathcal{L}(\mathbf{M})$ -formula where $|x| = 1$. Then \mathcal{S}_φ is eventually finite.*

For the benefit of the reader we now indicate the details of the proof of Proposition 6.6. Let \mathcal{Z} be a proper expansion of $(\mathbb{Z}, <, +)$. By a theorem of Michaux and Villemaire [69] (and an easy extra argument, given in the proof of [14, Theorem 10]), there is a subset U of \mathbb{Z} which is definable in \mathcal{Z} but not definable in $(\mathbb{Z}, <, +)$. By Simon's lemma, the family $\{a + U : a \in \mathbb{Z}\}$ is eventually finite; thus the subgroup A of \mathbb{Z} consisting of all $a \in \mathbb{Z}$ such that $a + U \approx U$ is non-zero, so $A = a\mathbb{Z}$ for some positive integer a . Let V be the union of all cosets of A which contain arbitrarily large elements of U ; then V is definable in $(\mathbb{Z}, <, +)$, hence it suffices to show that $U \approx V$. As $a + U \approx U$, we can take $\alpha \in \mathbb{Z}$ such that for every $u \in \mathbb{Z}$ with $u \geq \alpha$ we have $a + u \in U \iff u \in U$. One now proves easily that for every $u \in \mathbb{Z}$ with $u \geq \alpha$ we have $u \in U \iff u \in V$. \square

So for example, the expansion of $(\mathbb{Z}, <, +)$ by the set $b^{\mathbb{N}} = \{b^n : n \geq 0\}$ of powers of a natural number $b > 1$ is not dp-minimal, as is the expansion of $(\mathbb{Z}, <, +)$ by the set of factorials or by the set of Fibonacci numbers. In all these examples, the corresponding expansion of $(\mathbb{Z}, <, +)$ has quantifier elimination in a natural expansion of $\{<, +, U\}$ (see [19, 78]) and is NIP (as will be shown elsewhere).

6.3. Weakly quasi-o-minimal theories. In [53], T is called *weakly quasi-o-minimal* if for any $\mathbf{M} \models T$, any definable subset of M is a finite Boolean combination of convex subsets of M and \emptyset -definable sets. Every weakly quasi-o-minimal theory is NIP [53, Theorem 2.3]. In fact, the proof of Theorem 6.4 (mutatis mutandis) also shows more generally:

Theorem 6.8. *All weakly quasi-o-minimal theories have the VC1 property.*

This observation can be used to strengthen [92, Proposition 4.2], where it is shown that the complete theories of colored linearly ordered sets with monotone relations (shown to be NIP in [85]) are dp-minimal. A binary relation R on a set X is said to be monotone with respect to a linear ordering $<$ of X if

$$x' \leq xRy \leq y' \Rightarrow x'Ry' \quad \text{for all } x, x', y, y' \in X.$$

A *colored linearly ordered set with monotone relations* is a structure of the form $\mathbf{M} = (M, <, \{C_i\}_{i \in I}, \{R_j\}_{j \in J})$ where $<$ is a linear ordering on M , the C_i are unary predicates ("colors"), and the R_j are binary relations which are monotone (with respect to $<$). It was shown by Simon [92, Proposition 4.1] that every such colored linearly ordered set with monotone relations has quantifier elimination provided that each \emptyset -definable subset

of M is given by one of the predicates C_i and each monotone \emptyset -definable binary relation is given by one of the R_j .

Proposition 6.9. *Let \mathbf{M} be a colored linearly ordered set with monotone relations as above. Then $T = \text{Th}(\mathbf{M})$ is weakly quasi-o-minimal, and hence has the VC1 property.*

Proof. By the result of Simon just quoted, we may assume that \mathbf{M} admits quantifier elimination. Now for each $b \in M$ and $j \in J$ the set

$$\{x \in M : \mathbf{M} \models xR_jb\}$$

is an initial segment of M , and

$$\{x \in M : \mathbf{M} \models bR_jx\}$$

is a final segment of M (i.e., its complement is an initial segment of M). Hence any definable subset of M is a finite Boolean combination of initial segments of M and \emptyset -definable sets, so T is weakly quasi-o-minimal. \square

As in [85, 92] this leads to a result for (partially) ordered sets of finite width. Let $\mathbf{P} = (P, <)$ be an ordered set, i.e., a set P equipped with an irreflexive, asymmetric and transitive binary relation $<$ on P . A subset A of P is an antichain if for all $a \neq a'$ in A , neither $a < a'$ nor $a' < a$ holds, and $C \subseteq P$ is a chain if for all $c \neq c'$ in C , either $c < c'$ or $c' < c$. The *width* of \mathbf{P} is defined to be the supremum of the cardinalities of antichains in P , and denoted by $\text{width}(\mathbf{P})$. (Dually, the *height* of \mathbf{P} is defined to be the supremum of the cardinalities of a chain in P , denoted by $\text{height}(\mathbf{P})$.) A *colored ordered set* is a structure $\mathbf{P} = (P, <, (C_i)_{i \in I})$ where $(P, <)$ is an ordered set and each C_i is a unary predicate.

Corollary 6.10. *Let $\mathbf{P} = (P, <, (C_i)_{i \in I})$ be an infinite colored ordered set of finite width. Then $\text{vc}^{\text{Th}(\mathbf{P})}(m) = m$ for every m .*

Proof. Let $n = \text{width}(\mathbf{P})$ and let i, j range over $[n]$. By Dilworth's Theorem there is a partition $P = P_1 \cup \dots \cup P_n$ of P into disjoint chains P_i . Define a linear ordering \preceq on P by setting $a \preceq b$ iff either $a, b \in P_i$ for some i and $a \leq b$, or $a \in P_i, b \in P_j$ with $i < j$. For all i, j the binary relation

$$R_{ij} := \{(a, b) \in P : \exists a' \in P_i, b' \in P_j : a \preceq a' \leq b' \preceq b\}$$

is monotone with respect to \preceq , and the original ordering $<$ is \emptyset -definable in the linearly ordered set with monotone relations $(P, \preceq, (R_{ij})_{i,j})$, noting that $a \leq b$ iff $aR_{ii}aR_{ij}bR_{jj}b$ for some i and j . The claim now follows immediately from Proposition 6.9. \square

Question. Is every ordered set of finite width VC1?

By an interpretability argument, the previous corollary also leads to a (perhaps non-optimal) bound on the VC density for those distributive lattices with NIP theory. By [85, Theorem 6] these are exactly the distributive lattices of finite breadth. From Section 2.4 recall that a semilattice (L, \wedge) has breadth at most d if for all $b_1, \dots, b_{d+1} \in L$ there is some $i \in [d+1]$ such that $b_1 \wedge \dots \wedge b_{d+1} = b_1 \wedge \dots \wedge \widehat{b_i} \wedge \dots \wedge b_{d+1}$, and the smallest such d (if it exists) is called the breadth of (L, \wedge) .

Corollary 6.11. *Let $\mathbf{L} = (L, \wedge, \vee)$ be an infinite distributive lattice of breadth d . Then $\text{vc}^{\text{Th}(\mathbf{L})}(m) \leq dm$ for every m .*

Proof. Let $\mathbf{P} = (P, <)$ be an ordered set of width d , and let $\mathcal{A}(\mathbf{P})$ be the set of antichains of \mathbf{P} (so each element of $\mathcal{A}(\mathbf{P})$ is a subset of P of size $\leq d$). For $A, A' \in \mathcal{A}(\mathbf{P})$ let $A \wedge A'$ denote the set of minimal elements of $A \cup A'$ and $A \vee A'$ the set of maximal elements of $A \cup A'$; then $A \wedge A', A \vee A' \in \mathcal{A}(\mathbf{P})$, and $(\mathcal{A}(\mathbf{P}), \wedge, \vee)$ is a distributive lattice of breadth d . Moreover, one can choose \mathbf{P} such that the given distributive lattice \mathbf{L} is isomorphic to $\mathcal{A}(\mathbf{P})$ [84, Theorem 3]. Since $\mathcal{A}(\mathbf{P})$ is interpretable in \mathbf{P} on a definable subset of P^d , we have $\text{vc}^{\text{Th}(\mathbf{L})}(m) \leq \text{vc}^{\text{Th}(\mathbf{P})}(dm)$ by Corollary 3.17 and hence $\text{vc}^{\text{Th}(\mathbf{L})}(m) \leq dm$ by the previous corollary. \square

In [92] it is shown that the complete theory of each infinite tree \mathbf{T} (viewed as an ordered set) is dp-minimal. Here, a tree is an ordered set $\mathbf{T} = (T, <)$ with the property that for each $t \in T$ the set $\{t' \in T : t' < t\}$ is linearly ordered (by the restriction of $<$).

