VC-DIMENSION OF SHORT PRESBURGER FORMULAS

DANNY NGUYEN* AND IGOR PAK*

Abstract. We study VC-dimension of *short formulas* in Presburger Arithmetic, defined to have a bounded number of variables, quantifiers and atoms. We give both lower and upper bounds, which are tight up to a polynomial factor in the bit length of the formula.

1. Introduction

The notion of VC-dimension was introduced by Vapnik and Červonenkis in [VC71]. Although originally motivated by applications in probability and statistics, it was quickly adapted to computer science, learning theory, combinatorics, logic and other areas. We refer to [Vap98] for the extensive review of the subject, and to [Che16] for an accessible introduction to combinatorial and logical aspects.

1.1. **Definitions of VC-dimension and VC-density.** Let X be a set and $\mathcal{S} \subseteq 2^X$ be a family of subsets of X. For a subset $A \subseteq X$, let $\mathcal{S} \cap A \coloneqq \{S \cap A : S \in \mathcal{S}\}$ be the family of subsets of A cut out by $\mathcal{S}$. A subset $A \subseteq X$ is *shattered* by $\mathcal{S}$ if $\mathcal{S} \cap A = 2^A$, i.e., for every subset $B \subseteq A$, there is $S \in \mathcal{S}$ with $B = S \cap A$. The largest size $|A|$ among all subsets $A \subseteq X$ shattered by $\mathcal{S}$ is called the *VC-dimension* of $\mathcal{S}$, denoted by $\mathrm{VC}(\mathcal{S})$. If no such largest size $|A|$ exists, we write $\mathrm{VC}(\mathcal{S}) = \infty$.

The *shatter function* $\pi_{\mathcal{S}}$ is defined as follows:

$$\pi_{\mathcal{S}}(n) \,=\, \max\big\{|\mathcal{S} \cap A| : A \subseteq X, |A| = n\big\},$$

The *VC-density* of $\mathcal{S}$, denoted by $\mathrm{vc}(\mathcal{S})$ is defined as

$$\inf\Big\{r \in \mathbb{R}^+ \,:\, \operatorname{limsup}_{n\to\infty} \frac{\pi_{\mathcal{S}}(n)}{n^r} < \infty\Big\}.$$

The classical theorem of Sauer and Shelah [Sa72, Sh72] states that

$$\mathrm{vc}(\mathcal{S}) \,\le\, \mathrm{VC}(\mathcal{S}).$$

In other words, $\pi_{\mathcal{S}}(n) = O(n^d)$ in case $\mathcal{S}$ has finite VC-dimension d. In general, VC-density can be much smaller than VC-dimension, and also behaves a lot better under various operations on $\mathcal{S}$.

1.2. **NIP theories and bounds on VC-dimension/density.** It is of interest to distinguish the first-order theories in which VC-dimension and VC-density behave nicely. Let $\mathcal{L}$ be a first-order language and $\boldsymbol{M}$ be an $\mathcal{L}$-structure. Consider a *partitioned $\mathcal{L}$-formula* $F(\mathbf{x}; \mathbf{y})$ whose free variables are separated into two groups $\mathbf{x} \in M^m$ (objects) and $\mathbf{y} \in M^n$ (parameters). For each parameter tuple $\mathbf{y} \in M^n$, let

$$S_{\mathbf{y}} \,=\, \big\{\mathbf{x} \in M^m : \boldsymbol{M} \models F(\mathbf{x}; \mathbf{y})\big\}.$$

*Department of Mathematics, UCLA, Los Angeles, CA, 90095. Email: {ldnguyen,pak}@math.ucla.edu.
June 1, 2018.

Associated to F is the family $\mathcal{S}_F = \{S_{\mathbf{y}} : \mathbf{y} \in M^n\}$. We say that F is NIP, short for "F does *not* have the independence property", if $\mathcal{S}_F$ has finite VC-dimension. The structure $\boldsymbol{M}$ is called NIP if every partitioned $\mathcal{L}$-formula F is NIP in $\boldsymbol{M}$.

