

A Sauer–Shelah–Perles Lemma for Lattices

Stijn Cambie^{*} Bogdan Chornomaz[†] Zeev Dvir[‡] Yuval Filmus[§]

Shay Moran[¶]

December 31, 2019

Abstract

We study lattice-theoretical extensions of the celebrated Sauer–Shelah–Perles Lemma. We conjecture that a general Sauer–Shelah–Perles Lemma holds for a lattice L if and only if L is relatively complemented, and prove partial results towards this conjecture.

1 Introduction

Vapnik–Chervonenkis dimension [40, 41], or VC dimension for short, is a combinatorial parameter of major importance in discrete and computational geometry [12, 25, 27], statistical learning theory [9, 41], and other areas [3, 20, 23, 29]. The VC dimension of a family F of binary vectors, $F \subseteq \{0, 1\}^n$, is the largest cardinality of a set shattered by the family. In many of its applications, the power of this notion boils down to the Sauer–Shelah–Perles lemma [36, 37, 41], which states that the largest cardinality of a family on n points with VC dimension d is

$$\binom{n}{0} + \cdots + \binom{n}{d}, \quad (1)$$

a bound achieved by the family $\{S : |S| \leq d\}$. This lemma follows easily from the strengthening due to Pajor [35] and Aharoni and Holzman [1], which states that every family of binary vectors shatters at least as many sets as it has:

$$|F| \leq |\{X : X \text{ is shattered by } F\}| \quad (2)$$

This latter version will be of our primary concern and, for the sake of compliance with existing terminology, we will call it the SSP lemma.

VC dimension and corresponding lemmas have been extended to various settings, such as non-binary vectors [2, 24, 26, 39], integer vectors [42], Boolean matrices with forbidden configurations [4, 5], multivalued functions [24], continuous spaces [34], graph powers [11], and ordered variants [6]. In this paper, we study a generalization of the VC dimension to ranked lattices, and identify a rich class of lattices which satisfies the SSP lemma, including the lattice of subspaces of a finite vector space as well as all geometric lattices (flats of matroids). Moreover, we identify a necessary condition for a lattice to satisfy the SSP lemma: it must be free of an induced copy of the lattice $0 < 1 < 2$. We conjecture that this is the only obstruction; namely that a lattice satisfies the lemma if and only if it is *relatively complemented (RC)*, that is, does not contain an induced copy of $0 < 1 < 2$.

^{*}Department of Mathematics, Radboud University Nijmegen, Postbus 9010, 6500 GL Nijmegen, The Netherlands. Email: S.Cambie@math.ru.nl. Research supported by a Vidi Grant of the Netherlands Organization for Scientific Research (NWO), grant number 639.032.614.

[†]Department of Mathematics, Vanderbilt University. Email: bogdan.chornomaz@vanderbilt.edu.

[‡]Department of Computer Science and Department of Mathematics, Princeton University. Email: zeev.dvir@gmail.com. Research supported by NSF CAREER award DMS-1451191 and NSF grant CCF-1523816. Work on this paper was done while the author was a Von-Neumann fellow at the Institute for Advanced Study, supported by a Simons foundation fellowship.

[§]Computer Science Department, Technion — Israel Institute of Technology. Email: yuvalfi@cs.technion.ac.il. Taub Fellow — supported by the Taub Foundations. The research was funded by ISF grant 1337/16.

[¶]School of Mathematics, IAS. Email: shaymoran1@gmail.com. Research supported by the National Science Foundation under agreement No. CCF-1412958 and by the Simons Foundations.

VC dimension for ranked lattices. The definition of shattering for binary vectors is easier to formulate in terms of sets, where we identify an n -bit string with its set of 1's and vice versa. A family F of subsets of $\{1, \dots, n\}$ *shatters* a set S if for all $T \subseteq S$, the family F contains a set A with $A \cap S = T$. This definition readily generalizes to lattices (indeed, to meet-semilattices): a family F of elements in a lattice L *shatters* an element $s \in L$ if for all $t \leq s$, the family F contains an element a such that $a \wedge s = t$. If the lattice is ranked, it is natural to define the VC dimension of a family as the maximum rank of an element it shatters.

Given a ranked lattice L , the family F_d consisting of all elements of rank at most d , has VC dimension d . This family contains $\binom{L}{0} + \dots + \binom{L}{d}$ elements, where $\binom{L}{d}$ is the number of elements in L of rank d . The classical Sauer–Shelah–Perles lemma states that when L is the lattice of all subsets of $\{1, \dots, n\}$, the family F_d has maximum size among all families of VC dimension d .

The following generalization of Equation (1) appears in an unpublished (but well circulated) manuscript of Babai and Frankl [18].

Theorem 1.1 (Babai and Frankl [18]). *Let L be a ranked lattice of rank r with nonvanishing Möbius function, i.e. $\mu(x, y) \neq 0$ for all $x \leq y$. Then for all $0 \leq d \leq r$, any family $F \subseteq L$ of VC dimension d contains at most $\sum_{i=0}^d \binom{L}{i}$ elements. Furthermore, for every $d \leq r(L)$ the inequality is tight for some $F \subseteq L$ of VC dimension d .*

The condition of nonvanishing Möbius function is satisfied, for example, by a large class of lattices, namely geometric lattices, which are lattices whose elements are the flats of a finite matroid. Boolean lattices, which are the focus of the classical VC theory, also fall under this category. Another particularly compelling example of a geometric lattice is the lattice of all subspaces of \mathbb{F}_q^n , where \mathbb{F}_q is the finite field of order q .

Just as in the classical case, Theorem 1.1 follows from the next inequality corresponding to Equation (2):

Theorem 1.2 (Babai and Frankl [18]). *If L has nonvanishing Möbius function then every family $F \subseteq L$ shatters at least $|F|$ elements.*

For completeness, we provide a proof of Theorem 1.2 in Section 3.

1.1 Towards a Characterization of Sauer-Shelah-Perles Lattices

We will call a lattice satisfying the conclusion of Theorem 1.2 an SSP lattice. The main goal of this manuscript is to characterize such lattices. Note that the SSP property makes sense even for lattices which are not ranked. Having a nonvanishing Möbius function is sufficient for a lattice to be SSP, but this condition is not necessary. Two examples are given in Figure 1 and Figure 2. Both lattices are SSP, but the Möbius function vanishes on the interval consisting of the entire lattice. (The first example is simpler, but the lattice is not ranked.)