Problem. Determine the VC density function of each (infinite) tree.

(It is known [75] that a tree \mathbf{T} is stable iff \mathbf{T} has finite height, and then \mathbf{T} is superstable of U-rank $\leq \text{height}(\mathbf{T})$, so conceivably, the methods of [6] could be applied.)

7. A STRENGTHENING OF VC d , AND P -ADIC EXAMPLES

In this section we first introduce a strengthening of the VC d property defined and studied in Section 5, and we prove a more precise version of Theorem 5.7 for strong VC d structures. The extra precision afforded by this theorem is useful in situations where $\text{vc}(1) = 1$, yet we can only prove the VC d property for some $d > 1$. This is the case for P -minimal theories, which are discussed in the last subsection, where we prove Theorem 1.2 from the introduction.

7.1. The strong VC d property. In the following \mathbf{M} is a structure in a language \mathcal{L} and $\Delta = \Delta(x; y)$ is a finite non-empty set of partitioned \mathcal{L} -formulas. Let $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ be a uniform definition of $\Delta(x; B)$ -types over finite sets, where

$$\mathcal{F}_i = (\varphi_i(y; y_1, \dots, y_d))_{\varphi \in \Delta} \quad (i \in I).$$

Recall from Definition 5.1 above that this means the following: for every finite set $B \subseteq M^{|y|}$ and $q \in S^\Delta(B)$ there are $b_1, \dots, b_d \in B$ and some $i \in I$ such that $\mathcal{F}_i(y; b_1, \dots, b_d)$ defines q . If in addition for *every* choice of $b_1, \dots, b_d \in M^{|y|}$ and $i \in I$, the set

$$p_i(x; b_1, \dots, b_d) := \{\varphi(x; b) : \varphi \in \Delta, b \in M^{|y|}, \mathbf{M} \models \varphi_i(b; b_1, \dots, b_d)\}$$

of $\mathcal{L}(M)$ -formulas is consistent (with \mathbf{M}), then we say that \mathcal{F} is a *coherent* definition of $\Delta(x; B)$ -types over finite sets. (In this case, every restriction of $p_i(x; b_1, \dots, b_d)$ to a finite parameter set $B \subseteq M^{|y|}$ extends to a complete $\Delta(x; B)$ -type, but $\mathcal{F}_i(y; b_1, \dots, b_d)$ does not in general define such an extension.)

Remark. Often all our defining formulas φ_i have the syntactic form

$$\varphi_i(y; y_1, \dots, y_d) = \forall x (\chi_i(x; y_1, \dots, y_d) \rightarrow \varphi(x; y))$$

where χ_i is an \mathcal{L} -formula. In this case, the coherency condition for \mathcal{F} is automatically satisfied provided $\chi_i^{\mathbf{M}}(x; \bar{b}) \neq \emptyset$ for all $i \in I$ and $\bar{b} \in (M^{|y|})^d$. For example if \mathcal{S}_Δ has breadth d and is d -consistent, then Δ has a coherent definition of $\Delta(x; B)$ -types over finite sets with d parameters. (See Lemma 5.2.)

We say that \mathbf{M} has the *strong VCd property* if any $\Delta(x; y)$ with $|x| = 1$ has a coherent definition of $\Delta(x; B)$ -types over finite sets with d parameters. Clearly, the strong VCd property is a property of the elementary theory of \mathbf{M} . We say that a theory T has the *strong VCd property* if every model of T has the strong VCd property.

Remark. Suppose $|x| = 1$ and $\mathcal{F} = (\mathcal{F}_i)_{i \in I}$ is a coherent definition of $\Delta(x; B)$ -types over finite sets in \mathbf{M} , where $\mathcal{F}_i = (\varphi_i)_{\varphi \in \Delta}$. Then for every $\Delta' \subseteq \Delta$, the families $\mathcal{F}'_i := (\varphi_i)_{\varphi \in \Delta'}$ form a coherent definition of $\Delta'(x; B)$ -types over finite sets in \mathbf{M} . This shows in particular that in order to check that \mathbf{M} has the strong VCd property, one may restrict oneself to sets of \mathcal{L} -formulas Δ which are closed under negation.

In the rest of this subsection we assume that M is infinite. We have the following result on counting types in structures with the strong VCd property:

Theorem 7.1. *Suppose that \mathbf{M} has the strong VCd property, and let $r \in \mathbb{R}$ such that*

$$\pi_{\Delta}^*(t) = O(t^r) \quad \text{for every } \Delta(x; y) \text{ with } |x| = 1.$$

Then we have

$$\pi_{\Delta}^*(t) = O(t^{d(|x|-1)+r}) \quad \text{for every } \Delta(x; y).$$

The proof of Theorem 7.1 proceeds by counting types via sets of representatives (and with an induction supported by the strong VCd property): given a finite set $B \subseteq M^{|y|}$, we say that $R \subseteq M^{|x|}$ is a *set of representatives for $S^{\Delta}(B)$* if for every $q \in S^{\Delta}(B)$ there is $\alpha \in R$ realizing q . Equivalently, for every $a \in M^{|x|}$ there is $\alpha \in R$ such that for every $\varphi \in \Delta$ and every $b \in B$, $\mathbf{M} \models \varphi(a; b)$ if and only if $\mathbf{M} \models \varphi(\alpha; b)$. Thus $|S^{\Delta}(B)| \leq K$ iff there is a set of representatives for $S^{\Delta}(B)$ of size at most K .

Proof. The proof is similar to that of Theorem 5.7. We again induct on $m = |x|$, with the case $m = 1$ holding by hypothesis. For the inductive step write $x = (x_0, x')$ where $x' = (x_1, \dots, x_m)$, and let $\Delta(x; y)$ be given. We may assume that Δ is closed under negation. As in the proof of Theorem 5.7 let

$$\Delta_0(x_0; x', y) = \{\varphi(x_0; x', y) : \varphi(x; y) \in \Delta\}.$$

By the strong VCd property applied to Δ_0 , we can take finitely many families

$$\mathcal{F}_i = (\varphi_i(x', y; y_1, \dots, y_d))_{\varphi \in \Delta} \quad (i \in I)$$

of \mathcal{L} -formulas with the following two properties: for any $a' \in M^m$, any finite $B \subseteq M^{|y|}$ and any $q \in S^{\Delta_0}(a'B)$, there are $b_1, \dots, b_d \in B$ and $i \in I$ such that $\mathcal{F}_i(a', y; b_1, \dots, b_d)$ defines q , i.e., for all $\varphi \in \Delta$, $b \in B$:

$$\varphi(x_0; a', b) \in q \quad \iff \quad \mathbf{M} \models \varphi_i(a', b; b_1, \dots, b_d); \quad (7.1)$$

and for all $i \in I$, $a' \in M^m$ and $\bar{b} \in (M^{|y|})^d$, the set

$$p_i(x_0; a', \bar{b}) := \{\varphi(x_0; a', b) : \varphi \in \Delta, b \in M^{|y|}, \mathbf{M} \models \varphi_i(a', b; \bar{b})\}$$

of $\mathcal{L}(M)$ -formulas is consistent (with \mathbf{M}). In the rest of this proof let φ range over Δ and i over I . For each i , let

$$\Delta_i(x'; y, y_1, \dots, y_d) = \{\varphi_i(x'; y, y_1, \dots, y_d) : \varphi(x; y) \in \Delta\}$$

and apply the inductive hypothesis to each Δ_i . Thus there are constants K_i such that for any finite $C \subseteq (M^{|y|})^{(d+1)}$ there is a set of representatives for $S^{\Delta_i}(C)$ of size at most $K_i|C|^{d(m-1)+r}$.