One prominent example of an NIP structure is *Presburger Arithmetic* $\mathrm{PA} = (\mathbb{Z}, <, +)$, which is the first-order structure on $\mathbb{Z}$ with only addition and inequalities. The main result of this paper are the lower and upper bounds on the VC-dimensions of PA-formulas. These are contrasted with the following notable bounds on the VC-density:

Theorem 1 ([A+16]). *Given a* PA*-formula* $F(\mathbf{x}; \mathbf{y})$ *with* $\mathbf{y} \in \mathbb{Z}^n$*,* $\mathrm{vc}(\mathcal{S}_F) \le n$ *holds.*

In other words, VC-density in the setting of PA can be bounded solely by the dimension of the parameter variables $\mathbf{y}$. It cannot grow very large when we vary the number of object variables $\mathbf{x}$, quantified variables or the description of F. This follows from a more general result in [A+16], which says that every *quasi-o-minimal* structure satisfies a similar bound on the VC-density. We refer to [A+16] for the precise statement of this result and for the powerful techniques used to bound the VC-density.

Karpinski and Macintyre raised a natural question whether similar bounds would hold for the VC-dimension. In [KM97], they gave upper bounds for the VC-dimension in some *o-minimal* structures (PA is not one), which are polynomial in the parameter dimension n. Later, they extended their arguments in [KM00] to obtain upper bounds on the *VC-density*, this time linear in n. Also in [KM00], the authors claimed to have an effective bound on the VC-dimensions of PA-formulas. However, we cannot locate such an explicit bound in any papers. To our knowledge, no effective upper bounds on the VC-dimensions of general PA-formulas exist in the literature.

1.3. **Main results.** We consider PA-formulas with a fixed number of variables (both quantified and free). Clearly, this also restricts the number of quantifier alternations in F. The atoms in F are linear inequalities in these variables with some integer constants and coefficients (in binary). Given such a formula F, denote by $\ell(F)$ the length of F, i.e., the total bit length of all symbols, operations, integer coefficients and constants in F.

We can further restrict the form of a PA-formula by requiring that it does not contain too many inequalities. For fixed k and t, denote by $\text{Short-PA}_{k,t}$ the family of PA-formulas with at most k variables (both free and quantified) and t inequalities. When k and t are clear, a formula $F \in \text{Short-PA}_{k,t}$ is simply called a *short Presburger formula*. In this case, $\ell(F)$ is essentially the total length of a bounded number of integer coefficients and constants. Our main result is a lower bound on the VC-dimension of short Presburger formulas:

Theorem 2. *For every d, there is a short Presburger formula $F(x; y) = \exists \mathbf{u}\, \forall \mathbf{v}\ \Psi(x, y, \mathbf{u}, \mathbf{v})$ in the class* $\text{Short-PA}_{10,18}$ *with*

$$\ell(F) = O(d^2) \quad \textit{and} \quad \mathrm{VC}(F) \ge d.$$

Here x, y are singletons and $\mathbf{u} \in \mathbb{Z}^6, \mathbf{v} \in \mathbb{Z}^2$. The expression Ψ is quantifier-free, and can be computed in probabilistic polynomial time in d.

So in contrast with VC-density, the VC-dimension of a PA-formula F crucially depends on the actual length $\ell(F)$. For the formulas in the theorem, we have:

$$\mathrm{VC}(F) = \Omega\big(\ell(F)^{1/2}\big), \quad \text{and} \quad \mathrm{vc}(F) \le 1,$$

where the last inequality follows by Theorem 1. Note that if one is allowed an unrestricted number of inequalities in F, a similar lower bound to Theorem 2 can be easily established by an elementary combinatorial argument. However, since the formula F is short, we can only work with a few integer coefficients and constants.