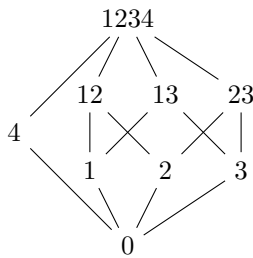


Figure 1: SSP lattice with vanishing Möbius function: $\mu(0, 1234) = 0$

On the other hand, we can identify a large class of lattices which are not SSP:

Theorem 1.3. *Let L be a lattice which is not relatively complemented, i.e., contains an induced copy of $0 < 1 < 2$. Then there exists a family $F \subseteq L$ shattering strictly fewer than $|F|$ elements.*

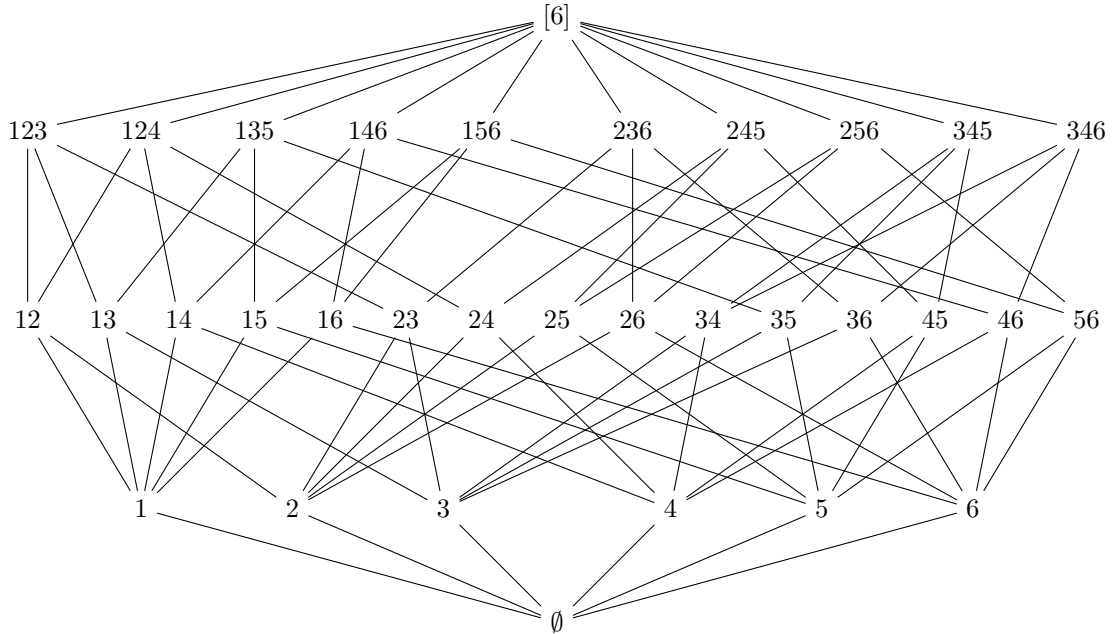


Figure 2: SSP lattice with vanishing Möbius function: $\mu(0, [6]) = 0$

The lattices in Figure 1 and Figure 2 are both relatively complemented. Indeed, we conjecture that induced copies of $0 < 1 < 2$ are the only obstructions for the SSP property:

Conjecture 1.4 ($SSP = RC$). A relatively complemented lattice is SSP.

Figure 3b on page 10 gives an example of a non-RC (and thus non-SSP) ranked lattice, which nevertheless satisfies the conclusion of Theorem 1.1. Characterization of such lattices is thus a separate problem, which we do not address in this paper.

As partial progress towards Conjecture 1.4, we prove it for lattices such that $\mu(x, y) \neq 0$ for all $x \leq y$ except possibly when x is the minimal element and y is the maximal element:

Theorem 1.5. *Let L be an RC lattice with minimal element 0 and maximal element e . If $\mu(x, y) \neq 0$ whenever $(x, y) \neq (0, e)$ then L is SSP.*

We also show that the SSP property is preserved under product:

Theorem 1.6. *If two lattices L and K are SSP, then so is $L \times K$.*

This theorem implies, via a structural result of Dilworth [16], that it suffices to prove Conjecture 1.4 for *simple* relatively complemented lattices (see Dilworth’s paper for a definition).

If we take a large power of any RC lattice satisfying the prerequisites of Theorem 1.5 (such as the ones in Figure 1 and Figure 2) then we get an SSP lattice whose Möbius function vanishes almost everywhere. This is a striking indication that the condition of nonvanishing Möbius function is far from being necessary for a lattice to be SSP.

We also verify the SSP property for specific families in RC lattices:

Theorem 1.7. *If L is an RC lattice and $F \subset L$ is a family for which $L \setminus \text{Str}(F)$ contains exactly one minimal element, then $|F| \leq |\text{Str}(F)|$.*

On the proof. All the results are stated for lattices, but they are true for meet-semilattices as well. Note that a meet-semilattice with a maximal element is a lattice, so Theorem 1.5 is still true in the meet-semilattice setting. The proofs of the other theorems also work in the meet-semilattice setting.

The original proofs of the Sauer–Shelah–Perles lemma used induction on n . Alon [2] and Frankl [17] gave an alternative proof using combinatorial shifting, and Frankl and Pach [19], Anstee [4], Gurvits [22],

Smolensky [38], and Moran and Rashtchian [33] gave other proofs using the polynomial method, which the presented proof also employs. The other proof methods — induction and shifting — seem to fail even for the particular case of subspace lattices.

2 Preliminaries

Posets. A *poset* is a partially ordered set. Unless mentioned otherwise, all posets we discuss are finite. We will use \leq to denote the partial order. An *antichain* is a collection of elements which are pairwise incomparable. An element x is *covered* by y , denoted $x \lessdot y$, if $x < y$ and no element z satisfies $x < z < y$. We can describe a poset using its *Hasse diagram*, which is a graph drawn on a plane, in which the vertices correspond to the elements, the edges to the covering relation, and lower elements are smaller.

A *meet-semilattice* is a poset in which any two elements x, y have an element $z \leq x, y$ such that $w \leq z$ whenever $w \leq x, y$. The element z is denoted $x \wedge y$, and is called the *meet* of x, y . The dual operation is the *join* $x \vee y$. A poset in which any two elements have both a meet and a join is known as a lattice. The meet of all elements in a meet-semilattice is called the *minimal element*, denoted by 0 . The join of all elements in a lattice (or join-semilattice) is called the *maximal element*, denoted by e . An atom of a lattice is an element x for which $0 < x$.

A meet-semilattice is *ranked* if each element x is associated with a non-negative integer rank $r(x)$, subject to the following two constraints (which completely specify the rank): $r(0) = 0$, and $r(y) = r(x) + 1$ if $x < y$. Not every meet-semilattice can be ranked. The rank of a meet-semilattice is the maximum rank of an element. We denote the number of elements of rank d in a poset L by $\binom{L}{d}$, and the number of elements of rank at most d by $\left[\begin{smallmatrix} L \\ \leq d \end{smallmatrix} \right]$.

The standard example of a lattice is the *Boolean lattice* of all subsets of $\{1, \dots, n\}$ ordered by inclusion. The meet of two elements is their intersection, and the join of two elements is their union. The rank of a subset is its cardinality.

The product $L \times K$ of two lattices L, K is a lattice whose elements are the elements of a Cartesian product of L and K , with the order relation $(\ell_1, k_1) \leq (\ell_2, k_2)$ iff $\ell_1 \leq \ell_2$ and $k_1 \leq k_2$.

Möbius function. The Möbius function of a finite poset is a function $\mu(x, y)$ defined for any two elements $x \leq y$ in the following way: $\mu(x, x) = 1$, and for $x < y$,

$$\mu(x, y) = - \sum_{z: x \leq z < y} \mu(x, z).$$

For example, on the Boolean lattice the Möbius function is $\mu(x, y) = (-1)^{|y \setminus x|}$, and on the integer lattice (the divisors of n ordered by divisibility) the Möbius function is $\mu(x, y) = \mu(y/x)$, where $\mu(\cdot)$ is the number-theoretic Möbius function.

