Shattered Sets and the Hilbert Function

Shay Moran[1] **and Cyrus Rashtchian**[2]

1 **Department of Computer Science, Technion, Israel; and Microsoft Research, Hertzelia, Israel; and Max Planck Institute for Informatics, Saarbrücken, Germany**
shaymoran1@gmail.com

2 **Department of Computer Science and Engineering, University of Washington, Seattle, WA, USA**
cyrash@cs.washington.edu

Abstract

We study complexity measures on subsets of the boolean hypercube and exhibit connections between algebra (the Hilbert function) and combinatorics (VC theory). These connections yield results in both directions. Our main complexity-theoretic result demonstrates that a large and natural family of linear program feasibility problems cannot be computed by polynomial-sized constant-depth circuits. Moreover, our result applies to a stronger regime in which the hyperplanes are fixed and only the directions of the inequalities are given as input to the circuit. We derive this result by proving that a rich class of extremal functions in VC theory cannot be approximated by low-degree polynomials. We also present applications of algebra to combinatorics. We provide a new algebraic proof of the Sandwich Theorem, which is a generalization of the well-known Sauer-Perles-Shelah Lemma. Finally, we prove a structural result about downward-closed sets, related to the Chvátal conjecture in extremal combinatorics.

1998 ACM Subject Classification F.1.3 Complexity Measures and Classes, G.2.3 Applications

Keywords and phrases VC Dimension, Shattered Sets, Sandwich Theorem, Hilbert Function, Polynomial Method, Linear Programming, Chvatal's Conjecture, Downward-closed Sets.

Digital Object Identifier 10.4230/LIPIcs.MFCS.2016.70

1 Introduction

Understanding the properties and structure of subsets of the boolean hypercube is a central theme in theoretical computer science and combinatorics. When studying a family of mathematical objects, endowing the objects with algebraic structure often sheds new light on interesting properties. This phenomena appears classically in areas such as algebraic topology and algebraic geometry. It also provides much utility when studying the boolean hypercube. Let $C \subseteq \{0,1\}^n$ be a subset of the boolean hypercube, and let $\mathbb{F}$ be a field. Consider the linear space of functions from C to $\mathbb{F}$, that is, $\mathbb{F}^C$. This is clearly a $|C|$-dimensional vector space over $\mathbb{F}$. Every function in this space can be represented as a multilinear polynomial with degree at most n. Interestingly, for certain sets C, smaller degree actually suffices. For example, when C is the standard basis, denoted $C = \{\vec{e}_1, \ldots, \vec{e}_m\}$, then any function $f : C \to \mathbb{F}$ can be expressed as the linear function $f(\vec{e}_1)x_1 + \ldots + f(\vec{e}_m)x_m$.

The Hilbert function, denoted $h_d(C, \mathbb{F})$, is the dimension of the space of functions $\{f : C \to \mathbb{F}\}$ that have representations as polynomials with degree at most d. This classical algebraic object will be useful in our study of how the structure of C affects the function space. In complexity theory, Smolensky [40] has used the Hilbert function to unify polynomial approximation lower bounds relating to bounded-depth circuits.

© Shay Moran and Cyrus Rashtchian;
licensed under Creative Commons License CC-BY
41st International Symposium on Mathematical Foundations of Computer Science (MFCS 2016).
Editors: Piotr Faliszewski, Anca Muscholl, and Rolf Niedermeier; Article No. 70; pp. 70:1–70:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

We establish new connections between the Hilbert function and VC theory. Our main technical contributions are bounds $h_d(C, \mathbb{F})$ in terms of basic concepts in VC theory, such as shattering, strong shattering, and down-shifts. Previous results on bounding the Hilbert function utilize a more intricate analysis and focus on symmetric sets, that is, unions of slices of the hypercube [40, 11]. In addition to giving new bounds on the Hilbert function, our connection between Algebra and Combinatorics allows us to derive results in both directions.

Our main complexity theoretical application is that determining feasibility of a large family of linear programs is hard for the class of bounded-depth circuits. More specifically, let $h_1, \ldots, h_m$ be affine functions. Each sign vector s in $\{\pm\}^m$ defines the following feasibility problem: does there exist $x \in \mathbb{R}^d$ such that $h_i(x) > 0$ when $s_i = +$, and $h_i(x) < 0$ when $s_i = -$, for all $i \in [m]$? This defines a boolean function that takes an input s and outputs one if and only if the problem is feasible. We prove that if $m = 2d + 1$, and the affine functions h_i are in general position, then this function cannot be approximated by low-degree polynomials, over any field. This implies a lower bound on the computability of this function by constant-depth circuits, due to the polynomial approximation technique introduced by Razborov [37] and Smolensky [41]. The above linear programming problem relates to the study of hyperplane arrangements (see the books of Matoušek [30] and Stanley [43] for more details and applications). Our results implicitly provide algebraic proofs of some known facts regarding the combinatorics of hyperplane arrangements.

As a combinatorial application of our bounds on the Hilbert function, we provide a short algebraic proof of the Sandwich Theorem. This theorem comes from VC theory and is a well-studied generalization of the Sauer-Shelah-Perles Lemma [29, 36, 15, 14, 18, 5, 8, 33, 28, 32, 35]. Similar proofs of related upper bounds have appeared previously, due to Frankl and Pach [7], Gurvits [24], and Smolensky [42]. We contribute new lower bounds and applications.

Facts we prove about the function space $\mathbb{F}^C$ also lead to a new result about downward-closed sets. A family D of subsets is *downward-closed* if $b \subseteq a$ and $a \in D$ implies $b \in D$. A theorem of Berge [10] implies that for any downward-closed set D there exists a bijection $\pi : D \to D$ such that $a \cap \pi(a) = \emptyset$ for all $a \in D$. We generalize his result to arbitrary, prescribed intersections. Let $\phi : D \to D$ have the property $\phi(a) \subseteq a$ for all $a \in D$. We show that there always exists a bijection $\pi : D \to D$ such that $a \cap \pi(a) = \phi(a)$. Note that choosing $\phi(a) = \emptyset$ for all a implies Berge's result.

Our algebra-combinatorics connection fits within the framework of the polynomial method. This method has been successful in providing elegant proofs of fundamental results in many areas, such as circuit complexity [41, 6, 37, 9], discrete geometry [25, 19, 39, 44], extremal combinatorics [1, 26, 7], and more.

The paper is organized as follows. We state our main theorems in Section 2. In Section 3, we prove our bounds on the Hilbert function. In Section 4, we use our Hilbert function bounds to prove that linear program feasibility is hard for bounded-depth circuits. Finally, in Section 5, we prove results about downward-closed sets. We now review preliminaries.