The construction in Theorem 2 uses a number-theoretic technique that employs continued fractions to encode a union of many arithmetic progressions. This technique was explored earlier in [NP17b] to show that various decision problems with short Presburger sentences are intractable. In this construction we need to pick a prime roughly larger than 4^d, which can be done in probabilistic polynomial time in d. This can be modified to a deterministic algorithm with run-time polynomial in d, at the cost of increasing $\ell(F)$:

Theorem 3. *For every d, there is a short Presburger formula $F(x; y) = \exists \mathbf{u}\, \forall \mathbf{v}\; \Psi(x, y, \mathbf{u}, \mathbf{v})$ in the class* Short-PA$_{10,18}$ *with*

$$\ell(F) \,=\, O(d^3) \quad \textit{and} \quad \mathrm{VC}(F) \geq d.$$

Here x, y are singletons and $\mathbf{u} \in \mathbb{Z}^6, \mathbf{v} \in \mathbb{Z}^2$. The expression Ψ is quantifier-free, and can be computed in deterministic polynomial time in d.

We conclude with the following polynomial upper bound for the VC-dimension of all (not necessarily short) Presburger formulas in a fixed number of variables:

Theorem 4. *For a Presburger formula $F(\mathbf{x}; \mathbf{y})$ with at most k variables (both free and quantified), we have:*

$$\mathrm{VC}(F) \,=\, O\big(\ell(F)^c\big),$$

where c and the $O(\cdot)$ constant depend only on k.

This upper bound implies that Theorem 2 is tight up to a polynomial factor. The proof of Theorem 4 uses an algorithm from [NP17a] for decomposing a semilinear set, i.e., one defined by a PA-formula, into polynomially many simpler pieces. Each such piece is a polyhedron intersecting a periodic set, whose VC-dimensions can be bounded by elementary arguments.

We note that the number of quantified variables is vital in Theorem 4. In §3.3, we construct PA-formulas $F(x; y)$ with x, y singletons and many quantified variables, for which $\mathrm{VC}(F)$ grows doubly exponentially compared to $\ell(F)$.

2. Proofs

We start with Theorem 3, and then show how it can be modified to give Theorem 2.

Proof of Theorem 3. Let $A = \{1, 2, \ldots, d\}$ and $\mathcal{S} = 2^A$. Since $\mathcal{S}$ contains all of the subsets of A, we have $\mathrm{VC}(\mathcal{S}) = d$. We order the sets in $\mathcal{S}$ lexicographically. In other words, for $S, S' \in \mathcal{S}$, we have $S < S'$ if $\sum_{i \in S} 2^i < \sum_{i \in S'} 2^i$. Thus, the sets in $\mathcal{S}$ can be indexed as $S_0 < S_1 < \cdots < S_{2^d - 1}$, where $S_0 = \varnothing, S_1 = \{1\}, \ldots, S_{2^d-1} = A$. Next, define:

$$T \coloneqq \bigsqcup_{0 \leq j < 2^d} \{i + dj \;:\; i \in S_j\}. \tag{2.1}$$

We show in Lemma 5 below that the set T is definable by a short PA formula $G_T(t)$ with only 8 quantified variables and 18 inequalities. Using this, it is clear that the parametrized formula

$$F_T(x; y) \;\coloneqq\; G_T(x + dy)$$

describes the family $\mathcal{S}$ (with y as the parameter), and thus has VC dimension d. We remark that G_T has only 1 quantifer alternation (see below). $\square$

Lemma 5. *The set T is definable by a short Presburger formula $G_T(t) = \exists \mathbf{u}\, \forall \mathbf{v}\ \Psi(t, \mathbf{u}, \mathbf{v})$ with $\mathbf{u} \in \mathbb{Z}^6, \mathbf{v} \in \mathbb{Z}^2$ and Ψ a Boolean combination of at most 18 inequalities in $t, \mathbf{u}, \mathbf{v}$ with binary length $\ell(\Psi) = O(d^3)$.*

Proof. Our strategy is to represent the set T as a union of arithmetic progressions (APs). In [NP17b], given d progressions $\mathrm{AP}_i = \{a_i, a_i + c_i, \ldots, a_i + b_i c_i\}$, we gave a method to define $\mathrm{AP}_1 \cup \cdots \cup \mathrm{AP}_d$ by a short Presburger formula of length polynomial in $\sum \log(a_i b_i c_i)$. For each $1 \leq i \leq d$, let $J_i = \{j : 0 \leq j < 2^d, i \in S_j\}$. From (2.1), we have:

$$T = \bigsqcup_{i=1}^{d} (i + dJ_i). \tag{2.2}$$

From the lexicographic ordering of the sets S_j, we can easily describe each set J_i as:

$$J_i = \{m + 2^{i-1} + 2^i n \ : \ 0 \leq m < 2^{i-1},\ 0 \leq n < 2^{d-i}\}. \tag{2.3}$$