The Möbius function is important due to the two *Möbius inversion formulas*:

Lemma 2.1 (Möbius inversion). *If f, g are two real-valued functions on a poset then*

$$f(x) = \sum_{y \geq x} g(y) \text{ for all } x \iff g(x) = \sum_{y \geq x} \mu(x, y) f(y) \text{ for all } x.$$

and

$$f(y) = \sum_{x \leq y} g(x) \text{ for all } y \iff g(y) = \sum_{x \leq y} \mu(x, y) f(x) \text{ for all } y.$$

We say that a poset has *nonvanishing Möbius function* if $\mu(x, y) \neq 0$ for all $x \leq y$ in the poset. For example, the Boolean lattice has nonvanishing Möbius function, and the integer lattice has nonvanishing Möbius function if and only if n is squarefree.

Relatively complemented lattices. In a lattice L which has minimal and maximal elements, the *complement* of an element y is an element z for which $z \wedge y = 0$ and $z \vee y = e$. A complemented lattice is a lattice with minimal and maximal elements in which every element has a complement. A lattice or meet-semilattice is called *relatively complemented* (RC) if every interval $[x, y] = \{z \mid x \leq z \leq y\}$, considered as a sublattice, is complemented. In particular, an RC lattice is complemented. We will mostly use the following equivalent characterization by Björner [8]:

Lemma 2.2 (Björner). *A finite lattice is RC if and only if it does not contain a 3-element interval, i.e. there are no two elements $x < y$ such that there is a unique z satisfying $x < z < y$.*

We will use another simple property of RC lattices due to Björner:

Lemma 2.3. *Finite RC lattices are atomic, that is, e is the join of all atoms. Equivalently, for any $x < e$ there is an atom a such that $a \not\leq x$.*

Deeper structural results on RC lattices can be found in [16].

Matroids and geometric lattices. A *matroid* over a finite set U is a finite non-empty collection of subsets of U called *independent sets*, satisfying the following two axioms: if a set is independent, then so are all its subsets; and if A, B are independent and $|A| > |B|$, then there exists an element $x \in A \setminus B$ such that $B \cup \{x\}$ is also independent.

The *rank* of a subset $S \subseteq U$ is the maximum cardinality of a subset of S which is independent. The rank of a matroid is the rank of U . A *flat* is a subset of U whose supersets all have higher rank.

Given a matroid, we can construct a poset whose elements are all flats of the matroid, ordered by inclusion. This poset forms a ranked lattice, and a lattice formed in this way is called a *geometric lattice*. The rank of an element in the lattice is the rank of the corresponding flat in the matroid. Weisner's theorem implies that geometric lattices have nonvanishing Möbius functions:

Theorem 2.4 (Weisner). *The Möbius function of a geometric lattice satisfies $(-1)^{r(y)-r(x)}\mu(x, y) > 0$ for all $x \leq y$.*

For a proof, see [21, Corollary 16.3].

The collection of all subsets of $\{1, \dots, n\}$ forms a matroid of rank n whose flats are all subsets of $\{1, \dots, n\}$. The corresponding geometric lattice is the Boolean lattice described above. A more interesting example of a matroid is the collection of all subsets of \mathbb{F}_q^n which are linearly independent, which forms a matroid of rank n whose flats are all subspaces of \mathbb{F}_q^n . The corresponding geometric lattice is called the *subspace lattice* of \mathbb{F}_q^n .

3 VC theory for lattices

3.1 Definitions

In order to develop VC theory for lattices we need to define two concepts: shattering and VC dimension. We start with the more basic concept, shattering:

Definition 3.1 (Shattering). Let L be a meet-semilattice. A set $F \subseteq L$ *shatters* an element $y \in L$ if for all $x \leq y$ there exists an element $z \in F$ such that $z \wedge y = x$.

We denote the set of all elements shattered by F by $\text{Str}(F)$.

We comment that the definition can be extended further to general posets: in this case, the condition $z \wedge y = x$ should be understood as follows: $z \wedge y$ exists, and equals x .

A basic property of shattering is that it is hereditary:

Lemma 3.2. *Let L be a meet-semilattice. If a set $F \subseteq L$ shatters an element $z \in L$ and $y \leq z$, then F also shatters y . In other words, $\text{Str}(F)$ is downward-closed.*

Proof. Let $x \leq y$. Since F shatters z and $x \leq z$, there exists an element $w \in F$ satisfying $w \wedge z = x$. Since $y \leq z$, the same element satisfies $w \wedge y = w \wedge (y \wedge z) = (w \wedge z) \wedge y = x \wedge y = x$. \square

Having defined shattering, the definition of VC dimension is obvious:

Definition 3.3. Let L be a ranked meet-semilattice. The *VC dimension* of a non-empty set $F \subseteq L$, denoted $\text{VC}(F)$, is the maximum rank of an element shattered by F .

These definitions specialize to the classical ones in the case of the Boolean lattice.

3.2 Proof of Theorem 1.2

Let \mathbb{F} be an arbitrary field of characteristic zero. We will prove Theorem 1.2 by giving a spanning set of size $|\text{Str}(F)|$ for the $|F|$ -dimensional vector space $\mathbb{F}[F]$ of \mathbb{F} -valued functions on F . Theorem 1.2 then follows, since the cardinality of any spanning set is at least the dimension. The spanning set we construct will consist of functions of the form given by the following definition:

Definition 3.4. For $x \in L$, the function $\chi_x: L \rightarrow \mathbb{F}$ is given by

$$\chi_x(y) = \mathbf{1}_{\{y \geq x\}},$$

that is, $\chi_x(y) = 1$ if $y \geq x$, and otherwise $\chi_x(y) = 0$.

For a set $G \subseteq L$,

$$X(G) = \{\chi_x : x \in G\}.$$

In the case of the Boolean lattice, we can think of the elements of the lattice as encoded by sets $S \subseteq \{1, \dots, n\}$ as well as by Boolean variables x_1, \dots, x_n . The reader can verify that

$$\chi_S = \prod_{i \in S} x_i.$$

Definition 3.4 extends this idea to general posets.

We will show that $\mathbb{F}[F]$ is spanned by $X(\text{Str}(F))$. The first step is showing that $X(L)$ is a basis for $\mathbb{F}[L]$, which for the Boolean lattice just states that every function on $\{0, 1\}^n$ can be expressed uniquely as a multilinear polynomial:

Lemma 3.5. *The set $X(L)$ is a basis for $\mathbb{F}[L]$.*

Proof. Since $|X(L)| = |L| = \dim \mathbb{F}[L]$, it suffices to show that $X(L)$ is linearly independent. Consider any linear dependency of the form $\ell := \sum_x c_x \chi_x = 0$. We will show that $c_x = 0$ for all $x \in L$, and so $X(L)$ is linearly independent.