1.1 Preliminaries

We begin with algebraic preliminaries. Let $C \subseteq \{0, 1\}^n$ and $\mathbb{F}$ be a field. Every $f : C \to \mathbb{F}$ can be expressed as a multilinear polynomial over variables $x_1, \ldots, x_n$ with coefficients in $\mathbb{F}$.

▶ **Definition 1.** For $d \in [n]$ the *Hilbert function* $h_d(C, \mathbb{F})$ is the dimension of the space of functions $f : C \to \mathbb{F}$ that can be represented as polynomials with degree at most d.

Notice that $h_d(C, \mathbb{F}) \leq \min\{\sum_{j=0}^{d} \binom{n}{j}, |C|\}$. A basic fact about the Hilbert function is that

$$1 = h_0(C, \mathbb{F}) \leq h_1(C, \mathbb{F}) \leq \ldots \leq h_n(C, \mathbb{F}) = |C|.$$

The final equality holds because all $f : C \to \mathbb{F}$ have representations with degree at most n.

It is natural to wonder when the Hilbert function attains its maximum and how the structure of C influences the Hilbert function. We introduce the following measure.

▶ **Definition 2.** The **interpolation degree of** C denoted $\mathsf{intdeg}(C, \mathbb{F})$ is the minimum d such that any $f : C \to \mathbb{F}$ can be expressed as a multilinear polynomial with degree at most d. In other words,

$$\mathsf{intdeg}(C, \mathbb{F}) = \min\{d \in [n] : h_d(C, \mathbb{F}) = |C|\}.$$

Intuitively, a smaller interpolation degree implies a less complex set.

We move on to combinatorial preliminaries. Our bounds on the Hilbert function use basic concepts in VC theory. We define these concepts now.

▶ **Definition 3.** A subset $I \subseteq [n]$ is **shattered** by $C \subseteq \{0, 1\}^n$ if for every pattern $s : I \to \{0, 1\}$ there is $c \in C$ that realizes s. In other words, the restriction of c to I equals s. A subset $I \subseteq [n]$ is **strongly shattered** by C if C contains all elements of some subcube on I. In other words, there exists a pattern $\bar{s} : ([n] \setminus I) \to \{0, 1\}$ such that all extensions of $\bar{s}$ to a vector in $\{0, 1\}^n$ are in C.

These definitions lead to natural families of sets, which will be important to our work.

▶ **Definition 4.** The **shattered sets** with respect to C are

$$\mathsf{str}(C) = \{I \subseteq [n] : I \text{ is shattered by } C\}.$$

The **strongly shattered sets** with respect to C are

$$\mathsf{sstr}(C) = \{I \subseteq [n] : I \text{ is strongly shattered by } C\}.$$

▶ **Definition 5.** The **VC dimension** of C is defined as $\mathsf{VC}(C) = \max\{|I| : I \in \mathsf{str}(C)\}$.

Note that $\mathsf{sstr}(C) \subseteq \mathsf{str}(C)$ and that both of these families are downward-closed.

We also lower bound the Hilbert function using down-shifts, a standard tool in extremal combinatorics. Let $C \subseteq \{0, 1\}^n$ and let $i \in [n]$. We denote as S_i the down-shift operator on the i'th coordinate. Obtain the set $S_i(C) \subseteq \{0, 1\}^n$ from C as follows. Replace every $c \in C$ such that both (i) $c_i = 1$, and (ii) the i'th neighbor[1] of c is not in C with the i'th neighbor of c. Authors have referred to this operation as “compression”, “switching”, and “polarization’. Previous works that use down-shifts include [27, 20, 13, 22, 23, 33].

An important property of down-shifts is that they transform an arbitrary subset of $\{0, 1\}^n$ into a downward-closed set, without changing cardinality. Specifically, if

$$D = S_n(S_{n-1}(\cdots S_1(C)))$$

is the result of sequentially applying S_i on C for each i, then D is downward-closed. It is also convenient in this context to think of D as a family of subsets of $[n]$ rather than a set of boolean vectors via the natural correspondence between boolean vectors and sets.

We move on to explaining our results in more formality and detail.

[1] Vectors $u, v \in \{0, 1\}^n$ are *i'th-neighbors* if they differ in coordinate i and are the same elsewhere.

2 Our Results

We start with the result about linear program feasibility. We then state the bounds on the Hilbert function in terms of shattered sets and down-shifts. We show this leads to bounded-depth circuit lower bounds. Finally, we state two combinatorial applications.

2.1 Linear Program Feasibility

We formalize and prove a strong version of the statement "linear programming feasibility can not be decided by polynomial-sized, constant-depth circuits." Clearly, linear programming being P-complete [17] implies a version of this statement for specific linear programs representing functions previously known not to have efficient bounded-depth circuits. We prove a stronger version stating that any linear feasibility problem, in which the number of constraints is roughly twice the number of variables and the constraints are non-degenerate, cannot be decided by an efficient bounded-depth circuit. For a set of hyperplanes $\mathcal{H}$ in $\mathbb{R}^k$ we will define a boolean function $f_{\mathcal{H}}$. It takes orientations as inputs and outputs one if and only if a certain polytope is nonempty. In particular, we establish hardness of this problem even when the hyperplanes are fixed in advance and only the orientations are given as input.

We express linear program feasibility as a boolean function as follows. Specify an arrangement of m hyperplanes $\mathcal{H} = \{h_1, \dots, h_m\}$ with normal vectors $\vec{n}_i$ and translation scalars b_i as

$$h_i = \{\vec{x} : \langle \vec{n_i}, \vec{x} \rangle = b_i\}.$$

A sign pattern $s \in \{-1, 1\}^m$ encodes the following linear programming feasibility problem:

Does there exist $\vec{x} \in \mathbb{R}^k$ satisfying $\mathsf{sign}(\langle x, \vec{n}_i \rangle - b_i) = s_i$ for all $i \in [m]$?

This corresponds to checking the feasibility of a linear program with m constraints and k variables. Define $f_{\mathcal{H}} : \{-1, 1\}^m \to \{0, 1\}$ as the boolean function such that $f_{\mathcal{H}}(s) = 1$ if and only if the linear program encoded by s is feasible.