Now let a finite $B \subseteq M^{|y|}$ be given. We let b range over B and $\bar{b} = (b_1, \dots, b_d)$ over B^d . For each \bar{b} and each i , let $R_i(B\bar{b})$ be a set of representatives for $S^{\Delta_i}(B\bar{b})$. Thus for any $a' \in M^m$ and i there is some $\alpha \in R_i(B\bar{b})$ such that for any φ and b ,

$$\mathbf{M} \models \varphi_i(a'; b, \bar{b}) \iff \mathbf{M} \models \varphi_i(\alpha; b, \bar{b}). \quad (7.2)$$

Notice that there are $|B|^d$ sequences \bar{b} , and $|B\bar{b}| = |B|$ for each \bar{b} . As above, we may suppose $|R_i(B\bar{b})| \leq K_i |B\bar{b}|^{d(m-1)+r} = K_i |B|^{d(m-1)+r}$.

For each i , given $\alpha \in M^m$ and \bar{b} , let $\delta_{i,\alpha,\bar{b}} \in M$ realize the restriction of the type $p_i(x_0; \alpha, \bar{b})$ to the (finite) parameter set αB . Let

$$R_\Delta = \{(\delta_{i,\alpha,\bar{b}}, \alpha) : i \in I, \bar{b} \in B^d, \alpha \in R_i(B\bar{b})\}$$

and observe that

$$|R_\Delta| \leq \sum_i |B|^d K_i |B|^{d(m-1)+r} = \left(\sum_i K_i \right) |B|^{md+r}.$$

Thus we are finished once we have shown that R_Δ is a set of representatives for $S^\Delta(B)$. For this, let $a = (a_0, a') \in M^{1+m}$ be given; we need to show that there is an element $(\delta_{i,\alpha,\bar{b}}, \alpha) \in R_\Delta$ such that for every φ and every b ,

$$\mathbf{M} \models \varphi(a; b) \iff \mathbf{M} \models \varphi(\delta_{i,\alpha,\bar{b}}, \alpha; b). \quad (7.3)$$

Let $q = \text{tp}^{\Delta_0}(a_0/a'B) \in S^{\Delta_0}(a'B)$. Take i and \bar{b} such that (7.1) holds for all φ and all b , and then take a representative $\alpha \in R_i(B\bar{b})$ for $\text{tp}^{\Delta_i}(a'/B\bar{b})$. Note that by (7.1) and since Δ is assumed to be closed under negation, for each φ and b , either $\mathbf{M} \models \varphi_i(a', b; \bar{b})$ or $\mathbf{M} \models \psi_i(a', b; \bar{b})$, where $\psi \in \Delta$ is equivalent to $\neg\varphi$; hence also either $\mathbf{M} \models \varphi_i(\alpha; b, \bar{b})$ or $\mathbf{M} \models \psi_i(\alpha; b, \bar{b})$, by (7.2), and thus

$$\mathbf{M} \models \varphi_i(\alpha; b, \bar{b}) \iff \mathbf{M} \models \varphi(\delta_{i,\alpha,\bar{b}}, \alpha; b), \quad (7.4)$$

by choice of $\delta_{i,\alpha,\bar{b}}$. Combining (7.1), (7.2) and (7.4) now yields (7.3) as required. \square

In each of the cases treated in Theorems 6.1 and 6.4 one can show that the theory in question has, indeed, the *strong* VC1 property. Since this is of limited interest for computing VC density (the VC1 property already gives the optimal result $\text{vc}(m) = m$), we do not give the details, and instead now turn to an application of Theorem 7.1 to p -adic examples.

7.2. P -minimal theories. Let $\mathcal{L}_{\text{rings}}$ be the language of rings, let $\mathcal{L}_p = \mathcal{L}_{\text{rings}} \cup \{P_n : n > 1\}$ where the P_n are unary predicates, and let \mathcal{L} be a language containing \mathcal{L}_p . Here and below, p is a fixed prime number. Let $p\text{CF}$ denote the \mathcal{L}_p -theory of \mathbb{Q}_p , where each P_n is interpreted as the set of n th powers in \mathbb{Q}_p :

$$\mathbb{Q}_p \models \forall x (P_n(x) \leftrightarrow \exists y (y^n = x)).$$

By a theorem of Macintyre [60], $p\text{CF}$ has elimination of quantifiers. Following [39], an \mathcal{L} -theory T containing $p\text{CF}$ is called *P -minimal* if, in every model of T , every definable subset in one variable is quantifier-free definable just using the language \mathcal{L}_p . (In fact, the setting of [39] also allowed for p -adically closed fields of arbitrary fixed p -rank, and our methods here could be adjusted to that.)

By [26], a motivating example of a P -minimal theory is the theory $p\text{CF}_{\text{an}}$ first investigated by Denef and van den Dries [22]. This is the theory of the p -adic numbers equipped with, for every $n > 0$ and power series $\sum_\nu a_\nu X^\nu \in \mathbb{Q}_p[[X]]$ such that $|a_\nu| \rightarrow 0$ as $|\nu| \rightarrow \infty$, a function symbol f of arity n taking value identically zero off \mathbb{Z}_p^n , and

such that $f(x) = \sum_{\nu} a_{\nu} x^{\nu}$ for all $x \in \mathbb{Z}_p^n$. (Here, $X = (X_1, \dots, X_n)$, $|a|$ denotes the p -adic norm of $a \in \mathbb{Q}_p$, $\nu = (\nu_1, \dots, \nu_n) \in \mathbb{N}^n$ is a multi-index, $|\nu| = \nu_1 + \dots + \nu_n$, and $x^{\nu} = x_1^{\nu_1} \dots x_n^{\nu_n}$.)

Our main result about VC density in P -minimal theories is:

Theorem 7.2. *Let T be a P -minimal \mathcal{L} -theory with definable Skolem functions. Then T has the strong VC2 property, and any finite set $\Delta(x; y)$ of \mathcal{L} -formulas has dual VC density at most $2|x| - 1$.*

Before proving this theorem, we introduce some notation and establish some auxiliary facts. We fix a model K of p CF, with valuation $v: K \rightarrow \Gamma_{\infty}$. We view \mathbb{Z} as a convex subgroup of Γ , by identifying 1 with $v(p)$. In the following, by a *ball* in K we always mean a closed ball, i.e., a set of the form

$$B = B_{\rho}(a) = \{x \in K : v(x - a) \geq \rho\} \quad \text{where } c \in K \text{ and } \rho \in \Gamma.$$

Its *radius*, denoted $\text{rad}(B)$, is ρ . By convention $\text{rad}(K) := -\infty$. Let \mathcal{B} denote the set of all balls in K . There is a natural semilinear partial order on \mathcal{B} , with $B \leq B'$ if and only if $B \supseteq B'$. A ball $B = B_{\rho}(a)$ as above has a unique immediate predecessor, namely $B_{\rho-1}(a)$, and p immediate successors, namely $B_{\rho+1}(a_i)$ where $a_i = a + ir$ for $i = 0, \dots, p-1$; here r is an arbitrary element of K with $v(r) = \rho$. Thus, if we form a graph with vertex set \mathcal{B} , with vertices B, B' adjacent if and only if one of B, B' is an immediate successor of the other in the partial order, each of its connected components is an unrooted tree of valency $p+1$. For $B, B' \in \mathcal{B}$ we write $\text{dist}(B, B') = d$ if B and B' are at distance d in this graph; for each $B \in \mathcal{B}$, there are $(p+1)^d$ balls at distance d to B , and $\beta_d := \sum_{i=0}^d (p+1)^i = \frac{1}{p}((p+1)^{d+1} - 1)$ balls at distance at most d to B . Note that dist is a metric on each connected component of \mathcal{B} .