Arrange the elements of L in an order x_1, \dots, x_N such that $x_i < x_j$ implies $i < j$. We prove that $c_{x_i} = 0$ by induction on i . Suppose that $c_{x_j} = 0$ for all $j < i$. Then in particular, $c_{x_j} = 0$ for all $x_j < x_i$, and therefore

$$0 = \ell(x_i) = \sum_j c_{x_j} \chi_{x_j}(x_i) = \sum_{j: x_j \leq x_i} c_{x_j} = c_{x_i}. \quad \square$$

The crucial step of the proof of Theorem 1.2 is an application of (generalized) inclusion-exclusion, which shows that if F does not shatter z then $\chi_z|_F$ can be expressed as a linear combination of $\chi_w|_F$ for $w < z$. In the case of the Boolean lattice, the argument is as follows. Suppose that F does not shatter S , say $A \cap S \neq T$ for all $A \in F$. Then all elements of F satisfy

$$\prod_{i \in T} x_i \prod_{j \in S \setminus T} (1 - x_j) = 0,$$

which implies that over F ,

$$\prod_{i \in S} x_i = \sum_{R \subsetneq S \setminus T} (-1)^{|S \setminus (T \cup R)|+1} \prod_{j \in T \cup R} x_j.$$

The argument for general posets is very similar, and uses Möbius inversion:

Lemma 3.6. *Suppose that $z \notin \text{Str}(F)$. There exist coefficients γ_y such that for all $p \in F$,*

$$\chi_z(p) = \sum_{y < z} \gamma_y \chi_y(p).$$

Proof. For an element $p \in F$, define the following two functions:

$$f_p(x) = \mathbf{1}_{\{x \leq p \wedge z\}}, \quad g_p(y) = \mathbf{1}_{\{y = p \wedge z\}}.$$

Clearly $f_p(x) = \sum_{y \geq x} g_p(y)$, and so Lemma 2.1 shows that $g_p(x) = \sum_{y \geq x} \mu(x, y) f_p(y)$. Since $f_p(y) = 0$ unless $y \leq z$, we can restrict the sum to the range $x \leq y \leq z$. When $y \leq z$, the condition $y \leq p \wedge z$ is equivalent to the condition $y \leq p$, and so we conclude that

$$g_p(x) = \sum_{x \leq y \leq z} \mu(x, y) f_p(y) = \sum_{x \leq y \leq z} \mu(x, y) \chi_y(p).$$

Since $z \notin \text{Str}(F)$, there exists an element $x \leq z$ such that $p \wedge z \neq x$ for all $p \in F$. In other words, $g_p(x) = 0$ for all $p \in F$. Therefore all $p \in F$ satisfy

$$\chi_z(p) = - \sum_{x \leq y < z} \frac{\mu(x, y)}{\mu(x, z)} \chi_y(p),$$

using the nonvanishing of the Möbius function. \square

We can now complete the proof, employing exactly the same argument used for the Boolean lattice.

Theorem 1.2. If L is a lattice in which $\mu(x, y) \neq 0$ for all $x \leq y$ then every family $F \subseteq L$ shatters at least $|F|$ elements.

Proof. Lemma 3.5 shows that $X(L)$ is a basis for $\mathbb{F}[L]$, and so the functions χ_x , restricted to the domain F , span $\mathbb{F}[F]$. We will show that every function in $\mathbb{F}[F]$ can be expressed as a linear combination of functions in $X(\text{Str}(F))$.

Consider any function $f \in \mathbb{F}[F]$. Since $X(L)$ spans $\mathbb{F}[f]$, there exist coefficients c_x such that $f = \sum_x c_x \chi_x$. Define the potential function

$$\Phi(\vec{c}) = \sum_{\substack{x \notin \text{Str}(F) \\ c_x \neq 0}} N^{r(x)},$$

where $N = |L| + 1$, and choose a representation which minimizes $\Phi(\vec{c})$. If $\Phi(\vec{c}) > 0$ then choose $z \notin \text{Str}(F)$ satisfying $c_z \neq 0$ of maximal rank. Lemma 3.6 shows that

$$f = \sum_{x \neq z} c_x \chi_x + \sum_{y < z} \gamma_y c_z \chi_y.$$

The corresponding coefficient vector \vec{d} satisfies $\Phi(\vec{d}) < \Phi(\vec{c})$, contradicting the choice of \vec{c} . We conclude that $\Phi(\vec{c}) = 0$, and so f is a linear combination of functions in $X(\text{Str}(F))$.

Concluding, we have shown that $X(\text{Str}(F))$ spans $\mathbb{F}[F]$. Hence $|\text{Str}(F)| = |X(\text{Str}(F))| \geq \dim \mathbb{F}[F] = |F|$. \square

3.3 Some corollaries

Theorem 1.2 immediately implies Theorem 1.1.

Theorem 1.1. Let L be a ranked lattice of rank r with nonvanishing Möbius function, i.e. $\mu(x, y) \neq 0$ for all $x \leq y$. Then for all $0 \leq d \leq r$, any family $F \subseteq L$ of VC dimension d contains at most $\left[\begin{smallmatrix} L \\ \leq \text{VC}(F) \end{smallmatrix} \right] = \left[\begin{smallmatrix} L \\ 0 \end{smallmatrix} \right] + \cdots + \left[\begin{smallmatrix} L \\ d \end{smallmatrix} \right]$ elements. Furthermore, for every $d \leq r(L)$ the inequality is tight for some $F \subseteq L$ of VC dimension d .

Proof. Suppose that $\text{VC}(F) = d$. If $|F| > \left[\begin{smallmatrix} L \\ \leq d \end{smallmatrix} \right]$ then, according to Theorem 1.2, also $|\text{Str}(F)| > \left[\begin{smallmatrix} L \\ \leq d \end{smallmatrix} \right]$. However, this implies that $\text{Str}(F)$ must contain a set of rank larger than d , contradicting the assumption $\text{VC}(F) = d$. This proves the inequality.

To show that the inequality is tight for all $d \leq r(L)$, consider the set $F_d = \{x : r(x) \leq d\}$. This is a set containing $\left[\begin{smallmatrix} L \\ \leq d \end{smallmatrix} \right]$ elements which shatters all elements of rank d but no element of rank $d + 1$, and so satisfies $\text{VC}(F_d) = d$. \square

We can generalize Theorem 1.1 to arbitrary antichains to obtain a further corollary.

Corollary 3.7. *Let L be a ranked lattice with nonvanishing Möbius function and let $A \subseteq L$ be a maximal antichain. If $F \subseteq L$ does not shatter any element of A then*

$$|F| \leq |F_A|, \text{ where } F_A = \{x \in L : x < y \text{ for some } y \in A\}.$$

Furthermore, F_A does not shatter any element of A .

Theorem 1.1 is the special case of Corollary 3.7 in which A consists of all elements of rank $\text{VC}(F) + 1$.

Proof of Corollary 3.7. Let us start by showing that $|F| \leq |F_A|$. If $|F| > |F_A|$ then, according to Theorem 1.2, also $|\text{Str}(F)| > |F_A|$. Therefore F shatters some element x such that $x \not< y$ for all $y \in A$. Since A is a maximal antichain, either $x \in A$ or $x \geq y$ for some $y \in A$. In both cases F shatters some element in A (in the second case, according to Lemma 3.2).