So each set J_i is not simply an AP, but the Minkowski sum of two APs. However, we can easily modify each J_i into an AP by defining:

$$J_i' = \{2^d(m + 2^{i-1}) + 2^i n \ : \ 0 \leq m < 2^{i-1},\ 0 \leq n < 2^{d-i}\}. \tag{2.4}$$

It is clear that J_i' is an AP that starts at 2^{d+i-1} and ends at $2^{d+i} - 2^i$ with step size 2^i. Let $\mathrm{AP}_i \coloneqq i + dJ_i'$ and

$$T' = \bigsqcup_{i=1}^{d} \mathrm{AP}_i. \tag{2.5}$$

This is a union of d arithmetic progressions. Using the construction from [NP17b], we can define T' by a short Presburger formula:

$$t' \in T' \quad \Longleftrightarrow \quad \exists \mathbf{w} \quad \forall \mathbf{v} \quad \Phi\big(t', \mathbf{w}, \mathbf{v}\big),$$

where $t' \in \mathbb{Z},\ \mathbf{w}, \mathbf{v} \in \mathbb{Z}^2$ and Φ is a Boolean combination of at most 10 inequalities. This construction works by finding a single continued fraction $\alpha = [a_0\,; b_0, a_1, b_1, \ldots, a_{2d-1}]$ whose successive convergents encode the starting and ending points of our $\mathrm{AP}_1, \ldots, \mathrm{AP}_d$. We refer to Section 4 in [NP17b] for the details. The largest term in each AP_i is $\gamma_i = i + d(2^{d+i} - 2^i)$, which has binary length $O(d)$. Each term a_k and b_k in the continued fraction α is at most the product of these γ_i. Since $\prod_{i=1}^{d} \gamma_i$ has binary length $O(d^2)$, and so does each term a_k and b_k. Therefore, the final continued fraction α is a rational number p/q with binary length $O(d^3)$. This implies that $\ell(\Phi) = O(d^3)$ as well.

To get a formula for T, note that from (2.2), (2.3), (2.4) and (2.5), we have:

$$\begin{aligned} t \in T \quad \Longleftrightarrow \quad \exists\, t', i, r, s \ \ : \ \ & t' \in T', \quad 1 \leq i \leq d, \quad 0 \leq s < 2^d, \\ & t' = i + d(2^d r + s), \quad t = i + d(r + s).^1 \end{aligned}$$

Here r and s respectively stand for $m + 2^{i-1}$ and $2^i n$ in (2.3). Using $\exists \mathbf{w}\, \forall \mathbf{v}\ \Phi(t', \mathbf{w}, \mathbf{v})$ to express $t' \in T'$, we get a formula $G_T(t)$ defining T with 8 quantified variables $t', i, r, s \in \mathbb{Z}$, $\mathbf{w}, \mathbf{v} \in \mathbb{Z}^2$ and 18 inequalities. Note that t', i, r, s and $\mathbf{w}$ are existential variables, so G_T has the form $\exists \mathbf{u}\, \forall \mathbf{v}\ \Psi(t, \mathbf{u}, \mathbf{v})$ with $\mathbf{u} \in \mathbb{Z}^6, \mathbf{v} \in \mathbb{Z}^2$ and Ψ quantifier-free. $\square$

[1] Each equality is a pair of inequalities.