Lemma 7.3. *Let $A \subseteq K$ be finite, $A \neq \emptyset$. Then there at most $|A| - 1$ distinct balls of the form $B_{v(a-b)}(a)$ where $a, b \in A$, $a \neq b$.*

Proof. We may assume $|A| > 1$. For $a \in A$ set

$$\mu_a := 1 + \max \{v(a - b) : b \in A, a \neq b\}.$$

Let \mathcal{B}_A be the smallest connected subgraph of \mathcal{B} containing all $B_{\mu_a}(a)$, $a \in A$. Note that the balls $B_{\mu_a}(a)$, $a \in A$, are pairwise disjoint; in particular, they are the leaves of the tree \mathcal{B}_A . The balls $B_{v(a-b)}(a)$, where $a, b \in A$, $a \neq b$ are vertices of \mathcal{B}_A , and all but one of them has degree greater than 2. Now use the fact that any (undirected) tree with finite vertex set V with $|V| > 1$ has $2 + \sum_{v \in V, \deg(v) > 2} (\deg(v) - 2)$ leaves. (This follows immediately from the well-known formula $2(|V| - 1) = \sum_{v \in V} \deg(v)$.) \square

We also recall the following basic fact (a consequence of the Newton formulation of Hensel's Lemma, see [39, Lemma 2.3]) about the subgroups $P_n^{\times} = P_n \setminus \{0\}$ of the multiplicative group $K^{\times} = K \setminus \{0\}$ of K :

Lemma 7.4. *Suppose $n > 1$, and let $x, y, a \in K$ with $v(y - x) > 2v(n) + v(y - a)$. Then $(x - a)(y - a)^{-1} \in P_n^{\times}$.*

Suppose now that T is an \mathcal{L} -theory satisfying the hypothesis of Theorem 7.2, and $K \models T$. Employing definability of Skolem functions and the explicit description of immediate predecessors and successors in the partial order of \mathcal{B} given above, one easily shows, by induction on d :

Lemma 7.5. *Let $(B_b)_{b \in K^m}$ be a \emptyset -definable family of subsets of K . Then there exist \emptyset -definable functions $c_i, r_i: K^m \rightarrow K$, $i \in \mathbb{N}$, with the following property: if $b \in K^m$ is such that $B = B_b$ is a ball in K , then the balls $B_{v(r_i(b))}(c_i(b))$, where $i = 1, \dots, \beta_d$, are exactly the balls of distance at most d to B .*

The assumption of definable Skolem functions also guarantees that any model of the theory T has cell decomposition [70]. Let $\Delta(x; y)$ be a finite set of \mathcal{L} -formulas, where $|x| = 1$, closed under negation. Then there are integers $N, n > 0$, and for each $i = 1, \dots, N$ there are \emptyset -definable functions $f_i, g_i, c_i: K^{|y|} \rightarrow K$ and elements λ_i of a fixed set of representatives of the cosets of the subgroup P_n^\times of K^\times with the following properties: for any $\varphi \in \Delta$ and $b \in K^{|y|}$, the set $\varphi(K; b)$ of realizations of $\varphi(x; b)$ is a finite union of some of the cells $U_1(b), \dots, U_N(b)$ defined by the data given above, i.e., sets of the form

$$U_i(b) = \{x \in K : v(f_i(b)) \square_{i1} v(x - c_i(b)) \square_{i2} v(g_i(b)) \& P_n(\lambda_i(x - c_i(b)))\} \quad (7.5)$$

where each symbol \square_{ij} is \leq , $<$, or no condition. Note that this includes the cases where $U_i(b) = \{c_i(b)\}$ is a singleton, or where

$$U_i(b) = \{x \in K : P_n(\lambda_i(x - c_i(b)))\}. \quad (7.6)$$

The center of the cell $U_i(b)$ is given by $c_i(b)$. Since the value group of K has smallest positive element $v(p)$, using the equivalences

$$\begin{aligned} v(a) < v(a') &\iff v(pa) \leq v(a') \\ v(a) \leq v(a') &\iff v(a) < v(pa'), \end{aligned}$$

valid for all $a, a' \in K$, not both zero, one sees that we may assume that in (7.5), \square_{i1} is \leq or no condition, and \square_{i2} is $<$ or no condition. From now on, we assume for convenience that all our cells have this particular form.

Based on the above data for the cell decomposition, we now describe a uniform definition of $\Delta(x; B)$ -types over finite sets in K , in terms of the graph of balls \mathcal{B} . A *special ball* is a ball having one of the following forms:

$$B_{v(c_i(b) - c_j(b'))}(c_i(b)), \quad B_{v(f_i(b))}(c_i(b)), \quad B_{v(g_i(b))}(c_i(b)) \quad (b, b' \in K^{|y|}).$$

Note that each special ball can be defined by using at most two parameter tuples. We say that $B_{v(c_i(b) - c_j(b'))}(c_i(b))$, $B_{v(f_i(b))}(c_i(b))$ and $B_{v(g_i(b))}(c_i(b))$ are special balls defined over $\{b, b'\}$. We also say that a special ball is defined over a subset B of K if it is defined over $\{b, b'\}$ where $b, b' \in B$. By Lemma 7.3, given a finite $B \subseteq K$, there are no more than $3N \cdot |B| - 1$ special balls defined over B .

Let us say that a ball B' is *near* a ball B if $\text{dist}(B, B') \leq n + 4v(n) + 2$; each ball has $\beta := \beta_{n+4v(n)+2}$ balls near it. In particular, given $b_1, b_2 \in K^{|y|}$ there are at most $M := (6N - 1) \cdot \beta$ balls near special balls defined over $\{b_1, b_2\}$. Set $I^{(1)} := [M] = \{1, \dots, M\}$ and $I^{(2)} = I^{(3)} := [N]$. By Lemma 7.5 there are \mathcal{L} -formulas $\chi_i(x; y_1, y_2)$, $i \in I^{(1)}$, such that for each $b_1, b_2 \in K^{|y|}$, the formulas $\chi_i(x; b_1, b_2)$ define exactly the balls near special balls defined over $\{b_1, b_2\}$. For each $i \in I^{(1)}$ and $\varphi \in \Delta$ define

$$\varphi_i^{(1)}(y; y_1, y_2) := \forall x (\chi_i(x; y_1, y_2) \rightarrow \varphi(x; y)).$$

So for $b, b_1, b_2 \in K^{|y|}$ we have:

$$K \models \varphi_i^{(1)}(b; b_1, b_2) \iff \varphi(K; b) \supseteq \chi_i(K; b_1, b_2).$$

For each $i \in I^{(2)} = I^{(3)}$ and $\varphi \in \Delta$ set

$$\varphi_i^{(2)}(y; y_1) := \varphi(c_i(y_1); y)$$

and

$$\varphi_i^{(3)}(y; y_1) := \forall x (P_N(\lambda_i(x - c_i(y_1))) \rightarrow \varphi(x; y)).$$

Now set $\mathcal{F}_i^{(j)} := (\varphi_i^{(j)}(y; y_1, y_2))_{\varphi \in \Delta}$ for each $i \in I^{(j)}$, $j = 1, 2, 3$. The first part of Theorem 7.2 will be proved once we show the following:

Claim. $\mathcal{F} = (\mathcal{F}_i^{(j)})$ is a coherent definition of $\Delta(x; B)$ -types over finite sets.

Since the coherency condition is obviously satisfied, it is enough to show that \mathcal{F} is a uniform definition of $\Delta(x; B)$ -types over finite sets. For this let $B \subseteq K^{|y|}$ be finite and non-empty, and let $q \in S^\Delta(B)$. Let

$$c(B) := \{c_i(b) : b \in B, i = 1, \dots, N\}$$

be the set of centers of the cells $U_i(b)$. In the following we let i range over $[N] = \{1, \dots, N\}$ and b (possibly with decorations) over B . By a “special ball” we always mean a special ball defined over B , and similarly a “near ball” is a ball near a special ball (defined over B).

We first eliminate two special cases (which are taken care of by the families $(\mathcal{F}_i^{(2)})$ and $(\mathcal{F}_i^{(3)})$): Suppose first that $q^K \cap c(B) \neq \emptyset$, say $c_i(b_1) \in q^K$ for some i and $b_1 \in B$. In this case, with such choice of i and b_1 , $\mathcal{F}_i^{(2)}(y; b_1)$ defines q . Similarly, if $|c(B)| = 1$, and for all i and b with $q^K \subseteq U_i(b)$, the condition \square_{i1} is vacuous and $g_i(b) = 0$, then all such cells $U_i(b)$ have the form as in (7.6), and for suitable i and b_1 , $\mathcal{F}_i^{(3)}(y; b_1)$ defines q .

So from now on we may assume that:

- (a) q^K is disjoint from $c(B)$; and
- (b) if $c(B)$ is a singleton, then for some i and b with $q^K \subseteq U_i(b)$ the condition \square_{i1} is \leq or $g_i(b) \neq 0$.