As an example, consider the following arrangement in $\mathbb{R}^2$. The three hyperplanes

$$h_1 = \{(x_1, x_2) : 5x_1 + 3x_2 = 3\},$$
$$h_2 = \{(x_1, x_2) : 8x_1 - x_2 = 8\},$$
$$h_3 = \{(x_1, x_2) : 4x_1 - 5x_2 = 0\}$$

form an arrangement of three lines in the plane. The vector $s = (+1, -1, +1)$ encodes the system

$$5x_1 + 3x_2 > 3 \qquad (s(1) = +1)$$
$$8x_1 - x_2 < 8 \qquad (s(2) = -1)$$
$$4x_1 + 5x_2 > 0 \qquad (s(3) = +1)$$

In the example, the system encoded by $(+1, -1, +1)$ is not satisfiable (see Figure 1). For more background material on hyperplane arrangements and related results, see the books by Stanley [43] and Matoušek [30].

We prove the following theorem.

▶ **Theorem 6.** *Let $\mathcal{H}$ be an arrangement of $2k + 1$ hyperplanes in $\mathbb{R}^k$ that are in general position. Any $AC^0[p]$ circuit, for a prime p, with depth d computing $f_{\mathcal{H}}$ requires $exp(\Omega(k^{1/2d}))$ gates.*

We prove Theorem 6 in Section 4, using the framework of Razborov [37] and Smolensky [41].

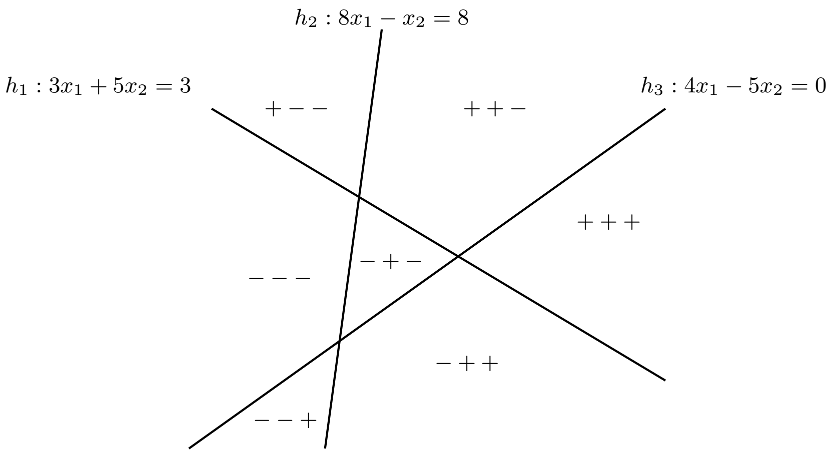

Figure 1 Three lines divide $\mathbb{R}^2$ into seven regions, each labeled by a feasible sign pattern.

Explicit Arrangements. The space of oriented hyperplanes is a rich and well-studied object. The books [30, 43] provide many facts and examples. The paper [2] and references therein give bounds on how many different boolean functions can be represented as $f_{\mathcal{H}}$ for some hyperplane arrangement $\mathcal{H}$.

General position hyperplane arrangements come from any $2k+1$ vectors in $\mathbb{R}^{k+1}$ such that every $k+1$ of them are linearly independent. For a vector $v \in \mathbb{R}^{k+1}$ the hyperplane has normal $(v_1, \ldots, v_k)$ and translation v_{k+1}. Explicit families of m vectors in $\mathbb{R}^d$ such that every d of them are independent are known for any m, d. For example, take the rows of an $m \times d$ Cauchy or Vandermonde matrix.

2.2 Hilbert Function Bounds

Our results are based on the following theorem.

▶ **Theorem 7.** *Any $C \subseteq \{0,1\}^n$ and any $d \in [n]$ satisfy the relationships*

$$|\{I \in \mathsf{sstr}(C) \ : \ |I| \leq d\}| \ \leq \ h_d(C, \mathbb{F}) \ \leq \ |\{I \in \mathsf{str}(C) \ : \ |I| \leq d\}|$$

and

$$\max\{|I| : I \in \mathsf{sstr}(C)\} \ \leq \ \mathsf{intdeg}(C, \mathbb{F}) \ \leq \ \max\{|I| : I \in \mathsf{str}(C)\}.$$

The upper bounds on interpolation degree are not new. Smolensky [42] derives the Sauer-Perles-Shelah Lemma using very similar polynomial-based arguments. The upper bounds on interpolation degree in terms of VC dimension also appear implicitly in the work of Frankl and Pach [7] and explicitly in Gurvits [24]. Our technical contributions center around the lower bounds and the applications. We prove Theorem 7 in Section 3.1.

We strengthen the lower bound on the Hilbert function in Theorem 7 using down-shifts.

▶ **Theorem 8.** *Let $C \subseteq \{0,1\}^n$ and let $D = S_n(S_{n-1}(\ldots S_1(C)))$. Then*

$$|\{I \in D \ : \ |I| \leq d\}| \ \leq \ h_d(C, \mathbb{F}) \quad \text{and} \quad \max\{|I| : I \in D\} \ \leq \ \mathsf{intdeg}(C, \mathbb{F}).$$

In Section 3.2 we prove this theorem and show that the parity function provides a tight example over $GF(2)$. We also discuss how Theorem 8 implies the lower bound in Theorem 7.

2.3 Low-Degree Polynomial Approximations

Classic results in bounded-depth circuit complexity reduce the task of proving circuit lower bounds to showing that a boolean function has no low-degree approximation [37, 41, 6]. Smolensky shows in [40] how to express all known degree lower bounds in terms of the Hilbert function. For a boolean function f consider the set $S = f^{-1}(1)$ as a subset of the boolean cube. Smolensky shows that if $h_d(S, \mathbb{F})$ is large, then f is hard to approximate.

▶ **Theorem 9** ([40]). *Consider* $f : \{0,1\}^n \to \{0,1\}$ *and* $p : \{0,1\}^n \to \mathbb{F}$. *Define* $S = f^{-1}(1)$ *and fix* $d = \lfloor (n - \deg_{\mathbb{F}}(p) - 1)/2 \rfloor$. *Then,*

$$\Pr_x[p(x) \neq f(x)] \geq \frac{2 \cdot h_d(S, \mathbb{F}) - |S|}{2^n},$$

where x is uniform over $\{0,1\}^n$.

Theorem 7 implies the following corollary in terms of strongly shattered sets.

▶ **Corollary 10.** *Assume n is odd. Consider* $f : \{0,1\}^n \to \{0,1\}$. *If* $|f^{-1}(1)| = 2^{n-1}$ *and* $\mathsf{sstr}(f^{-1}(1)) = \{I \subseteq [n] : |I| \leq \frac{n-1}{2}\}$, *then for any polynomial* $p \in \mathbb{F}[x_1, \ldots, x_n]$ *we have*

$$\Pr_x[p(x) \neq f(x)] \geq \frac{1}{2} - \frac{10 \deg_{\mathbb{F}}(p)}{\sqrt{n}},$$

where x is uniform over $\{0,1\}^n$.