Proof of Theorem 2. Note that the above construction of F_T and G_T is deterministic with run-time polynomial in d. For Theorem 2, only the existence of a short PA formula with high VC-dimension is needed. In this case, our lower bound can be improved to $\mathrm{VC}(F) \geq c\sqrt{\ell(F)}$, for some $c > 0$, as follows. Recall that $\gamma_i = i + d(2^{d+i} - 2^i)$ is the largest term in $\mathrm{AP}_i = i + dJ'_i$ in (2.5). Pick the smallest prime p larger than $\max(\gamma_1, \ldots, \gamma_d) \approx d4^d$. This prime p can substitute for the large number M in Section 4.1 of [NP17b], which was (deterministically) chosen as $1 + \prod_{i=1}^d \gamma_i$, so that it is larger and coprime to all γ_i's. The rest of the construction follows verbatim. Note that $\log p = O(d)$ by Chebyshev's theorem. So the final continued fraction $\alpha = [a_0; b_0, a_1, b_1, \ldots, a_{2d-1}]$ has length $O(d^2)$, because now each term a_k, b_k has length at most $\log p$. This completes the proof. □

Proof of Theorem 4. Let $F(\mathbf{x};\mathbf{y})$ be any PA formula with free variables $\mathbf{x} \in \mathbb{Z}^m$, $\mathbf{y} \in \mathbb{Z}^n$ and n' other quantified variables, where m, n, n' are fixed. Let $k = m + n + n'$. In [NP17a] (Theorem 5.2), we gave the following polynomial decomposition on the semilinear set defined by F:

$$\Sigma_F \;\coloneqq\; \big\{(\mathbf{x},\mathbf{y}) \in \mathbb{Z}^{m+n} \;:\; F(\mathbf{x};\mathbf{y}) = \text{true}\big\} \quad = \quad \bigsqcup_{j=1}^{r} R_j \cap \boldsymbol{T}_j. \tag{2.6}$$

Here each R_j is a polyhedron in $\mathbb{R}^{m+n}$, and each $\boldsymbol{T}_j \subseteq \mathbb{Z}^{m+n}$ is a periodic set, i.e., a union of several cosets of some lattice $\mathcal{T}_j \subseteq \mathbb{Z}^{m+n}$. In other words, the set defined by F is a union of r pieces, each of which is a polyhedron intersecting a periodic set. Our decomposition is algorithmic, in the sense that the pieces R_j and lattices $\mathcal{T}_j$ can be found in time $O\big(\ell(F)^c\big)$, with c and $O(\cdot)$ depending only on k. The algorithm describes each piece R_j by a system of inequalities and each lattice $\mathcal{T}_j$ by a basis. Denote by $\ell(R_j)$ and $\ell(\mathcal{T}_j)$ the total binary lengths of these systems and basis vectors, respectively. These also satisfy:

$$\sum_{j=1}^{r} \ell(R_j) + \ell(\mathcal{T}_j) = O\big(\ell(F)^c\big). \tag{2.7}$$

Each R_j can be written as the intersection $H_{j1} \cap \cdots \cap H_{jf_j}$, where each H_{jk} is a half-space in $\mathbb{R}^{m+n}$, and f_j is the number of facets of R_j. Note that $f_j \leq \ell(R_j) = O\big(\ell(F)^c\big)$. We rewrite (2.6) as:

$$\Sigma_F \quad = \quad \bigsqcup_{j=1}^{r} H_{j1} \cap \cdots \cap H_{jf_j} \cap \boldsymbol{T}_j. \tag{2.8}$$

Therefore, the set Σ_F is a Boolean combination of $f_1 + \cdots + f_r$ half-spaces and r periodic sets. In total, there are

$$f_1 + \cdots + f_r + r = O\big(\ell(F)^c\big) \tag{2.9}$$

of those basic sets.

For a set $\Gamma \subseteq \mathbb{R}^{m+n}$ and $\mathbf{y} \in \mathbb{Z}^n$, denote by $\Gamma_{\mathbf{y}}$ the subset $\{\mathbf{x} \in \mathbb{Z}^m : (\mathbf{x},\mathbf{y}) \in \Gamma\}$ and by $\mathcal{S}_\Gamma$ the family $\{\Gamma_{\mathbf{y}} : \mathbf{y} \in \mathbb{Z}^n\}$. For a half-space $H \subset \mathbb{R}^{m+n}$, it is easy to see that $\mathrm{VC}(\mathcal{S}_H) = 1$. For each periodic set $\boldsymbol{T}_j$ with period lattice $\mathcal{T}_j$, the family $\mathcal{S}_{\boldsymbol{T}_j}$ has cardinality at most $\det(\mathcal{T}_j \cap \mathbb{Z}^n) \leq 2^{O(\ell(\mathcal{T}_j))}$. Thus, we have

$$\mathrm{VC}(\mathcal{S}_{\boldsymbol{T}_j}) \leq \log|\mathcal{S}_{\boldsymbol{T}_j}| = O\big(\ell(\mathcal{T}_j)\big). \tag{2.10}$$