Under these assumptions, it is enough to show: *there is a near ball D such that $D \subseteq q^K$.* We first note:

Lemma 7.6. *Let $a \in K \setminus c(B)$, and let B_1 be a ball containing a which is maximal subject to the condition $B_1 \cap c(B) = \emptyset$; that is, $B_1 = B_\delta(a)$ where*

$$\delta = 1 + \max \{v(a - c) : c \in c(B)\}.$$

Let also $B_0 := B_{\delta+2v(n)}(a)$. Then for all $x \in K$ and $c \in c(B)$ we have:

- (1) $x \in B_1 \Rightarrow v(x - c) = v(a - c)$;
- (2) $x \in B_0 \Rightarrow (x - c)(a - c)^{-1} \in P_n$.

In particular, if a cell $U_i(b)$ contains a then it contains B_0 .

The proof of (1) is obvious, and to deduce (2) from (1) use Lemma 7.4.

Let $a \in K$ realize q , and define δ , B_0 and B_1 as in the previous lemma. Also, take $c \in c(B)$ such that $\delta = 1 + v(a - c)$.

Lemma 7.7. *Let B_2 be a ball. Then*

- (1) B_2 properly contains B_1 iff it contains both a and c ;
- (2) if B_2 contains c but not a , then $\text{dist}(B_1, B_2) = \text{rad}(B_2) - \delta + 2$.

The proof of (1) is clear, and for (2) note that if $c \in B_2$ and $a \notin B_2$, then $B_1, B_2 \subseteq B_{\delta-1}(a)$.

We first assume that there is a special ball E such that $\text{dist}(B_1, E) \leq 2v(n) + n + 1$. Then by Lemma 7.6, $D := B_0 = B_{\delta+2v(n)}(a)$ is contained in those cells $U_i(b)$ which contain a ; hence $D \subseteq q^K$. Also, $\text{dist}(D, E) \leq n + 4v(n) + 2$, so D is near E . Hence the ball D has the required properties.

So from now on, we may suppose that for any special ball E we have $\text{dist}(B_1, E) > 2v(n) + n + 1$. We distinguish two cases:

Case 1: there is a special ball which contains B_1 . We let C be the smallest such special ball, with radius ρ . We have $c \in C$ (since C properly contains B_1) and hence $C = B_\rho(c)$. The idea now is to replace a and the ball $D = B_0$ by another realization $a' \in C$ of q and a ball D' which is contained in and near the special ball C . As Γ is a \mathbb{Z} -group, there is a unique $\delta' \in \Gamma$ such that

$$\rho + 2v(n) + 1 < \delta' \leq \rho + 2v(n) + 1 + n \quad \text{and} \quad \delta' \equiv \delta \pmod{n}.$$

By assumption we have

$$2v(n) + n + 1 < \text{dist}(B_1, C) = \delta - \rho$$

and hence $\delta > \delta' > \rho + 1$. Now choose $d \in P_n$ with $v(d) = \delta' - \delta$. (From now on until the end of the proof of Theorem 7.2 we temporarily suspend our promise of d always denoting a natural number.) Put

$$D' := B_{\delta'+2v(n)+1}(a') \quad \text{where } a' := d(a - c) + c.$$

Then D' is contained in the special ball C , and D' is near C . Indeed, $D' \subseteq B_{\delta'-1}(a')$, and these two balls are at distance $2v(n) + 2$. The latter ball also contains c , so

$$\text{dist}(B_{\delta'-1}(a'), C) = \text{dist}(B_{\delta'-1}(c), B_\rho(c)) = \delta' - 1 - \rho \leq 2v(n) + n,$$

hence $\text{dist}(D', C) \leq 4v(n) + 2 + n$. Thus D' has the right properties, provided we manage to show:

Claim 1. Let $U_i(b)$ be a cell as in (7.5) which contains a . Then $D' \subseteq U_i(b)$.

Towards the proof of this claim, we first show two auxiliary claims:

Claim 2. Let $c' \in c(B)$. Then

$$v(a' - c') = \begin{cases} \delta' - 1 & \text{if } v(c - c') \geq v(a - c), \\ v(c - c') \leq \rho < \delta' - 1 & \text{otherwise.} \end{cases}$$

Proof. If $v(c - c') \geq v(a - c)$ then

$$v(c - c') \geq v(a - c) = \delta - 1 > \delta' - 1$$

and hence

$$v(a' - c') = v(d(a - c) + (c - c')) = \delta' - 1.$$

So suppose $v(c - c') < v(a - c)$. We have $v(a - c) > v(c - c')$, so $v(a - c') = v(c - c')$ and hence a is contained in the special ball $E := B_{v(c-c')}(c)$. In fact, for every $x \in B_1$ we have

$$v(x - a) \geq \delta = v(a - c) + 1 > v(a - c) > v(c - c')$$

and hence $B_1 \subseteq E$. By minimality of C thus $C \subseteq E$. This yields $\delta' - 1 > \rho \geq v(c - c')$ and thus $v(a' - c') = v(c - c') < \delta' - 1$. \square

Claim 3. For every $c' \in c(B)$ we have $v(a - c') \geq v(a' - c')$.

Proof. Certainly, $v(a - c') \geq \min\{v(a - a'), v(a' - c')\}$. But the minimum is always achieved by $v(a' - c')$, as

$$v(a - a') = v((d - 1)(a - c)) = \delta' - 1 \geq v(a' - c')$$

by Claim 2. □

By Claim 2 and Lemma 7.6 (applied to a' in place of a), in order to show Claim 1, it is enough to prove that $a' \in U_i(b)$. We abbreviate $c' = c_i(b)$. Suppose that \square_{i1} is \leq . Then $f_i(b) \neq 0$, and by Lemma 7.6, (1), all elements x of B_1 satisfy the condition $v(f_i(b)) \leq v(x - c')$; hence C is contained in the special ball $B_{v(f_i(b))}(c')$, by the minimality of C . Since $D' \subseteq C$, all elements x of D' also satisfy $v(f_i(b)) \leq v(x - c')$; in particular, of course, $v(f_i(b)) \leq v(a' - c')$. If \square_{i2} is $<$, then by Claim 3, $v(a' - c') \leq v(a - c') < v(g_i(b))$, as required. It remains to check that $a - c'$ and $a' - c'$ lie in the same coset of P_n^\times . We distinguish two cases. If $v(c - c') \leq \rho$ then $v(a - c) = \delta - 1 > \rho \geq v(c - c')$, hence $v(a - c') = v(c - c')$ and thus

$$v(a - a') = \delta' - 1 > 2v(n) + \rho \geq 2v(n) + v(a - c');$$

therefore $a - c'$ and $a' - c'$ are in the same P_n^\times -coset, by Lemma 7.4. Suppose $v(c - c') > \rho$. Then by Claim 2 we have $v(c - c') \geq v(a - c) = \delta - 1$ and $v(a' - c') = \delta' - 1$. Now consider the special ball $E := B_{v(c - c')}(c)$. Note that $a \notin E$: otherwise $v(c - c') = v(a - c) = \delta - 1$ and hence $B_1 = B_\delta(a) \subseteq E$ with $\text{dist}(B_1, E) = 1$, contrary to our initial assumption (made before Case 1). Thus, by Lemma 7.7, (2) and said assumption, we obtain $2v(n) + n + 1 < v(c - c') - \delta + 2$. Hence

$$v(c - c') > 2v(n) + \delta - 1 = 2v(n) + v(a - c) \geq 2v(n) + v(a' - c'),$$

with the last inequality by Claim 3. So by Lemma 7.4, $a - c'$ and $a - c$ are in the same P_n^\times -coset, as are $a' - c$ and $a' - c'$. Certainly, as $a' - c = d(a - c)$ and $d \in P_n^\times$, the elements $a - c$ and $a' - c$ lie in the same coset of P_n^\times . Hence $a - c'$ and $a' - c'$ also lie in the same coset of P_n^\times . This finishes the proof of Claim 1, and hence of Case 1. □

Case 2: no special ball contains B_1 . In this case, for every $c' \in c(B)$, the special ball $C = B_{v(c - c')}(c)$ does not contain a , so $v(c' - c) > v(a - c) = \delta - 1$ and hence

$$v(a - c') = \min\{v(a - c), v(c - c')\} = v(a - c) = \delta - 1.$$

Since C is of distance greater than $2v(n) + n + 1$ from B_1 , by part (2) of Lemma 7.7 we also obtain

$$v(c - c') > \delta + 2v(n) + n - 1. \tag{7.7}$$

Similarly, since each special ball $B_{v(g_i(b))}(c_i(b))$ does not contain a , we get

$$v(g_i(b)) > \delta + 2v(n) + n - 1,$$

and since $a \notin B_{v(f_i(b))}(c_i(b))$, the condition \square_{i1} is vacuous for each i and b with $a \in U_i(b)$.