Proof. Since $\mathsf{sstr}(f^{-1}(1)) = \{I \subseteq [n] : |I| \leq \frac{n-1}{2}\}$, we have that

$$|\{I \in \mathsf{sstr}(C) \ : \ |I| \leq d\}| = \sum_{j=0}^{d} \binom{n}{j}$$

for all $d = 0, 1, \ldots, (n-1)/2$. Theorem 7 implies that $h_d(f^{-1}(1), \mathbb{F}) = \sum_{j=0}^{d} \binom{n}{j}$ as well. Plugging these into Theorem 9, along with $|f^{-1}(1)| = 2^{n-1}$, gives the corollary. ◀

Bernasconi and Egidi [11] thoroughly characterize the Hilbert function for symmetric sets and prove that any nearly-balanced, symmetric boolean function is hard to approximate. They leave as an open question deriving bounds for non-symmetric sets. Our connection to VC theory leads to new families of functions satisfying the conditions of Corollary 10. Many of these functions, such as the linear programming feasibility functions from Section 2.1, are non-monotone and non-symmetric. As a final remark, recent work shows that Smolensky's lower bound (and thus our result) extends to nonclassical polynomials [12].

2.4 The Sandwich Theorem

The following relationship which is a generalization of the Sauer-Perles-Shelah Lemma was discovered several times and independently [14, 36, 18, 5].

▶ **Theorem 11** (Sandwich Theorem). *For any* $C \subseteq \{0,1\}^n$ *we have* $|\mathsf{sstr}(C)| \leq |C| \leq |\mathsf{str}(C)|$.

Since $|\mathsf{str}(C)| \leq \sum_{i=0}^{\mathsf{VC}(C)} \binom{n}{i}$, this implies the Sauer-Perles-Shelah Lemma.

Theorem 7 yields a new algebraic proof of the Sandwich Theorem. Indeed, this follows from examining the case of $d = n$ and observing that $h_n(C, \mathbb{F}) = |C|$.

The Sandwich Theorem is tight in the sense that there are sets that achieve equality in both of its inequalities[2]. These sets are calles *shattering extremal sets*. For example, downward-closed sets are shattering extremal. Shattering extremal sets have been rediscovered and studied in different contexts [29, 15, 14, 18, 8, 33, 28, 32, 35]. In our context, Corollary 10 says that shattering extremal sets S of size $|S| = 2^{n-1}$ and VC dimension $\frac{n-1}{2}$ correspond to boolean functions that cannot be approximated by low-degree polynomials.

2.5 Downward-closed Sets and Chvátal's Conjecture

Downward-closed sets have a well-studied, rich combinatorial structure. A theorem of Berge [10] implies the following fact. For any downward-closed set D, there is a bijection $\pi : D \to D$ such that $a \cap \pi(a) = \emptyset$, for all $a \in D$. We refer to such a bijection as a *pseudo-complement*. We prove the following generalization of the existence of a pseudo-complement.

▶ **Theorem 12.** *Let D be any downward-closed set. Fix any mapping $\phi : D \to D$ with the property that $\phi(a) \subseteq a$ for all $a \in D$. Then there exists a bijection $\pi : D \to D$ satisfying the condition that $a \cap \pi(a) = \phi(a)$ for all $a \in D$.*

Note that choosing $\phi(a) = \emptyset$ for all a implies the existence of a pseudo-complement.

In topology, downward-closed sets correspond to simplicial complexes. We think of the ϕ as prescribing intersections. For simplicial complexes, this corresponds to prescribing that complexes intersect in certain faces. We prove Theorem 12 in Section 5. Our proof proceeds by proving that a certain matrix is invertible. A non-zero determinant implies that the matrix contains a permutation matrix that yields the desired bijection.

We next discuss the result by Berge for the existence of pseudo-complements and its connections with Chvátal's conjecture in extremal combinatorics [16]. Berge's result about pseudo-complements follows from the following stronger theorem that he proved.

▶ **Theorem 13** ([10]). *If D is a downward-closed set, then either D or $D \setminus \emptyset$ can be partitioned into pairs of disjoint sets.*

We need two definitions to explain Berge's motivation. A family B of subsets of $[n]$ is called a *star* if there is an element $x \in [n]$ such that $x \in b$ for all $b \in B$. It is called an *intersecting family* if every pair of sets in B intersects. Chvátal's conjecture is the following.

▶ **Conjecture 14** (Chvátal's conjecture). *If D is a downward closed set, then the cardinality of the largest star in D is equal to the cardinality of the largest intersecting family in D.*

This conjecture remains open, aside from partial results, such as the following corollary of Berge's theorem.

▶ **Corollary 15.** *In a downward-closed set D, any intersecting family has cardinality at most $|D|/2$.*

We contrast Berge's theorem and our Theorem 12. Berge's pair decomposition induces a permutation π such that $\pi(\pi(a)) = a$, whereas a permutation decomposes D into disjoint cycles with unspecified lengths. Many people have observed that the above corollary only needs the pseudo-complement result, instead of the stronger statement in Berge's theorem [4]. Indeed, consider each disjoint cycle in the guaranteed permutation, and note that at most half of the sets in the cycle may mutually intersect. Therefore, our Theorem 12 implies the above corollary.

[2] In fact, it is well known (see for example [33]) that any set achieving equality in one of the inequalities, also achieves equality in the other.

3 The Hilbert Function for Subsets of the Boolean Cube

We prove upper and lower bounds on the Hilbert function. First, we prove the bounds in Theorem 7 involving the shattered and the strongly shattered sets. Then, we prove the bounds in Theorem 8 using shifting. Finally we consider an example of applying these bounds to analyze the Hilbert function of the parity function.

3.1 Bounding the Hilbert Function Using Shattered Sets

The high-level idea of the proof of Theorem 7 is to define a vector space V with $\dim(V) = |C|$ and prove that $|\mathsf{sstr}(C)| \leq \dim(V) \leq |\mathsf{str}(C)|$. We sandwich the dimension $\dim(V)$ by finding a linearly independent set of size $|\mathsf{sstr}(C)|$ and a spanning set of size $|\mathsf{str}(C)|$.

We analyze the $|C|$-dimensional vector space $\{f : C \to \mathbb{F}\}$. Evaluation on C induces a natural mapping from $P \in \mathbb{F}[x_1, \ldots, x_n]$ to the restriction $P|_C \in \{f : C \to \mathbb{F}\}$. The following lemma provides the desired sets of spanning monomials and linearly independent monomials.