Next, let us show that F_A does not shatter any element of A . Suppose that F_A shatters some element $a \in A$. Then some $x \in F_A$ satisfies $x \wedge a = a$, that is, $x \geq a$. Since $x \in F_A$, we know that $x < y$ for some $y \in A$. Put together, this implies that $a < y$, contradicting the fact that A is an antichain. \square

A final corollary is a *dichotomy theorem*, a direct consequence of the Sauer–Shelah–Perles lemma which is the source of many of its applications. Before describing our generalized dichotomy theorem, let us briefly describe the classical one. Let $F \subseteq \{0, 1\}^X$, where X is infinite. For every finite $I \subseteq X$, we can consider the projection of F to the coordinates of I , denoted $F|_I$. The *growth function* of F is

$$\Pi_F(n) = \max_{\substack{I \subseteq X \\ |I|=n}} |F|_I|.$$

The Sauer–Shelah–Perles lemma immediately implies the following polynomial versus exponential dichotomy for the growth function:

- Either $\text{VC}(F) = \infty$, in which case $\Pi_F(n) = 2^n$;
- or $\text{VC}(F) = d < \infty$, in which case $\Pi_F(n) \leq \binom{n}{\leq d} \leq 2n^d$.

For example, it implies that there is no F for which $\pi_F(n) = \Theta(2^{\log^2 n})$.

We can extend this result to vector spaces (we leave extensions to more general domains to the reader). Let \mathbb{F}_q be a finite field, let X be an infinite set, let \mathcal{V} denote the linear space of all functions $v: X \rightarrow \mathbb{F}_q$ with a finite support (i.e. $v(x) = 0$ for all but finitely many $x \in X$), and let \mathcal{L} denote the (infinite) lattice of all finite dimensional subspaces of \mathcal{V} . Let $F \subseteq \mathcal{L}$ be a family of subspaces. For every $I \in \mathcal{L}$, we can consider the projection $F|_I = \{V \cap I : V \in F\}$. The growth function of F is defined as in the classical case, with dimension replacing cardinality:

$$\Pi_F(n) = \max_{\substack{I \in \mathcal{L} \\ \dim(I)=n}} |F|_I|.$$

Theorem 1.1 immediately implies a dichotomy as in the classical case. In order to understand the resulting orders of growth, we need to be able to estimate $\begin{bmatrix} L \\ d \end{bmatrix}$ for subspace lattices L .

Lemma 3.8. *Let $L = \mathbb{F}_q^n$. For all $d \leq n$,*

$$q^{d(n-d)} \leq \begin{bmatrix} L \\ \leq d \end{bmatrix} \leq 2n^d q^{dn}.$$

In particular, $|L| \geq q^{(n^2-1)/4}$.

Proof. The number of elements of L of rank d is the q -binomial coefficient $\begin{bmatrix} n \\ d \end{bmatrix}_q$. There are many formulas for $\begin{bmatrix} n \\ d \end{bmatrix}_q$. The one we use is

$$\begin{bmatrix} n \\ d \end{bmatrix}_q = \sum_{\substack{A \subseteq \{1, \dots, n\} \\ |A|=d}} q^{\sum_{i \in A} i - d(d+1)/2}.$$

Calculation shows that the summand with highest exponent, corresponding to $A = \{n - d + 1, \dots, n\}$, has exponent $d(n - d)$. Therefore

$$q^{d(n-d)} \leq \binom{n}{d}_q \leq \binom{n}{d} q^{d(n-d)} \leq n^d q^{dn}.$$

This implies that

$$\left[\begin{array}{c} L \\ \leq d \end{array} \right] \leq \sum_{e=0}^d n^e q^{en} \leq n^d q^{dn} \sum_{e=0}^d (nq^n)^{-e}.$$

We can assume that $n \geq 1$, and so $nq^n \geq 2$, implying that $\sum_{e=0}^{\infty} (nq^n)^{-e} \leq 2$. This proves the main inequalities. The lower bound on $|L|$ follows from taking $d = \lfloor n/2 \rfloor$. \square

Combining the lemma with Theorem 1.1 specialized to the subspace lattice, we immediately obtain the following dichotomy theorem:

Theorem 3.9. *For every family $F \subseteq \mathcal{L}$, exactly one of the following holds:*

- *Either $\text{VC}(F) = \infty$, in which case $\Pi_F(n) \geq q^{(n^2-1)/4}$;*
- *or $\text{VC}(F) = d < \infty$, in which case $\Pi_F(n) \leq 2n^d q^{dn}$.*

4 Partial results towards the SSP=RC conjecture

4.1 Proof of Theorem 1.3

We start by showing that an SSP lattice must be RC.

Theorem 1.3. Let L be a lattice which is not relatively complemented. Then there exists a family $F \subseteq L$ shattering strictly fewer than $|F|$ elements.

Proof. Since L is not relatively complemented, there exist two elements x, y in L such that there is a unique element z satisfying $x < z < y$. Let $F = \{w \mid w \leq y\} \setminus \{x\}$. It is easy to check that $\text{Str}(F) \subseteq \{w \mid w \leq y\}$. On the other hand, y, z are not shattered since $x \notin F$. Therefore $|\text{Str}(F)| \leq |F| - 1$. \square

Figure 3 shows two examples of ranked lattices that are not relatively complemented, only one of the two satisfying the conclusion of Theorem 1.1.

The first example L_1 is the path lattice $0 < 1 < 2$, appearing in Figure 3a. The family $F = \{1, 2\}$ shatters only 0 and hence has VC dimension 0, while $|F| > \binom{L_1}{0}$.

The second example L_2 , appearing in Figure 3b, is more complex. One can check that $|\text{Str}(F)| \geq |F|$ unless $F = L_2 \setminus 4, L_2 \setminus 5$. Both of these families shatter 12, and so have VC dimension 2. Since $\binom{L_2}{\leq 2} = |L_2| - 1$, the lattice satisfies the conclusion of Theorem 1.1.

The lattice L_2 thus separates the SSP condition from the weaker condition given by Theorem 1.1. It seems hard to characterize the lattices that are “weakly SSP”, that is, satisfy $|\text{Str}(F)| \leq \binom{L}{\leq \text{VC}(F)}$ for all $F \subseteq L$.

4.2 Proof of Theorem 1.5

In this subsection we prove Theorem 1.5, which concerns RC lattices for which $\mu(x, y) \neq 0$ holds whenever $(x, y) \neq (0, e)$ (recall that 0 is the minimal element and e is the maximal element). Such lattices do exist, for example the two lattices given in the introduction (Figure 1 and Figure 2).

The crucial tool for the proof will be the following analog of Lemma 3.6 for this case.

