DIPLOMA THESIS

S-extremal set systems and Gröbner bases

Tamás Mészáros

Advisor: **Lajos Rónyai**
professor
Institute of Mathematics
Department of Algebra

BME
2010.

Contents

1 Introduction

We say that a set system $\mathcal{F} \subseteq 2^{[n]}$ shatters a given set $S \subseteq [n]$ if $2^S = \{F \cap S : F \in \mathcal{F}\}$. In general, a set system $\mathcal{F}$ shatters at least $|\mathcal{F}|$ sets. This inequality was proved by various authors (Aharoni and Holzman [10], Pajor [11], Sauer [12], Shelah [13]), and studied by many others. We concentrate on the case of equality. A set system is called S-extremal if it shatters exactly $|\mathcal{F}|$ sets. Our aim is to characterize these combinatorial objects and to study their properties. In contrast to earlier studies, which used mostly combinatorial methods, we take a different approach and develop algebraic methods based on the algebraic interpretation of shatteringm, introduced by Anstee, Rónyai and Sali in [1]. When considering a set system as a set of characteristic vectors, one can define the corresponding vanishing ideal of polynomials. We study the standard monomials and Gröbner bases of these ideals. Of course, later the precise definitions will be given, for the present we say that a Gröbner basis of an ideal of polynomials is a special generating system, possessing good properties, and the standard monomials of an ideal form a basis - also possessing good properties - of the corresponding quotient structure. They both will be useful to characterize S-extremal set systems.

After the introduction, as a preparation, we first present general results concerning the topic. In Section 2 we investigate the notion of standard monomials and Gröbner bases. We present some well known results, all of them can be found in [7], [9] and in Hungarian in [8]. In Section 3 we study the special case of vanishing ideals. We present the lex game, introduced by Felszeghy, Ráth and Rónyai, which gives a good description of the standard monomials of vanishing ideals. We investigate some results from [2], among others a fast, linear time, algorithm for computing the lexicographic standard monomials.

In the second part of our study we deal with the special case of ideals vanishing on $0-1$ vectors, i.e. with the case of set systems. In Section 4 we introduce the already mentioned notion of shattering and extramality. First we present our preliminary results concerning shattered sets and one of our main results, a characterization of extremal set systems using Gröbner

bases. After this we give an efficient algorithm that determines whether a set system $\mathcal{F} \subseteq 2^{[n]}$ is extremal or not. The running time of the algorithm is $O(n^2|\mathcal{F}|)$, which is an improvement over the previous bound of Greco in [6]. In Section 5 we deal with different set system operations and their relation to extremality. The most common operation, the downshift operation, is discussed in Section 6. Downshifting compresses the set system toward zero in a specific coordinate. It was already studied by various authors, among others by Bollobás and Radcliffe in [3]. After presenting their results we prove several propositions necessary for our own results concerning downshifts and standard monomials.

There is a straightforward graph theoretical interpretation of the topic. Our study in Section 7 is mainly based on the the work of Greco in [6]. We discuss the properties of isometrically and strongly isometrically embedded set systems, give a novel proof for one statement proposed by Greco and develop own results as well. The main result of the section uses notions from both, [3] and [6]. In Section 8 we make some remarks on the Vapnik-Chervonenkis dimension, to present a widely known and used notion in mathematics which is in close connection with the notion of shattering. We give a new algorithm using standard monomials for computing the Vapnik-Chervonenkis dimension of a set system with the same time bound like in [15].

The last part of our study deals with the generalization of our results to the case when the vectors are not necessarily binary. We generalize the notion of shattering like Shinohara in [22] and the downshift operation like it was introduced by Bollobás, Leader and Radcliffe in [5]. Several of our results can be generalized, among others one of our main results concerning the standard monomials of a vanishing ideal.

2 Standard monomials and Gröbner bases

Before getting started with the main definitions, we introduce some notations. Throughout the study $\mathbb{F}$ will stand for an ordinary field, and n will be a positive integer. The set $\{1, 2, \ldots, n\}$ will be referred to shortly as $[n]$, the power set of it as $2^{[n]}$, the family of subsets of size k as $[n]_k$ and for the ring of polynomials in n variables over $\mathbb{F}$ we will use the usual notation $\mathbb{F}[x_1, \ldots, x_n]$.

A monomial is a polynomial in $\mathbb{F}[x_1, \ldots, x_n]$ of the form $x_1^{w_1} x_2^{w_2} \ldots x_n^{w_n}$. The ring of polynomials $\mathbb{F}[x_1, \ldots, x_n]$ can be considered as a vector space over $\mathbb{F}$ with a natural basis that consists of the monomials.