▶ **Lemma 16.** *For all fields $\mathbb{F}$ and sets $C \subseteq \{0,1\}^n$ the following two facts hold.*
1. *The monomials $\prod_{i \in I} x_i$ for $I \in \mathsf{str}(C)$ span $\{f : C \to \mathbb{F}\}$.*
2. *The monomials $\prod_{i \in I} x_i$ for $I \in \mathsf{sstr}(C)$ are linearly independent in $\{f : C \to \mathbb{F}\}$.*

Proof. For $I \subseteq [n]$, let x_Y denote the monomial $x_I = \prod_{i \in I} x_i$. For the first item, we express every $f : C \to \mathbb{F}$ as a linear combination of monomials $(x_I)|_C$ where $I \in \mathsf{str}(C)$. It suffices to express the monomials $(x_I)|_C$ for all $I \subseteq [n]$. We prove this by induction. For the base case, if $I \in \mathsf{str}(C)$, we are done. Otherwise, I is not shattered by C and there exists $s \in \{0,1\}^I$ such that for all $c \in C$, we have $c|_I \neq s$. Consider

$$P = \prod_{i \in I} (x_i - (1 - s_i)) .$$

Note that $P(c) = 0$ for all $c \in C$ and hence $P|_C = 0|_C$. Specifically, by expanding the product $\prod_{i \in I} (x_i - (1 - s_i))$ we see

$$0|_C = P = (x_I)|_C + (Q)|_C,$$

where the degree of Q is smaller than $|I|$. By induction, we can write Q as a combination of $x_{I'}$ for $I' \in \mathsf{str}(C)$. Since $(x_I)|_C = (-Q)|_C$ we get that x_I is in this span as well.

We now prove the second item. Consider a linear combination

$$P = \sum_{I \in \mathsf{sstr}(C)} \alpha_I x_I$$

such that not all α_I are zero. We want to show that there is $c \in C$ such that $P(c) \neq 0$. Let $Z \in \mathsf{sstr}(C)$ be a maximal set such that $\alpha_Z \neq 0$. Since Z is strongly shattered by C, there is some $\bar{s} : ([n] \setminus Z) \to \{0,1\}$ such that all extensions of it in $\{0,1\}^n$ are in C. Let $Q(x_i)_{i \in Z}$ be the polynomial obtained by plugging in the values of $\bar{s}$ in the variables of $([n] \setminus Z)$. By maximality of Z it follows that the coefficient of x_Z in Q is $\alpha_Z \neq 0$, and so Q is not the 0 polynomial. Therefore there is $s \in \{0,1\}^Z$ such such that $Q(s) \neq 0$. Pick $c \in C$ such that

$$c_i = \begin{cases} s_i & i \in Z, \\ \bar{s}_i & i \in ([n] \setminus Z). \end{cases}$$

It follows that $P(c) = Q(s) \neq 0$, which finishes the proof. ◀

We use this lemma to prove bounds on the Hilbert function and interpolation degree.

Proof of Theorem 7. For the upper bound on $h_d(C,\mathbb{F})$, the above proof shows how to express all monomials of degree d using monomials of equal or smaller degree. For the lower bound on $h_d(C,\mathbb{F})$, linear independence still holds after restricting set size.

The upper bound on $\mathsf{intdeg}(C,\mathbb{F})$ is immediate. For the lower bound on $\mathsf{intdeg}(C,\mathbb{F})$, since $\mathsf{sstr}(C)$ is downward-closed, the linear independence of the monomials in $\mathsf{sstr}(C)$ implies any maximal degree monomial in $\{(x_I)|_C : I \in \mathsf{sstr}(C)\}$ cannot be expressed solely by lower degree monomials. ◀

3.2 Down-shifts, Downward-closed Bases, and the Hilbert Function

We prove Theorem 8. We also use the theorem to analyze the Hilbert function for the parity function. Theorem 8 is a direct corollary of the following theorem.

▶ **Theorem 17.** *Let $C \subseteq \{0,1\}^n$ and let $D = S_n(S_{n-1}(\ldots S_1(C)))$. Then the set of monomials $\{\prod_{i\in I} x_i : I \in D\}$ is a basis for the vector space of functions $\{f : C \to \mathbb{F}\}$.*

A theorem, equivalent in content, but expressed with respect to Gröbner bases, is proved in [31]. For completeness we include an elementary proof in the full version of this paper.

The lower bound given in Theorem 8 subsumes the lower bound in Theorem 7. This is a direct corollary of the following simple lemma.

▶ **Lemma 18.** *Let $C \subseteq \{0,1\}^n$ and let $D = S_n(S_{n-1}(\ldots S_1(C)))$. We have that $\mathsf{sstr}(C) \subseteq D$, where we associate $\{0,1\}^n$ with subsets of $[n]$ in the natural way.*

Proof. Since D is downward-closed, it follows that it is shattering extremal and therefore $\mathsf{sstr}(D) = D$. So, it is enough to show that $\mathsf{sstr}(C) \subseteq \mathsf{sstr}(D)$. To this end, it suffices to show that for every class C', $\mathsf{sstr}(C') \subseteq \mathsf{sstr}(S_i(C'))$. Let $I \in \mathsf{sstr}(C')$. Therefore C' contains a subcube B in coordinates I. During the down-shift, B is either shifted or stays in place, but in any case also $S_i(C')$ contains a subcube in coordinates I and therefore $I \in \mathsf{sstr}(S_i(C'))$. ◀

The Hilbert Function of Parity. A simple example which demonstrates an application of Theorem 8 is the parity function. Let P denote the set of all vectors of even hamming weight. Notice that P does not contain subcubes other than $\emptyset$. Therefore, $\mathsf{sstr}(P) = \{\emptyset\}$. As a consequence, the lower bound on the Hilbert function in Theorem 7 reveals no information in this case. In contrast, shifting gives a better bound. If we down-shift P, say on the first coordinate, we get the set $S_1(P) = D = \{v : v_1 = 0\}$. Therefore, as D is downward closed, shifting it on other coordinates does not change it. Thus, $S_n(S_{n-1}(\ldots S_1(P))) = D$. By Theorem 8 we have that $h_d(P,\mathbb{F}) \geq \binom{n-1}{\leq d} = \binom{n-1}{d} + \binom{n-1}{d-1} + \ldots + \binom{n-1}{0}$.