Fix a special ball E of the form $B_{v(g_i(b))}(c_i(b))$ with minimal radius $\gamma = v(g_i(b))$, if there is such a special ball; otherwise let $\gamma = \infty$. Also, if $|c(B)| > 1$, let C be a special ball of the form $B_{v(c - c')}(c)$, where $c' \in c(B)$, with minimal radius $\rho = v(c - c')$; we set $\rho = \infty$ if $|c(B)| = 1$. Note that by our general assumption (made before Lemma 7.6), not both of γ and ρ are ∞ . We now distinguish two subcases:

Case 2a: $\rho - 2v(n) \leq \gamma$. Let $\delta' \in \Gamma$ such that

$$\rho - 2v(n) - n < \delta' \leq \rho - 2v(n), \quad \delta' \equiv \delta \pmod{n},$$

choose $d \in P_n$ with $v(d) = \delta' - \delta$, and set

$$D' := B_{\delta'+2v(n)}(a') \quad \text{where } a' := d(a - c) + c.$$

By (7.7) we have $\delta' > \delta$. Moreover, for each $c'' \in c(B)$ we have $v(c - c'') \geq \rho > \delta' - 1 = v(d(a - c))$ and hence $v(a' - c'') = \delta' - 1$. Note that the ball $B_{\delta'-1}(a')$ contains D' and is of distance $2v(n) + 1$ to D' . The ball $B_{\delta'-1}(a')$ contains c , hence

$$\text{dist}(B_{\delta'-1}(a'), C) \leq \rho - (\delta' - 1) < 2v(n) + n + 1$$

and thus $\text{dist}(D', C) < 4v(n) + n + 2$, so D' is near C . Let $U_i(b)$ be a cell containing a ; it remains to show that then $a' \in U_i(b)$. We already noted that condition \square_{i1} is vacuous. As to \square_{i2} , suppose that condition is $<$. Writing $c' = c_i(b)$ we then have

$$v(a' - c') = \delta' - 1 < \rho - 2v(n) \leq \gamma \leq v(g_i(b))$$

as required. Finally, by (7.7) and Lemma 7.4, $a - c'$ and $a - c$ are in the same P_n^\times -coset, and since

$$v(c - c') \geq \rho > 2v(n) + \delta' - 1 = 2v(n) + v(a' - c'),$$

the elements $a' - c$ and $a' - c'$ are also in the same P_n^\times -coset. As $a' - c = d(a - c)$ and $a - c$ are in the same P_n^\times -coset, finally $a' - c'$ and $a - c'$ are in the same P_n^\times -coset, as required. \square

Case 2b: $\rho - 2v(n) > \gamma$. In this case we let $\delta' \in \Gamma$ be such that

$$\gamma - 2v(n) - n < \delta' \leq \gamma - 2v(n), \quad \delta' \equiv \delta \pmod{n},$$

and with this choice of δ' define d , a' and D' as in Case 2a. Note that $\rho > \gamma$, so for each $c'' \in c(B)$ we have $v(c - c'') > \delta' - 1 = v(d(a - c))$ and hence $v(a' - c'') = \delta' - 1$. Since

$$\text{dist}(B_{\delta'-1}(a'), E) \leq \gamma - (\delta' - 1) < 2v(n) + n + 1$$

we see, similarly as in Case 2a, that $\text{dist}(D', E) < 4v(n) + n + 2$, so D' is near E . Let $U_i(b)$ be a cell containing a , and suppose \square_{i2} is $<$. Then $v(a' - c') = \delta' - 1 < \gamma \leq v(g_i(b))$, and as at the end of Case 2a one sees that $a' - c'$ and $a - c'$ are in the same coset of P_n^\times . \square

To complete the proof of the theorem, we apply Theorem 7.1 with $r = 1$. By what we have shown above, for every finite non-empty $B \subseteq M^{|y|}$, each type in $S^\Delta(B)$ is uniquely determined by either a center $c_i(b)$, where $b \in B$, or a near ball. However, there are at most $N|B| = O(|B|)$ centers, and at most $(3N|B| - 1) \cdot \beta = O(|B|)$ near balls; thus $|S^\Delta(B)| = O(|B|)$ as required. \square

From Theorem 7.2 and Corollary 5.13 we obtain:

Corollary 7.8. *Every P -minimal theory with definable Skolem functions is dp -minimal.*

Remark 7.9. In [24, Section 6], Dolich, Goodrick and Lippel already showed that $p\text{CF} = \text{Th}(\mathbb{Q}_p)$ is dp -minimal. By 3.6 of [22], the P -minimal theory $p\text{CF}_{\text{an}}$ has definable Skolem functions. (Formally, [22] handles the corresponding subanalytic structure on \mathbb{Z}_p , but the translation is straightforward.) Note that the proof of [22, 3.6] takes place in the ground model \mathbb{Z}_p , where all elements are named by constants, so ‘definable’ means ‘ \emptyset -definable’, and curve selection, as stated there, gives definable Skolem functions. Hence the conclusion of Theorem 7.2 and Corollary 7.8 apply to it and its reducts. (See also Lemma 3.6.) Cell decomposition in $p\text{CF}_{\text{an}}$ is also proved in [21].