Let $\Gamma_1, \ldots, \Gamma_t \subseteq \mathbb{Z}^{m+n}$ be any t sets with $\mathrm{VC}(\mathcal{S}_{\Gamma_i}) = d_i$. By an application of the Sauer-Shelah lemma ([Sa72, Sh72]), if Σ is any Boolean combination of $\Gamma_1, \ldots, \Gamma_t$, then we can bound $\mathrm{VC}(\mathcal{S}_\Sigma)$ as:

$$\mathrm{VC}(\mathcal{S}_\Sigma) = O\big((d_1 + \cdots + d_t)\log(d_1 + \cdots + d_t)\big).$$

Applying this to (2.8), we get $\mathrm{VC}(\mathcal{S}_{\Sigma_F}) = O(\ell \log \ell)$, where

$$\ell \;=\; \sum_{j=1}^{r} \left(\mathrm{VC}(\mathcal{S}_{\boldsymbol{T}_j}) + \sum_{j'=1}^{f_j} \mathrm{VC}(\mathcal{S}_{H_{jj'}}) \right) \;\leq\; \sum_{j=1}^{r} \mathrm{VC}(\mathcal{S}_{\boldsymbol{T}_j}) + f_j.$$

By (2.7), (2.9) and (2.10), we have $\ell = O\big(\ell(F)^c\big)$. We conclude that $\mathrm{VC}(F) = O\big(\ell(F)^{2c}\big)$. $\square$

3. Final remarks and open problems

3.1. The proof of Theorem 2 is almost completely effective except for finding a small prime p larger than a given integer N. This problem is considered to be computationally very difficult in the deterministic case, and only exponential algorithms are known (see [LO87, TCH12]).

3.2. Our constructed short formula F is of the form $\exists\forall$. It is interesting to see if similar polynomial lower bounds are obtainable with existential short formulas. For such a formula $F(\mathbf{x}; \mathbf{y}) = \exists \mathbf{z}\, \Phi(\mathbf{x}, \mathbf{y}, \mathbf{z})$, the quantifier-free expression $\Phi(\mathbf{x}, \mathbf{y}, \mathbf{z})$ captures the set of integer points Γ lying in a union of some polyhedra P_i's. Note that the total number of polyhedra and their facets should be bounded, since we are working with short formulas. Therefore, F simply capture the pairs $(\mathbf{x}, \mathbf{y})$ in the projection of Γ along the $\mathbf{z}$ direction. Denote this set by $\mathrm{proj}(\Gamma)$. The work of Barvinok and Woods [BW03] shows that $\mathrm{proj}(\Gamma)$ has a *short generating function*, and can even be counted efficiently in polynomial time. In our construction, the set that yields high VC-dimension is a union arithmetic progressions, which cannot be counted efficiently unless $\mathsf{P} = \mathsf{NP}$ (see [SM73]). This difference indicates that $\mathrm{proj}(\Gamma)$ has a much simpler combinatorial structure, and may not attain a high VC-dimension.

3.3. One can ask about the VC-dimension of a general PA-formula with no restriction on the number of variables, quantifier alternations or atoms. Fischer and Rabin famously showed in [FR74] that PA has decision complexity at least doubly exponential in the general setting. For every $\ell > 0$, they constructed a formula $\mathrm{Prod}_\ell(a, b, c)$ of length $O(\ell)$ so that for every triple

$$0 \leq a, b, c \,<\, 2^{2^{2^\ell}},$$

we have $\mathrm{Prod}_\ell(a, b, c) = \text{true}$ if and only if $ab = c$. Using this "partial multiplication" relation, one can easily construct a formula $F_\ell(x; y)$ of length $O(\ell)$ and VC-dimension at least 2^{2^ℓ}. This can be done by constructing a set similar to T in (2.1) with d replaced by 2^{2^ℓ} using Prod_ℓ. We leave the details to the reader.