This lower bound is tight when the field has characteristic two and $d \leq n/2$. It suffices to show every polynomial q of degree at most d can be expressed by a polynomial of degree at most d that does not depend on x_1. Therefore the $\binom{n-1}{\leq d}$ multilinear monomials that do not depend on x_1 span the space of degree at most d polynomials with domain P. Note that $(x_1 + \ldots + x_n)|_P = 0$, and therefore every appearance of x_1 can be replaced by $x_2 + \ldots + x_n$. This transforms q to a polynomial that does not depend on x_1 without changing the represented function.

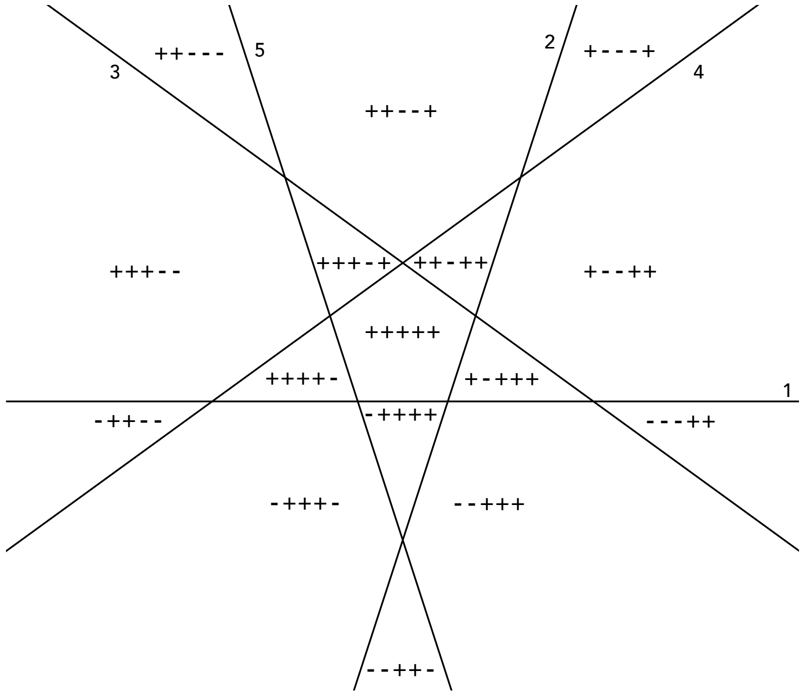

Figure 2 Five hyperplanes divide $\mathbb{R}^2$ into 16 cells. Cell labels in $\{-,+\}^5$ correspond to oriented hyperplane feasibility. Notice that every two coordinates are strongly shattered, but no three coordinates are shattered. This provides a proof-by-picture of Proposition 20 for $m = 5$ and $d = 2$.

4 Linear Programming and Low-degree Polynomial Approximations

We now prove Theorem 6. By the Razborov-Smolensky framework, it suffices to prove that $f_{\mathcal{H}}$ cannot be approximated by a low-degree polynomial over any field.[3]

▶ **Theorem 19.** *Let $\mathcal{H}$ be an arrangement of $2k+1$ hyperplanes in $\mathbb{R}^k$ that are in general position. For any any polynomial $p \in \mathbb{F}[x_1, \ldots, x_{2k+1}]$ we have*

$$\Pr_s[p(s) \neq f_{\mathcal{H}}(s)] \geq \frac{1}{2} - \frac{10 \deg_{\mathbb{F}}(p)}{\sqrt{2k+1}},$$

where s is uniform over $\{-1,1\}^{2k+1}$.

The proof of Theorem 19 proceeds via a reduction to Corollary 10. Let

$$S_{\mathcal{H}} = \{s \in \{-1,1\}^n : \ f_{\mathcal{H}}(s) = 1\}.$$

To apply Corollary 10 on $f_{\mathcal{H}}$ we will show $|S_{\mathcal{H}}| = 2^{2k}$ and $\mathsf{sstr}(S_{\mathcal{H}}) = \{I \subseteq [2k+1] : |I| \leq k\}$. We establish this by the following proposition. The facts we need about hyperplane arrangements follow from standard arguments [21, 43]. For intuition about the following proposition, see Figure 2 for a pictorial proof in $\mathbb{R}^2$.

▶ **Proposition 20.** *For any m hyperplanes $\mathcal{H}$ in $\mathbb{R}^d$ in general position*

$$\mathsf{sstr}(S_{\mathcal{H}}) = \mathsf{str}(S_{\mathcal{H}}) = \{I \subseteq [m] : |I| \leq d\}.$$

---

[3] We state the following theorem for $\{-1,1\}$ inputs to $f_{\mathcal{H}}$. This only makes sense for fields containing these elements. When $\mathbb{F} = \mathbb{F}_2$ simply replace $\{-1,1\}$ with $\{0,1\}$ in the definition of $f_{\mathcal{H}}$.

Proof. In the full version of the paper we include two lemmas that characterize the shattered and strongly shattered sets of $S_{\mathcal{H}}$ when $\mathcal{H}$ is in general position. The first lemma shows $\mathsf{str}(S_{\mathcal{H}}) \subseteq \{I \subseteq [m] : |I| \leq d\}$. The second lemma shows $\{I \subseteq [m] : |I| \leq d\} \subseteq \mathsf{sstr}(S_{\mathcal{H}})$. Since $\mathsf{sstr}(S_{\mathcal{H}}) \subseteq \mathsf{str}(S_{\mathcal{H}})$ these two lemmas combine to finish the proof. ◀

Proposition 20 implies Theorem 6. The equality $\mathsf{sstr}(S_{\mathcal{H}}) = \mathsf{str}(S_{\mathcal{H}})$ along with the Sandwich Theorem implies that $|S_{\mathcal{H}}| = |\mathsf{sstr}(S_{\mathcal{H}})|$. Let k be the ambient dimension in Theorem 6. The above proposition for $m = 2k+1$ and $d = k$ gives $|S_{\mathcal{H}}| = 2^{2k}$ and also $\mathsf{sstr}(S_{\mathcal{H}}) = \{I \subseteq [2k+1] : |Y| \leq k\}$. Thus $f_{\mathcal{H}}$ satisfies the premises of Corollary 10, and Theorem 6 follows.

5 Downward-closed Sets and Prescribed Intersections

We prove Theorem 12. Let $D \subseteq \{0,1\}^n$ be a downward-closed set. Fix $\phi : D \to D$ with the property that $\phi(a) \subseteq a$ for all $a \in D$. We will show that there exists a bijection $\pi : D \to D$ satisfying the condition that $a \cap \pi(a) = \phi(a)$ for all $a \in D$. We first prove two lemmas about the function space $\{f : D \to GF(2)\}$ and then use these to prove the existence of π. The first lemma holds for all subsets of the boolean cube.