REFERENCES

- [1] K. A. S. Abdel-Ghaffar, *Maximum number of edges joining vertices on a cube*, Inform. Process. Lett. **87** (2003), no. 2, 95–99.
- [2] H. Adler, *Theories controlled by formulas of Vapnik-Chervonenkis codimension 1*, preprint (2008).
- [3] ———, *An introduction to theories without the independence property*, preprint (2008).
- [4] M. Anthony, G. Brightwell, and C. Cooper, *The Vapnik-Chervonenkis dimension of a random graph*, Discrete Math. **138** (1995), no. 1-3, 43–56.
- [5] M. Aschenbrenner, L. van den Dries, *Closed asymptotic couples*, J. Algebra **225** (2000), 309–358.
- [6] M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson, S. Starchenko, *Vapnik-Chervonenkis density in some theories without the independence property*, II, preprint (2011).
- [7] P. Assouad, *Densité et dimension*, Ann. Inst. Fourier (Grenoble) **33** (1983), no. 3, 233–282.
- [8] ———, *Observations sur les classes de Vapnik-Chervonenkis et la dimension combinatoire de Blei*, in: *Seminaire d'Analyse Harmonique, 1983–1984*, pp. 92–112, Publications Mathématiques d'Orsay, vol. 85-2, Université de Paris-Sud, Département de Mathématiques, Orsay, 1985.
- [9] J. Baldwin, S. Shelah, *Randomness and semigenericity*, Trans. Amer. Math. Soc. **349** (1997), no. 4, 1359–1376.
- [10] J. Balogh, B. Bollobás, *Unavoidable traces of set systems*, Combinatorica **25** (2005), no. 6, 633–643.
- [11] S. Basu, R. Pollack, M.-F. Roy, *On the number of cells defined by a family of polynomials on a variety*, Mathematika **43** (1996), no. 1, 120–126.
- [12] L. Bélair, *Types dans les corps valués munis d'applications coefficients*, Illinois J. Math. **43** (1999), no. 2, 410–425.
- [13] L. Bélair, M. Bousquet, *Types dans les corps valués*, C. R. Acad. Sci. Paris Sér. I Math. **323** (1996), no. 8, 841–844.
- [14] O. Belegradek, Y. Peterzil, F. Wagner, *Quasi-o-minimal structures*, J. Symbolic Logic **65** (2000), no. 3, 1115–1132.
- [15] O. Belegradek, V. Verbovskiy, F. Wagner, *Coset-minimal groups*, Ann. Pure Appl. Logic **121** (2003), no. 2-3, 113–143.
- [16] G. Birkhoff, *Lattice Theory*, 3rd ed., American Mathematical Society Colloquium Publications, vol. XXV, American Mathematical Society, Providence, R.I., 1967.
- [17] H. Brönnimann, M. T. Goodrich, *Almost optimal set covers in finite VC-dimension*, Discrete Comput. Geom. **14** (1995), no. 4, 463–479.
- [18] E. Casanovas, M. Ziegler, *Stable theories with a new predicate*, J. Symbolic Logic **66** (2001), no. 3, 1127–1140.
- [19] G. Cherlin, F. Point, *On extensions of Presburger arithmetic*, in: B. I. Dahn (ed.), *Proceedings of the fourth Easter conference on model theory (Gross Körös, 1986)*, pp. 17–34, Seminarberichte, vol. 86, Humboldt Universität, Sektion Mathematik, Berlin, 1986.
- [20] A. Chernikov, P. Simon, *Externally definable sets and dependent pairs*, preprint (2010), available online at <http://front.math.ucdavis.edu/1007.4468>.
- [21] R. Cluckers, *Analytic p -adic cell decompositions and integrals*, Trans. Amer. Math. Soc. **356** (2004), 1489–1499.
- [22] J. Denef, L. van den Dries, *p -adic and real subanalytic sets*, Ann. Math. **128** (1988), 70–138.
- [23] M. A. Dickmann, *Elimination of quantifiers for ordered valuation rings*, J. Symbolic Logic **52** (1987), 116–128.
- [24] A. Dolich, J. Goodrick, D. Lippel, *D_p -minimal theories: basic facts and examples*, Notre Dame J. Formal Logic **52** (2011), no. 3, 267–288.
- [25] L. van den Dries, *Tame Topology and O -minimal Structures*, London Mathematical Society Lecture Note Series, vol. 248, Cambridge University Press, Cambridge, 1998.
- [26] L. van den Dries, D. Haskell and H. D. Macpherson, *One-dimensional p -adic subanalytic sets*, J. London Math. Soc. (2) **59** (1999), 1–20.
- [27] R. M. Dudley, *A course on empirical processes*, in: P. L. Hennequin (ed.), *École d'été de probabilités de Saint-Flour XII*, pp. 1–142, Lecture Notes in Mathematics, vol. 1097, Springer-Verlag, Berlin, 1984.
- [28] ———, *Uniform Central Limit Theorems*, Cambridge Studies in Advanced Mathematics, vol. 63, Cambridge University Press, Cambridge, 1999.
- [29] J.-L. Duret, *Les corps faiblement algébriquement clos non séparablement clos ont la propriété d'indépendance*, in: L. Pacholski et al. (eds.), *Model Theory of Algebra and Arithmetic* (Proc.

- Conf., Karpacz, 1979), pp. 136–162, Lecture Notes in Mathematics, vol. 834, Springer-Verlag, Berlin, 1980.
- [30] H. Edelsbrunner, *Algorithms in Combinatorial Geometry*, EATCS Monographs on Theoretical Computer Science, vol. 10, Springer-Verlag, Berlin, 1987.
- [31] Gy. Elekes, *SUMS versus PRODUCTS in number theory, algebra and Erdős geometry*, in: G. Halász et al. (eds.), *Paul Erdős and his Mathematics*, II (Budapest, 1999), pp. 241–290, Bolyai Soc. Math. Stud., vol. 11, János Bolyai Math. Soc., Budapest, 2002.
- [32] P. Erdős, R. Rado, *A partition calculus in set theory*, Bull. Amer. Math. Soc. **62** (1956), 427–489.
- [33] Z. Füredi, J. Pach, *Traces of finite sets: extremal problems and geometric applications*, in: P. Frankl et al. (eds.), *Extremal Problems for Finite Sets* (Visegrád, 1991), pp. 251–282, Bolyai Soc. Math. Stud., vol. 3, János Bolyai Math. Soc., Budapest, 1994.
- [34] V. Guingona, *On uniform definability of types over finite sets*, preprint (2010), available online at <http://front.math.ucdavis.edu/1005.4924>.
- [35] V. Guingona, C. Hill, *Local dp-rank and VC-density over indiscernible sequences*, preprint (2011), available online at <http://front.math.ucdavis.edu/1108.2554>.
- [36] Y. S. Gurevich, P. H. Schmitt, *The theory of ordered abelian groups does not have the independence property*, Trans. Amer. Math. Soc. **284** (1984), 171–182.
- [37] J. Hartman, *The homeomorphic embedding of K_n in the m -cube*, Discrete Math. **16** (1976), no. 2, 157–160.
- [38] D. Haskell, H. D. Macpherson, *Cell decompositions of C-minimal structures*, Ann. Pure Appl. Logic **66** (1994), no. 2, 113–162.
- [39] ———, *A version of o-minimality for the p-adics*, J. Symb. Logic **62** (1997), 1075–1092.
- [40] ———, *VC density in real closed valued fields*, Prépublications de la séminaire de structures algébriques ordonnées **83** (2008–2009), Equipe de logique, Université Paris VII.
- [41] D. Haussler, *Sphere packing numbers for subsets of the Boolean n -cube with bounded Vapnik-Chervonenkis dimension*, J. Combin. Theory Ser. A **69** (1995), no. 2, 217–232.
- [42] W. Hodges, *Model Theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993.
- [43] J. E. Holly, *Canonical forms for definable subsets of algebraically closed and real closed valued fields*, J. Symbolic Logic **60** (1995), no. 3, 843–860.
- [44] E. Hrushovski, D. Kazhdan, *Integration in valued fields*, in: V. Ginzburg (ed.), *Algebraic Geometry and Number Theory*, pp. 261–405, Progress in Mathematics, vol. 253, Birkhäuser Boston, Inc., Boston, MA, 2006.
- [45] U. Hrushovski, A. Pillay, *Weakly normal groups*, in: Ch. Berline et al. (eds.), *Logic Colloquium '85 (Orsay, 1985)*, pp. 233–244, Stud. Logic Found. Math., vol. 122, North-Holland, Amsterdam, 1987.
- [46] G. Jeronimo, J. Sabia, *On the number of sets definable by polynomials*, J. Algebra **227** (2000), no. 2, 633–644.
- [47] H. R. Johnson, M. C. Laskowski, *Compression schemes, stable definable families, and o-minimal structures*, Discrete Comput. Geom. **43** (2010), no. 4, 914–926.
- [48] I. Kaplan, A. Onshuus, A. Usvyatsov, *Additivity of the dp-rank*, preprint (2011), available online as no. 251 at <http://www.logique.jussieu.fr/modnet/Publications/Preprint%20server/>.
- [49] M. Karpinski, A. Macintyre, *Polynomial bounds for VC dimension of sigmoidal and general pfaffian neural networks*, J. Comput. System Sci. **54** (1997), 169–176.
- [50] ———, *Approximating volumes and integrals in o-minimal and P-minimal theories*, in: A. Macintyre (ed.), *Connections between Model Theory and Algebraic and Analytic Geometry*, pp. 149–177, Quad. Mat., vol. 6, Dept. Math., Seconda Univ. Napoli, Caserta, 2000.
- [51] T. Kövari, V. T. Sós, and P. Turán, *On a problem of K. Zarankiewicz*, Colloquium Math. **3** (1954), 50–57.
- [52] K. Kudaïbergenov, *On the independence property*, Siberian Math. J. **41** (2000), no. 1, 113.
- [53] ———, *Weakly quasi-o-minimal models*, Siberian Adv. Math. **20** (2010), no. 4, 285–292.
- [54] F.-V. Kuhlmann, *Abelian groups with contractions, II: Weak o-minimality*, in: A. Facchini, C. Menini (eds.): *Abelian Groups and Modules*, Kluwer, Dordrecht (1995), 323–342.
- [55] M. C. Laskowski, *Vapnik-Chervonenkis classes of definable sets*, J. London Math. Soc. (2) **45** (1992), no. 2, 377–384.
- [56] ———, unpublished notes.
- [57] L. Lipshitz, *Rigid subanalytic sets*, Amer. J. Math. **115** (1993), no. 1, 77–108.