Lemma 4.1. *Let $F \subseteq L$ be a family different from L and $L \setminus 0$. Let $z \in L$ be such that $z \notin \text{Str}(F)$. There exist coefficients λ_y such that for all $p \in F$ we have*

$$\chi_z(p) = \sum_{y < z} \lambda_y \chi_y(p).$$

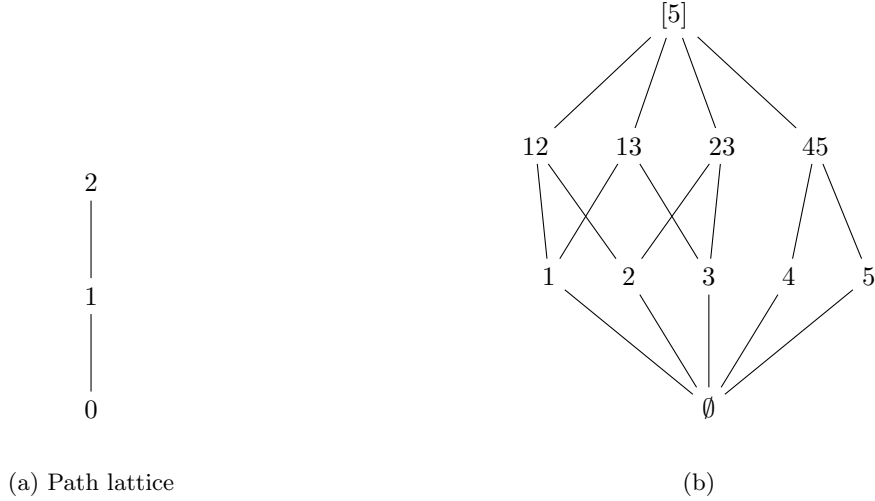


Figure 3: Lattices that are not RC

Proof. If μ is nonvanishing or $z \neq e$ then $\mu(x, z) \neq 0$ for all $x \leq z$, and so the original proof of Lemma 3.6 still applies. It therefore suffices to handle the case $z = e$ and $\mu(0, e) = 0$.

For every $p \in L$, define a function v_p on $L \setminus e$ by $v_p(y) = \chi_y(p)$ (in other words, if we think of the functions χ_y for $y < e$ as the rows of a matrix, then the functions v_p are its columns); so $v_p(y)$ indicates the condition “ $y \leq p$ ”. We will show that the space of linear dependencies of $\{v_p : p \in L\}$ is one-dimensional. More explicitly, we show that if $\sum_{a \in L} c_a v_a = 0$ then $c_a = \mu(a, e)c_e$ for all $a \in L$.

The proof is by backwards induction. The base case, $a = e$, is trivial. Now suppose that $c_b = \mu(b, e)c_e$ for all $b > a$. Then

$$0 = \sum_{b \in L} c_b v_b(a) = c_a + \sum_{b > a} \mu(b, e)c_e = c_a - \mu(a, e)c_e,$$

by definition of the Möbius function.

As $F \neq L, L \setminus 0$, there must be some element $a \neq 0$ missing from F . We claim that the functions $\{v_p : p \in F\}$ are linearly independent. Indeed, any linear dependency $\sum_{p \in F} c_p v_p = 0$ lifts to a linear dependency of $\{v_p : p \in L\}$, with $c_p = 0$ for $p \notin F$. In this linear dependency, $c_a = 0$. Since $a \neq 0$, we have $\mu(a, e) \neq 0$, and so $c_e = c_a / \mu(a, e) = 0$, which implies that the linear dependency is trivial.

Since the functions $\{v_p : p \in F\}$ are linearly independent, the $(L \setminus 0) \times F$ matrix whose columns are v_p has full rank $|F|$, and so its rows (which are just the functions $\chi_y|_F$ for $y < e$) span $\mathbb{F}[F]$. In particular, some linear combination of the rows equals $\chi_e|_F$. \square

Theorem 1.5 follows by a straightforward adaptation of the proof of Theorem 1.2.

Theorem 1.5. Let L be an RC lattice with minimal element 0 and maximal element e . If $\mu(x, y) \neq 0$ whenever $(x, y) \neq (0, e)$ then every family $F \subseteq L$ shatters at least $|F|$ elements.

Proof. When F is different from L and $L \setminus 0$, the conclusion holds exactly as in the proof of Theorem 1.2 in Subsection 3.2, by using Lemma 4.1 instead of Lemma 3.6. When $F = L$, we obviously have $|F| \leq |\text{Str}(F)|$ as $F = \text{Str}(F)$. We argue that when $F = L \setminus 0$, then $\text{Str}(F) = L \setminus e$. Indeed, take $y \in L \setminus e$. Any $s \leq y$ different from 0 belongs to F , and $y \wedge s = s$. By Lemma 2.3, there is an atom $a \in F$ such that $a \not\leq y$. Thus $y \wedge a = 0$, and so y is shattered. \square

4.3 Proof of Theorem 1.6

In this subsection we prove that the SSP property is closed under taking products.

Theorem 1.6. If two lattices L and K are SSP, then $L \times K$ is SSP.

Proof. Let $F \subseteq K \times L$. We want to show that $|\text{Str}(F)| \geq |F|$. For $k \in K$ and $\ell \in L$, define $F_k, I_k \subseteq L$ and $G_\ell, J_\ell \subseteq K$ as follows:

$$\begin{aligned}
F_k &= \{\ell \in L \mid (k, \ell) \in F\}; \\
I_k &= \text{Str}_L(F_k); \\
G_\ell &= \{k \in K \mid \ell \in I_k\}; \\
J_\ell &= \text{Str}_K(G_\ell).
\end{aligned}$$

Construction of these sets is illustrated in Figure 4 below.

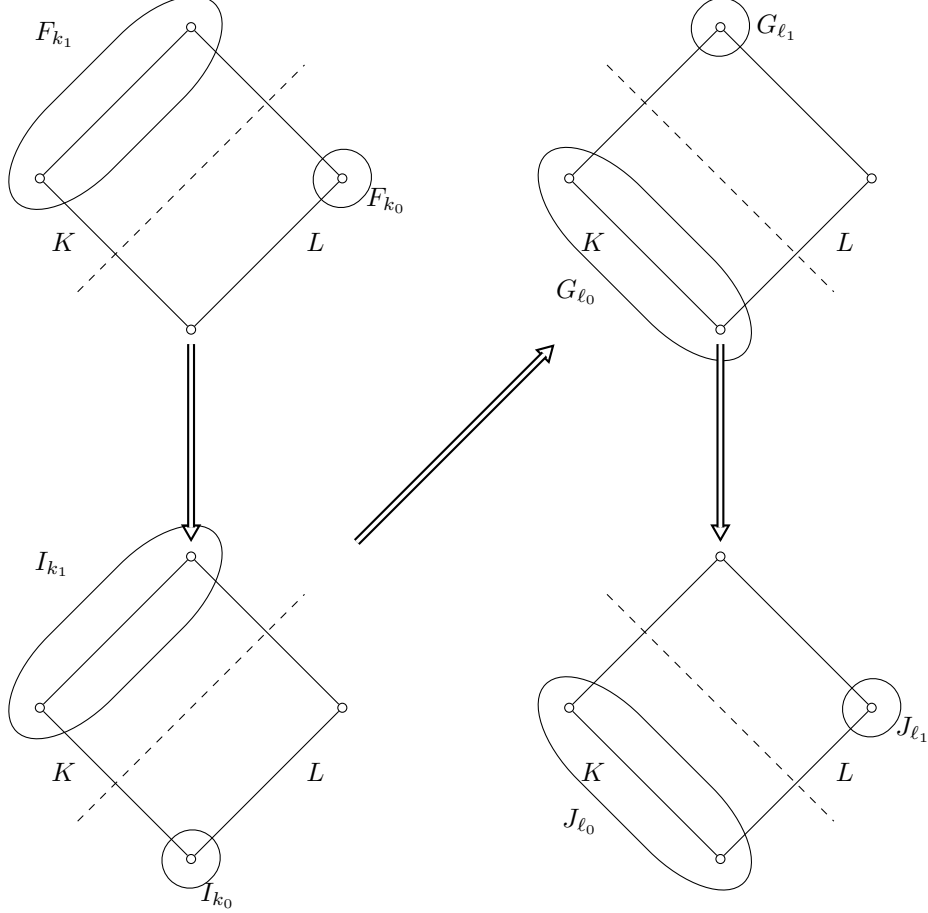


Figure 4: Construction of F_k , I_k , G_ℓ and J_ℓ .