▶ **Lemma 21.** *Let $C \subseteq \{0,1\}^n$ be a subset of the boolean hypercube. The monomials*

$$\prod_{i \in a} x_i \text{ for } a \in C$$

form a basis for $\{f : C \to GF(2)\}$.

Proof. We proceed using induction on $|C|$. When $C = \{a\}$ for $a \in \{0,1\}^n$ the function space has dimension one and the monomial $\prod_{i\in a} x_i$ represent the constant "1" function in this space, which spans it. Let $z \in C$ denote a maximal Hamming weight element in C. Notice $\prod_{i\in z} x_i$ is an indicator function in $\{f : C \to GF(2)\}$ for the input z. By the inductive hypothesis on $(C \setminus \{z\})$, we know the set of monomials $\prod_{i\in a} x_i$ for $a \in (C \setminus \{z\})$ form a basis for $\{f : (C \setminus \{z\}) \to GF(2)\}$. Since $\prod_{i\in z} x_i$ is an indicator function, we may add it to the basis for $\{f : (C \setminus \{z\}) \to GF(2)\}$ and achieve a basis for $\{f : C \to GF(2)\}$. ◀

We remark that if C is downward-closed, then it is shattering extremal, and the above lemma is a corollary of the Sandwich theorem. We prove the following stronger claim as well.

▶ **Lemma 22.** *Let $D \subseteq \{0,1\}^n$ be a downward-closed set. Fix any mapping $\phi : D \to D$ with the property that $\phi(a) \subseteq a$ for all $a \in D$. The functions*

$$\prod_{i \in \phi(a)} x_i \prod_{i \in a \setminus \phi(a)} (1 + x_i)$$

for $a \in D$ form a basis for $\{f : D \to GF(2)\}$.

Proof. Let $\mathcal{B}$ denote the set of polynomials that we wish to show is a basis. Since the cardinality of $\mathcal{B}$ is $|D|$ it is enough to show that it is a spanning set. By Lemma 21, it is enough to show that every monomial of the form $\prod_{i\in a} x_i$ for $a \in D$ can be expressed as a linear combination of polynomials in $\mathcal{B}$. We proceed by induction on the size of a. The case of $a = \emptyset$ is trivial. For the induction step, let $a \in D$ be non-empty. Expand the polynomial

$$\prod_{i \in \phi(a)} x_i \prod_{i \in a \setminus \phi(a)} (1 + x_i) = \left(\prod_{i \in a} x_i \right) + r,$$

where r is a linear combination of monomials $\prod_{i \in b} x_i$ for $b \subseteq a$ and $b \neq a$. Since D is downward-closed, by induction hypothesis r is in the span of $\mathcal{B}$. Thus,

$$\prod_{i \in a} x_i = \left(\prod_{i \in \phi(z)} x_i \prod_{i \in a \setminus \phi(a)} (1 + x_i) \right) + r$$

is also in the span of $\mathcal{B}$, and we are done. ◀

Proof of Theorem 12. We show there exists a bijection $\pi : D \to D$ such that $a \cap \pi(a) = \phi(a)$ for all $a \in D$, for the given map ϕ. Consider the $|D| \times |D|$ boolean matrix M defined as follows. Index the rows and columns both by D, and define the element in location $(a, b) \in D \times D$ to be one if and only if $a \cap b = \phi(a)$. We claim that M is nonsingular. Indeed, the rows of M correspond to the functions in Lemma 22. Since they form a basis, the row space of M is $|D|$-dimensional. This implies the determinant of M is nonzero. There must exist a permutation $\pi : [n] \to [n]$ such that $\prod_{i=1}^{|D|} M_{i,\pi(i)} = 1$. By the definition of M, we found the bijection π we were looking for. ◀

6 Conclusion

We exhibited a connection between algebra and combinatorics. We provided a general way to lower bound the Hilbert function. We showed a new family of functions cannot be approximated by low-degree polynomials. We provided a polynomial method proof of the Sandwich theorem and for a new theorem about prescribed intersections.

6.1 Open Directions

Our work suggests that the interpolation degree is a useful complexity measure on subsets of the boolean hypercube. Therefore, an open direction is to better understand the structure of sets with low interpolation degree. As noted by Remscrim [38], one can equivalently define interpolation degree in terms of the rank of a certain incidence matrix. The matrix corresponds to the monomials in our Lemma 21 with a cut-off on the degree. For the case of interpolation degree one, this characterization is particularly simple.

▶ **Proposition 23.** *A set $C \subseteq \{0,1\}^n$ has $\mathsf{intdeg}(C, \mathbb{F}) = 1$ if and only if the boolean vectors corresponding to C are affinely independent in $\mathbb{F}^n$.*

We are curious if other properties of the vectors in C correspond to implications for the interpolation degree. Even for interpolation degree two, the algebraic/matrix description becomes more opaque and less intuitive than the characterization in the above proposition. Since $\mathsf{intdeg}(C, \mathbb{F}) \leq \mathsf{VC}(V)$, any combinatorial characterization may also shed new light on the structure of sets with VC dimension two, for which our understanding is lacking [3, 34].

Acknowledgements. We thank Paul Beame, Eli Ben-Sasson, Yuval Filmus, Ariel Gabizon, Sivaramakrishnan Natarajan Ramamoorthy, Anup Rao, Mert Sağlam, and Amir Yehudayoff for helping us improve the presentation of the paper. We also thank the anonymous reviewers for their comments.