- [58] L. Lipshitz, Z. Robinson, *One-dimensional fibers of rigid subanalytic sets*, J. Symbolic Logic **63** (1998), 83–88.
- [59] L. Lovász and B. Szegedy, *Regularity partitions and the topology of graphons*, in: I. Bárány et al. (eds.), *An Irregular Mind. Szemerédi is 70*, pp. 415–446, Bolyai Society Mathematical Studies, vol. 21, Springer-Verlag, Berlin; János Bolyai Mathematical Society, Budapest, 2010.
- [60] A. J. Macintyre, *On definable subsets of p -adic fields*, J. Symbolic Logic **41** (1976), 605–610.
- [61] D. Macpherson, D. Marker and C. Steinhorn, *Weakly o -minimal structures and real closed fields*, Trans. Amer. Math. Soc. **352** (2000), no. 12, 5435–5483.
- [62] D. Macpherson, C. Steinhorn, *On variants of o -minimality*, Ann. Pure Appl. Logic **79** (1996), no. 2, 165–209.
- [63] D. Marker, *Model Theory*, Graduate Texts in Mathematics, vol. 217, Springer-Verlag, New York, 2002.
- [64] J. Matoušek, *Tight upper bounds for the discrepancy of half-spaces*, Discrete Comput. Geom. **13** (1995), no. 3-4, 593–601.
- [65] ———, *Geometric set systems*, in: A. Balog et al. (eds.), *European Congress of Mathematics, II (Budapest, 1996)*, 1–27, Progr. Math., vol. 169, Birkhäuser, Basel, 1998.
- [66] ———, *Lectures on Discrete Geometry*, Graduate Texts in Mathematics, vol. 212, Springer-Verlag, New York, 2002.
- [67] ———, *Bounded VC-dimension implies a fractional Helly theorem*, Discrete Comput. Geom. **31** (2004), no. 2, 251–255.
- [68] J. Matoušek, E. Welzl, L. Wernisch, *Discrepancy and approximations for bounded VC-dimension*, Combinatorica **13** (1993), no. 4, 455–66.
- [69] C. Michaux, R. Villemaire, *Presburger arithmetic and recognizability of sets of natural numbers by automata: new proofs of Cobham’s and Semenov’s theorems*, Ann. Pure Appl. Logic **77** (1996), no. 3, 251–277.
- [70] M.-H. Mourgues, *Cell decomposition for P -minimal fields*, MLQ Math. Log. Q. **55** (2009), no. 5, 487–492.
- [71] A. Onshuus, A. Usvyatsov, *On dp -minimality, strong dependence and weight*, J. Symbolic Logic **76** (2011), no. 3, 737–758.
- [72] J. Pach, M. Sharir, *Repeated angles in the plane and related problems*, J. Combin. Theory Ser. A **59** (1992), no. 1, 12–22.
- [73] ———, *On the number of incidences between points and curves*, Combin. Probab. Comput. **7** (1998), no. 1, 121–127.
- [74] ———, *Geometric incidences*, in: J. Pach (ed.), *Towards a Theory of Geometric Graphs*, pp. 185–223, Contemporary Mathematics, vol. 342, American Mathematical Society, Providence, RI, 2004.
- [75] M. Parigot, *Théories d’arbres*, J. Symbolic Logic **47** (1982), no. 4, 841–853.
- [76] A. Pillay, *The model-theoretic content of Lang’s conjecture*, in: E. Bouscaren (ed.), *Model Theory and Algebraic Geometry*, pp. 101–106, Lecture Notes in Mathematics, vol. 1696, Springer-Verlag, Berlin, 1998.
- [77] K.-P. Podewski, M. Ziegler, *Stable graphs*, Fund. Math. **100** (1978), no. 2, 101–107.
- [78] F. Point, *On decidable extensions of Presburger arithmetic: from A. Bertrand numeration systems to Pisot numbers*, J. Symbolic Logic **65** (2000), no. 3, 1347–1374.
- [79] F. Point, F. O. Wagner, *Essentially periodic ordered groups*, Ann. Pure Appl. Logic **105** (2000), no. 1–3, 261–291.
- [80] B. Poizat, *Cours de Théorie des Modèles*, Nur al-Mantiq wal-Marifah, Villeurbanne (1985).
- [81] R. Pollack, M.-R. Roy, *On the number of cells defined by a set of polynomials*, C. R. Acad. Sci. Paris Sér. I Math. **316** (1993), no. 6, 573–577.
- [82] M. Prest, *Model Theory and Modules*, London Mathematical Society Lecture Note Series, vol. 130, Cambridge University Press, Cambridge, 1988.
- [83] N. Sauer, *On the density of families of sets*, J. Combinatorial Theory Ser. A **13** (1972), 145–147.
- [84] J. H. Schmerl, *\aleph_0 -categorical distributive lattices of finite breadth*, Proc. Amer. Math. Soc. **87** (1983), no. 4, 707–713.
- [85] ———, *Partially ordered sets and the independence property*, J. Symbolic Logic **54** (1989), no. 2, 396–401.
- [86] S. Shelah, *Stability, the f.c.p., and superstability; model theoretic properties of formulas in first-order theory*, Ann. Math. Logic **3** (1971), no. 3, 271–362.

- [87] ———, *A combinatorial problem; stability and order for models and theories in infinitary languages*, Pacific J. Math. **41** (1972), 247–261.
- [88] ———, *Classification Theory and the Number of Nonisomorphic Models*, 2nd ed., Studies in Logic and the Foundations of Mathematics, vol. 92, North-Holland Publishing Co., Amsterdam, 1990.
- [89] ———, *Dependent first order theories, continued*, Israel J. Math. **173** (2009), 1–60.
- [90] ———, *Strongly dependent theories*, Israel J. Math., to appear, available online at <http://arxiv.org/abs/math/0504197>.
- [91] S. Shelah, J. Spencer, *Zero-one laws for sparse random graphs*, J. Amer. Math. Soc. **1** (1988), no. 1, 97–115.
- [92] P. Simon, *On dp-minimal ordered structures*, J. Symbolic Logic **76** (2011), no. 2, 448–460.
- [93] J. Solymosi, T. Tao, *An incidence theorem in higher dimensions*, preprint (2011), available online at <http://front.math.ucdavis.edu/1103.2926>.
- [94] J. Spencer, *The Strange Logic of Random Graphs*, Algorithms and Combinatorics, vol. 22, Springer-Verlag, Berlin, 2001.
- [95] J. Spencer, E. Szemerédi, W. T. Trotter, Jr., *Unit distances in the Euclidean plane*, in: B. Bollobás (ed.), *Graph Theory and Combinatorics*, pp. 293–303, Academic Press, Inc., London, 1984.
- [96] E. Szemerédi, W. T. Trotter, Jr., *A combinatorial distinction between the Euclidean and projective planes*, European J. Combin. **4** (1983), no. 4, 385–394.
- [97] C. Tóth, *The Szemerédi-Trotter theorem in the complex plane*, preprint (2003), available online at <http://front.math.ucdavis.edu/0305.5283>.
- [98] V. N. Vapnik, A. Ja. Červonenkis, *The uniform convergence of frequencies of the appearance of events to their probabilities*, Theor. Probability Appl. **16** (1971), 264–280.
- [99] V. Weispfenning, *Elimination of quantifiers for certain ordered and lattice-ordered abelian groups*, Bull. Soc. Math. Belg. Sér. B **33** (1981), 131–155.
- [100] J. Wierzejewski, *On stability and products*, Fund. Math. **93** (1976), no. 2, 81–95.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, LOS ANGELES, BOX 951555, LOS ANGELES, CA 90095-1555, U.S.A.

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

DEPARTMENT OF MATHEMATICS, EAST STROUDSBURG UNIVERSITY, SCIENCE & TECHNOLOGY CENTER, ROOM 118, EAST STROUDSBURG, PA 18301, U.S.A.

DEPARTMENT OF MATHEMATICS AND STATISTICS, MCMASTER UNIVERSITY, 1280 MAIN ST W, HAMILTON ON L8S 4K1, CANADA

E-mail address: haskell@math.mcmaster.ca

SCHOOL OF MATHEMATICS, UNIVERSITY OF LEEDS, LEEDS LS2 9JT, U.K.

E-mail address: h.d.macpherson@leeds.ac.uk

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF NOTRE DAME, 255 HURLEY BUILDING, NOTRE DAME, IN 46556-4618, U.S.A.

E-mail address: starchenko.1@nd.edu