Now, by SSP of K and L , $|I_k| \geq |F_k|$ and $|J_\ell| \geq |G_\ell|$, for all $k \in K$, $\ell \in L$. Also, by definition, $\bigsqcup_{k \in K} \{k\} \times I_k = \bigsqcup_{\ell \in L} G_\ell \times \{\ell\}$. Let us define $J = \bigsqcup_{\ell \in L} J_\ell \times \{\ell\}$. Then

$$|J| = \sum_{\ell \in L} |J_\ell| \geq \sum_{\ell \in L} |G_\ell| = \sum_{k \in K} |I_k| \geq \sum_{k \in K} |F_k| = |F|.$$

To conclude the proof, we show that F shatters all elements of J .

Take any $(k, \ell) \in J$, and let $k' \leq k$ and $\ell' \leq \ell$. By construction, $k \in J_\ell$, and so k is shattered by G_ℓ . In particular, some $u \in G_\ell$ satisfies $k \wedge_K u = k'$. Since $u \in G_\ell$, by construction $\ell \in I_u$, that is, ℓ is shattered by F_u . In particular, some $v \in F_u$ satisfies $\ell \wedge_L v = \ell'$. In total, $(u, v) \in F$ satisfies $(k, \ell) \wedge (u, v) = (k', \ell')$. \square

4.4 Proof of Theorem 1.7

We close by identifying a class of families which satisfies the SSP property in any RC lattice.

Theorem 1.7. Let L be an RC lattice and $F \subset L$ be a family for which $L \setminus \text{Str}(F)$ contains exactly one minimal element, then $|F| \leq |\text{Str}(F)|$.

Proof. Denote the set of elements not shattered by F by $N = L \setminus \text{Str}(F)$. The set N is closed upwards: if F does not shatter u then it also does not shatter any $v \geq u$. Therefore, if x is the unique minimal element of N , then $N = [x] = \{u \mid x \leq u\}$.

Since x is not shattered by F , there exists some $y \leq x$ such that no $u \in F$ satisfies $u \wedge x = y$. Denote $D = \{u \mid x \wedge u = y\}$, so D and F are disjoint. We will show that $|N| \leq |D|$. This implies the lemma since

$$|F| = |L| - |L \setminus F| \leq |L| - |D| \leq |L| - |N| = |L| - |L \setminus \text{Str}(F)| = |\text{Str}(F)|.$$

Now, to prove $|N| \leq |D|$, let us take an arbitrary $a \in N = [x]$, that is, an arbitrary a satisfying $x \leq a$. Then $y \leq x \leq a$, but, as L is RC, the interval $[y, a]$ is complemented, and we can pick a complement $c(a)$ of x in $[y, a]$. Note that this applies, in particular, when $y = x = a$, in which case $c(a) = a$.

By definition, $c(a) \wedge x = y$, that is, $c(a) \in D$. Also, $c(a) \vee x = a$. This implies that the map $a \mapsto c(a)$ is one-to-one on N . Indeed, if $c(a) = c(b)$ then $a = x \vee c(a) = x \vee c(b) = b$. The existence of a one-to-one mapping from N to D proves that $|N| \leq |D|$, finishing the argument. \square

5 Conclusion and open problems

VC dimension and shattering were extensively studied in the classical case. Theorem 1.2 and, if true, Conjecture 1.4, enable us to ask related questions in an extended setting of lattices with nonvanishing Möbius function, or of RC lattices. We end the paper by outlining a list of possible questions, which we consider to be interesting.

Families with small VC dimension and inclusion-maximality. An explicit description of all sets of VC dimension 1 exists in the classical case: they correspond to forests [7, 30]. Having an exhaustive description of these families in the extended setup would be desirable. A subset of an SSP-lattice of VC dimension k is called *inclusion-maximal* if adding an element to it increases its VC dimension. Naturally, study of families with small VC dimension boils down to study of inclusion-maximal families of small VC dimension.

Question 5.1. Give a description of inclusion-maximal families of VC dimension 1 for SSP lattices.

It is known that, in a classical setup, every set of VC dimension 1 can be extended to a set of size $n + 1$ without increasing its VC dimension. The same does not hold in the case of vector spaces, as the following example shows:

Lemma 5.2. Consider a subspace lattice $L = \mathbb{F}_q^n$, where $n \geq 2$. Let U be a subspace of \mathbb{F}_q^n of dimension $n - 1$. The set $F = \{0, \mathbb{F}_q^n\} \cup \{x : x \notin U\}$ is an inclusion-maximal set of VC dimension 1, and it contains $q^{n-1} + 2$ subspaces; in comparison, $\left[\begin{smallmatrix} L \\ \leq 1 \end{smallmatrix} \right] = 1 + \frac{q^n - 1}{q - 1}$, which is larger by a factor of roughly $\frac{q}{q-1}$.

Question 5.3. Is there a “nice” characterization of lattices for which all inclusion-maximal sets of VC dimension 1 have at least $n + 1$ elements? Here n is the number of join-irreducible elements, which in the classical setup corresponds to the size of the base set.

Shattering-extremality. A subset of an SSP-lattice is called *shattering-extremal* if it shatters as many elements as it has, that is, if it obtains equality in SSP inequality. A number of results exists for shattering-extremal families in the classical setup, in particular, for small VC dimension [30–32].

Question 5.4. What can we say about shattering-extremal sets of small VC dimension for SSP lattices?

In the classical setting there is a notion of *strong shattering*, which is in some sense dual to shattering. The analog of the SSP lemma for strong shattering states that a family strongly shatters at most as many elements as it has. Thus the size of the family is sandwiched between the number of elements it strongly shatters and the number of elements it shatters [10]. It turns out that equality holds in the SSP lemma iff it holds in the strong SSP lemma.

Question 5.5. Can strong shattering be reasonably defined in the lattice setting? Will shattering-extremality be equivalent to strong shattering-extremality, as in the classical setting?

Shattering-extremal families which are closed under intersection are precisely *convex geometries* [14].

Question 5.6. Characterize the class of shattering-extremal meet-subsemilattices of SSP lattices. Matroids and their duals, being SSP, are in this class, as are duals of antimatroids, which are convex geometries.

Matroids and antimatroids are known examples of *greedoids* [28]. Are shattering-extremal families precisely duals of greedoids, or is there some other connection?