References

1 Noga Alon. Combinatorial Nullstellensatz. *Combinatorics Probability and Computing*, 8(1):7–30, 1999.
2 Noga Alon, Peter Frankl, and Vojtech Rödl. Geometrical Realization of Set Systems and Probabilistic communication Complexity. In *FOCS*, pages 277–280. IEEE, 1985.
3 Noga Alon, Shay Moran, and Amir Yehudayoff. Sign rank, VC dimension and spectral gaps. *Electronic Colloquium on Computational Complexity (ECCC)*, 21:135, 2014. URL: `http://eccc.hpi-web.de/report/2014/135`.
4 Ian Anderson. Combinatorics of Finite sets. *Bull. Amer. Math. Soc.*, 1988.
5 Richard P. Anstee, Lajos Rónyai, and Attila Sali. Shattering news. *Graphs and Combinatorics*, 18(1):59–73, 2002. `doi:10.1007/s003730200003`.
6 J. Aspnes, R. Beigel, M. Furst, and S. Rudich. The expressive power of voting polynomials. *Combinatorica*, 14(2):135–148, 1994. `doi:10.1007/BF01215346`.
7 L Babai and P Frankl. *Linear Algebra Methods in Combinatorics*. University of Chicago, 1992.
8 Hans-Jürgen Bandelt, Victor Chepoi, Andreas W. M. Dress, and Jack H. Koolen. Combinatorics of lopsided sets. *Eur. J. Comb.*, 27(5):669–689, 2006.
9 Richard Beigel. The Polynomial Method in Circuit Complexity. In *In Proceedings of the 8th IEEE Structure in Complexity Theory Conference*. Citeseer, 1993.
10 C Berge. A Theorem Related to the Chvátal conjecture. In *Proceedings, 5th British Combinatorial Conference*, 1976.
11 Anna Bernasconi and Lavinia Egidi. Hilbert Function and Complexity Lower Bounds for Symmetric Boolean Functions. *Information and Computation*, 153(1):1–25, 1999.
12 Abhishek Bhowmick and Shachar Lovett. Nonclassical Polynomials as a Barrier to Polynomial Lower Bounds. In *30th Conference on Computational Complexity*, page 72, 2015.
13 Béla Bollobás and Imre Leader. Sums in the grid. *Discrete Mathematics*, 162(1-3):31–48, 1996. `doi:10.1016/S0012-365X(96)00303-2`.
14 Béla Bollobás and A. J. Radcliffe. Defect sauer results. *J. Comb. Theory, Ser. A*, 72(2):189–208, 1995.
15 Béla Bollobás, A. J. Radcliffe, and Leader I. Reverse kleitman inequalities. *Proc. London Math. Soc., Ser. A*, (3) 58:153–168, 1989.
16 Vašek Chvátal and Donald Knuth. *Selected Combinatorial Research Problems*. Stanford University, 1972.
17 David P Dobkin and Steven P Reiss. The Complexity of Linear Programming. *Theoretical Computer Science*, 11:1–18, 1980.
18 A.W.M. Dress. Towards a theory of holistic clustering. *DIMACS Ser. Discrete Math. Theoret. Comput. Sci.*, 37 Amer. Math. Soc.:271–289, 1997.
19 Zeev Dvir. On the size of Kakeya sets in finite fields. *Journal of the American Mathematical Society*, 22(4):1093–1097, 2009.
20 Per Enflo. On the nonexistence of uniform homeomorphisms betweenl p-spaces. *Arkiv för Matematik*, 8(2):103–105, 1970. `doi:10.1007/BF02589549`.
21 Bernd Gärtner and Emo Welzl. Vapnik-Chervonenkis dimension and (pseudo-) hyperplane arrangements. *Discrete & Computational Geometry*, 12(1):399–432, 1994.
22 Oded Goldreich, Shafi Goldwasser, Eric Lehman, Dana Ron, and Alex Samorodnitsky. Testing monotonicity. *Combinatorica*, 20(3):301–337, 2000. `doi:10.1007/s004930070011`.
23 Ben Joseph Green and Terence Tao. Freiman's theorem in finite fields via extremal set theory. *Combinatorics, Probability & Computing*, 18(3):335–355, 2009. `doi:10.1017/S0963548309009821`.
24 Leonid Gurvits. Linear algebraic proofs of VC-dimension based inequalities. In *Computational Learning Theory*, pages 238–250. Springer, 1997.

25 Larry Guth and Nets Hawk Katz. On the Erdos Distinct Distances Problem in the Plane. *Annals of Mathematics*, 181(1):155–190, 2015.

26 Stasys Jukna. *Extremal Combinatorics: with Applications in Computer Science*. Springer Science & Business Media, 2011.

27 Daniel J. Kleitman. Families of non-disjoint subsets. *Journal of Combinatorial Theory*, 1(1):153–155, 1966. `doi:10.1016/S0021-9800(66)80012-1`.

28 László Kozma and Shay Moran. Shattering, graph orientations, and connectivity. *Electr. J. Comb.*, 20(3):P44, 2013. URL: `http://www.combinatorics.org/ojs/index.php/eljc/article/view/v20i3p44`.

29 J. Lawrence. Lopsided sets and orthant-intersection by convex sets. *Pac. J. Math.*, 104(1):155–173, 1983.

30 Jiří Matoušek. *Lectures on discrete geometry*, volume 212. Springer New York, 2002.

31 Tamás Mészáros. S-extremal set systems and Gröbner Bases. Master's thesis, Budapest University of Technology and Economics, 2009.

32 Tamás Mészáros and Lajos Rónyai. Shattering-extremal set systems of VC dimension at most 2. *Electr. J. Comb.*, 21(4):P4.30, 2014. URL: `http://www.combinatorics.org/ojs/index.php/eljc/article/view/v21i4p30`.

33 S. Moran. Shattering Extremal Systems. Master's thesis, Saarland University, Saarbrücken, Germany, 2012.

34 Shay Moran, Amir Shpilka, Avi Wigderson, and Amir Yehudayoff. Teaching and compressing for low vc-dimension. In *FOCS*, 2015. URL: `http://arxiv.org/abs/1502.06187`.

35 Shay Moran and Manfred K. Warmuth. Labeled compression schemes for extremal classes. *CoRR*, abs/1506.00165, 2015. URL: `http://arxiv.org/abs/1506.00165`.

36 A. Pajor. Sous-espaces l_1^n des espaces de banach. *Travaux en Cours. Hermann, Paris*, 1985.

37 Alexander A Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical Notes*, 41(4):333–338, 1987.

38 Zachary Remscrim. The Hilbert Function, Algebraic Extractors, and Recursive Fourier Sampling. *ECCC Technical Report TR 16-020*, 2016.

39 Micha Sharir, Adam Sheffer, and Joshua Zahl. Improved Bounds for Incidences Between Points and Circles. *Combinatorics, Probability and Computing*, 24(03):490–520, 2015.

40 R. Smolensky. On Representations by Low-degree Polynomials. In *FOCS*, 1993.

41 Roman Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the nineteenth annual ACM symposium on Theory of computing*, pages 77–82. ACM, 1987.

42 Roman Smolensky. Well-Known Bound for the VC-Dimension Made Easy. *Computational Complexity*, 6(4):299–300, 1997. `doi:10.1007/BF01270383`.

43 Richard P Stanley. Introduction to Hyperplane Arrangements. *Lecture notes, IAS/Park City Mathematics Institute*, 2004.

44 Terence Tao. Algebraic Combinatorial Geometry: the Polynomial Method in Arithmetic Combinatorics, Incidence Combinatorics, and Number Theory. *EMS Surveys in Mathematical Sciences*, 1(1):1–46, 2014.