Regarding upper bound, Oppen showed in [Opp78] that any PA-formula F of length ℓ is equivalent to a quantifier-free formula G of length $2^{2^{2^{c\ell}}}$ for some universal constant $c > 0$. This implies that $\mathrm{VC}(G)$, and thus $\mathrm{VC}(F)$, is at most triply exponential in $\ell(F)$. We conjecture that a doubly exponential upper bound on $\mathrm{VC}(F)$ holds in the general setting.

It is unlikely that such an upper bound could be established by straightforward quantifier elimination, which generally results in triply exponential blow up (see [Wei97, Thm 3.1]).

Acknowledgements. We are grateful to Matthias Aschenbrenner and Artëm Chernikov for many interesting conversations and helpful remarks. This paper was finished while both authors were visitors at MSRI; we are thankful for the hospitality, great work environment and its busy schedule. The second author was partially supported by the NSF.

References

[A+16] M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson and S. Starchenko, Vapnik-Chervonenkis density in some theories without the independence property, I, *Trans. AMS* **368** (2016), 5889–5949.

[BW03] A. Barvinok and K. Woods, Short rational generating functions for lattice point problems, *Jour. AMS* **16** (2003), 957–979.

[Che16] A. Chernikov, *Models theory and combinatorics*, course notes, UCLA; available electronically at https://tinyurl.com/y8ob6uyv.

[FR74] M. J. Fischer and M. O. Rabin, Super-Exponential Complexity of Presburger Arithmetic, in *Proc. SIAM-AMS Symposium in Applied Mathematics*, AMS, Providence, RI, 1974, 27–41.

[KM97] M. Karpinski and A. Macintyre, Polynomial bounds for VC dimension of sigmoidal and general Pfaffian neural networks, *J. Comput. System Sci.* **54** (1997), 169–176.

[KM00] M. Karpinski and A. Macintyre, Approximating volumes and integrals in o-minimal and p-minimal theories, in *Connections between model theory and algebraic and analytic geometry*, Seconda Univ. Napoli, Caserta, 2000, 149–177.

[NP17a] D. Nguyen and I. Pak, Enumeration of integer points in projections of unbounded polyhedra, *SIAM J. Discrete Math.* **32** (2018), 986–1002.

[NP17b] D. Nguyen and I. Pak, Short Presburger Arithmetic is hard, in *Proc. 58th FOCS*, IEEE, Los Alamitos, CA, 2017, 37–48; `arXiv:1708.08179`.

[LO87] J. C. Lagarias and A. M. Odlyzko, Computing $\pi(x)$: an analytic method, *J. Algorithms* **8** (1987), 173–191.

[Opp78] D. C. Oppen, A $2^{2^{2^{pn}}}$ upper bound on the complexity of Presburger arithmetic, *J. Comput. System Sci.* **16** (1978), 323–332.

[Sa72] N. Sauer, On the density of families of sets, *J. Combin. Theory, Ser. A* **13** (1972), 145–147.

[Sh72] S. Shelah, A combinatorial problem; stability and order for models and theories in infinitary languages, *Pacific J. Math.* **41** (1972), 247–261.

[SM73] L. J. Stockmeyer and A. R. Meyer, Word problems requiring exponential time: preliminary report, in *Proc. Fifth STOC*, ACM, New York, 1973, 1–9.

[TCH12] T. Tao, E. Croot and H. Helfgott, Deterministic methods to find primes, *Math. Comp.* **81** (2012), 1233–1246.

[VC71] V. N. Vapnik and A. Ja. Červonenkis, The uniform convergence of frequencies of the appearance of events to their probabilitie, *Theor. Probability Appl.* **16** (1971), 264–280.

[Vap98] V. N. Vapnik, *Statistical learning theory*, John Wiley, New York, 1998.

[Wei97] V. D. Weispfenning, Complexity and uniformity of elimination in Presburger arithmetic, in *Proc. 1997 ISSAC*, ACM, New York, 1997, 48–53.