It is well-known that for antimatroids there is a characterization in terms of *forbidden projections*, also called *circuits* [15]. This characterization can be recast, in a straightforward manner, for convex geometries. A similar characterization also exists for shattering-extremal families in general [13].

Question 5.7. Is there a “forbidden projections” characterization of shattering-extremal families of SSP lattices?

References

- [1] Ron Aharoni and Ron Holzman. Unpublished.
- [2] Noga Alon. On the density of sets of vectors. *Discrete Math.*, 46(2):199–202, 1983.
- [3] Noga Alon, Shay Moran, and Amir Yehudayoff. Sign rank versus Vapnik–Chervonenkis dimension. *Sbornik: Mathematics*, 208(12):1724, 2017.
- [4] R. P. Anstee. General forbidden configuration theorems. *J. Combin. Theory Ser. A*, 40(1):108–124, 1985.
- [5] R. P. Anstee and Balin Fleming. Two refinements of the bound of Sauer, Perles and Shelah, and of Vapnik and Chervonenkis. *Discrete Math.*, 310(23):3318–3323, 2010.
- [6] R. P. Anstee, Lajos Rónyai, and Attila Sali. Shattering news. *Graphs Combin.*, 18(1):59–73, 2002.
- [7] Shai Ben-David. 2 notes on classes with Vapnik–Chervonenkis dimension 1. *ArXiv e-prints*, July 2015.
- [8] Anders Björner. On complements in lattices of finite length. *Discrete Math.*, 36(3):325–326, 1981.
- [9] Anselm Blumer, Andrzej Ehrenfeucht, David Haussler, and Manfred K. Warmuth. Learnability and the Vapnik–Chervonenkis dimension. *J. Assoc. Comput. Mach.*, 36(4):929–965, 1989.
- [10] Béla Bollobás and Andrew J Radcliffe. Defect sauer results. *Journal of Combinatorial Theory, Series A*, 72(2):189–208, 1995.
- [11] Nicolò Cesa-Bianchi and David Haussler. A graph-theoretic generalization of the Sauer-Shelah lemma. *Discrete Appl. Math.*, 86(1):27–35, 1998.
- [12] Bernard Chazelle and Emo Welzl. Quasi-optimal range searching in spaces of finite VC-dimension. *Discrete Comput. Geom.*, 4(5):467–489, 1989.
- [13] Bogdan Chornomaz. Unpublished.
- [14] Bogdan Chornomaz. Convex geometries are extremal for the generalized sauer-shelah bound. *The Electronic Journal of Combinatorics*, 25(2):2–35, 2018.
- [15] Brenda L Dietrich. A circuit set characterization of antimatroids. *Journal of Combinatorial Theory, Series B*, 43(3):314–321, 1987.
- [16] R. P. Dilworth. The structure of relatively complemented lattices. *Ann. of Math. (2)*, 51:348–359, 1950.
- [17] Peter Frankl. On the trace of finite sets. *J. Combin. Theory Ser. A*, 34(1):41–45, 1983.

- [18] Peter Frankl and László Babai. *Linear algebra methods in combinatorics*. Preliminary version 2 edition, 1992.
- [19] Peter Frankl and János Pach. On the number of sets in a null t -design. *European J. Combin.*, 4(1):21–23, 1983.
- [20] Zoltán Füredi and János Pach. Traces of finite sets: extremal problems and geometric applications. In *Extremal problems for finite sets (Visegrád, 1991)*, volume 3 of *Bolyai Soc. Math. Stud.*, pages 251–282. János Bolyai Math. Soc., Budapest, 1994.
- [21] Chris Godsil. An introduction to the Moebius function. *ArXiv e-prints*, March 2018.
- [22] Leonid Gurvits. Linear algebraic proofs of VC-dimension based inequalities. In *Computational Learning Theory*, pages 238–250. Springer, 1997.
- [23] Hamed Hatami and Yingjie Qian. Teaching dimension, VC dimension, and critical sets in Latin squares. *J. Comb.*, 9(1):9–20, 2018.
- [24] David Haussler and Philip M. Long. A generalization of Sauer’s lemma. *J. Combin. Theory Ser. A*, 71(2):219–240, 1995.
- [25] David Haussler and Emo Welzl. ϵ -nets and simplex range queries. *Discrete Comput. Geom.*, 2(2):127–151, 1987.
- [26] M. G. Karpovsky and V. D. Milman. Coordinate density of sets of vectors. *Discrete Math.*, 24(2):177–184, 1978.
- [27] János Komlós, János Pach, and Gerhard Woeginger. Almost tight bounds for ϵ -nets. *Discrete Comput. Geom.*, 7(2):163–173, 1992.
- [28] Bernhard Korte, László Lovász, and Rainer Schrader. *Greedoids*, volume 4. Springer Science & Business Media, 2012.
- [29] Ilan Kremer, Noam Nisan, and Dana Ron. On randomized one-round communication complexity. *Comput. Complexity*, 8(1):21–49, 1999.
- [30] Tamás Mészáros and Lajos Rónyai. Shattering-extremal set systems of small VC dimension. *ISRN Comb.*, 2013.
- [31] Tamás Mészáros and Lajos Rónyai. Shattering-extremal set systems of VC dimension at most 2. *Electron. J. Combin.*, 21(4):Paper 4.30, 17, 2014.
- [32] Shay Moran. Shattering extremal sets. Master’s thesis, Saarland University, Saarbrücken, Germany, 2012.
- [33] Shay Moran and Cyrus Rashtchian. Shattered sets and the hilbert function. In *MFCS*, volume 58 of *LIPICs*, pages 70:1–70:14. Schloss Dagstuhl — Leibniz-Zentrum fuer Informatik, 2016.
- [34] B. K. Natarajan. On learning sets and functions. *Mach. Learn.*, 4(1):67–97, Oct 1989.
- [35] Alain Pajor. *Sous-espaces l_1^n des espaces de Banach*, volume 16 of *Travaux en Cours [Works in Progress]*. Hermann, Paris, 1985. With an introduction by Gilles Pisier.
- [36] Norbert Sauer. On the density of families of sets. *J. Combinatorial Theory Ser. A*, 13:145–147, 1972.
- [37] Saharon Shelah. A combinatorial problem; stability and order for models and theories in infinitary languages. *Pacific J. Math.*, 41:247–261, 1972.
- [38] Roman Smolensky. Well-known bound for the vc-dimension made easy. *Computational Complexity*, 6(4):299–300, 1997.
- [39] J. Michael Steele. Existence of submatrices with all possible columns. *J. Combin. Theory Ser. A*, 24(1):84–88, 1978.

- [40] V. N. Vapnik and A. Ja. Červonenkis. The uniform convergence of frequencies of the appearance of events to their probabilities. *Dokl. Akad. Nauk SSSR*, 181:781–783, 1968.
- [41] V. N. Vapnik and A. Ya. Chervonenkis. The uniform convergence of frequencies of the appearance of events to their probabilities. *Theor. Probability Appl.*, 16:264280, 1971.
- [42] Roman Vershynin. Integer cells in convex sets. *Adv. Math.*, 197(1):248–273, 2005.