For vectors we use boldface letters, and we denote their coordinates by the same letter indexed by respective numbers, e.g. $\mathbf{w} = (w_1, \ldots, w_n)$. Similarly, we will note by $f(\mathbf{x})$ the polynomial $f(x_1, x_2, \ldots, x_n) \in \mathbb{F}[x_1, \ldots, x_n]$ and by $\mathbf{x}^{\mathbf{w}}$ the monomial $x_1^{w_1} x_2^{w_2} \ldots x_n^{w_n}$.

2.1 Term orders

Definition 2.1 *The relation $\prec$ is a term order on the monomials if it is a linear order with 1 as minimal element and it is monotone with respect to multiplication.*

We will always talk about standard monomials with respect to a fixed $\prec$ term order. For different term orders the standard monomials may differ. This fact will be essential in some of our future results.

An example of a term order is the lexicographic ordering of monomials. We say that $\mathbf{x}^w$ is smaller than or equal to $\mathbf{x}^u$ according to the lexicographic order if for the first index i such that $w_i \neq u_i$, we have $w_i < u_i$. This is clearly a term order.

For example for $n = 2$ the lexicographic ordering of the first some monomials is the following

$$1 \prec x_2 \prec x_2^2 \prec x_2^3 \cdots \prec x_1 \prec x_1 x_2 \prec x_1 x_2^2 \prec \ldots x_1^2 \prec x_1^2 x_2 \prec x_1^2 x_2^2 \prec \ldots$$

By reordering the variables, we can get another lexicographic order, so we will talk about a lexicographic term order based on some permutation of the variables $x_1, x_2, \ldots, x_n$.

Generally we can say that term orders are in close connection with the divisibility of monomials:

Proposition 2.1 *(Dickson's lemma, see [7, p.21].) Every $\prec$ term order is the refinement of the divisibility between monomials and is a well-ordering.*

Proof: For the first part, let us suppose that $\mathbf{x}^u|\mathbf{x}^v$. Then $\frac{\mathbf{x}^v}{\mathbf{x}^u}$ is a monomial as well, so $1 \preceq \frac{\mathbf{x}^v}{\mathbf{x}^u}$. And then when multiplying by $\mathbf{x}^u$ we get the desired inequality.
For the proof of the second part see [7, p.21]. ■

2.2 Standard and leading monomials

Let us fix a $\prec$ term order. The leading monomial of a nonzero polynomial $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is the greatest monomial according to $\prec$ appearing in $f(\mathbf{x})$ with nonzero coefficient, and is denoted by $lm(f(\mathbf{x}))$. For an ideal of polynomials I we denote by $Lm(I)$ the set of leading monomials of the polynomials in I:

$$Lm(I) = \{lm(f(\mathbf{x})) : f(\mathbf{x}) \in I, f \neq 0\}.$$

A monomial which is not a leading monomial of any polynomial in I is called a standard monomial. The set of standard monomials is denoted by $Sm(I)$:

$$Sm(I) = \{\mathbf{x}^\mathbf{w} \in \mathbb{F}[\mathbf{x}]\} \backslash Lm(I) = \{\mathbf{x}^\mathbf{w} : \nexists f(\mathbf{x}) \in I, \text{ for which } lm(f) = \mathbf{x}^\mathbf{w}\}.$$

The standard monomials of an ideal will be henceforth of great importance. Now we will discuss some general and well known properties of standard monomials.

Definition 2.2 *A set $S \subseteq \{\mathbf{x}^\mathbf{w} \in \mathbb{F}[\mathbf{x}]\}$ is downward (upward) closed with respect to divisibility or shortly a down-set (up-set) if $\mathbf{x}^\mathbf{w} \in S$ and $\mathbf{x}^\mathbf{u}|\mathbf{x}^\mathbf{w}$ ($\mathbf{x}^\mathbf{w}|\mathbf{x}^\mathbf{u}$) imply that $\mathbf{x}^\mathbf{u} \in S$.*

Proposition 2.2 *$Sm(I)$ is a down-set and $Lm(I)$ is an up-set.*

Proof: If $\mathbf{x}^v|\mathbf{x}^u$ and $\mathbf{x}^v \in Lm(I)$, then there exists a polynomial $p(\mathbf{x})$ in I with $\mathbf{x}^v$ as its leading monomial. Since I is an ideal, and $\frac{\mathbf{x}^u}{\mathbf{x}^v}$ is a monomial, $q(\mathbf{x}) = \frac{\mathbf{x}^u}{\mathbf{x}^v}p(\mathbf{x}) \in I$. However, because of the properties of a term order, the leading monomial of $q(\mathbf{x})$ is $\mathbf{x}^u$, and so $\mathbf{x}^u \in Lm(I)$. The other part of the statement follows from the fact that the complementary of an up-set is a down-set. ■

Proposition 2.3 *The canonical image of $Sm(I)$ is a basis of $\mathbb{F}(\mathbf{x})/I$ as an $\mathbb{F}$ vector space.*

Proof: Clearly there are no two elements of $Sm(I)$ belonging to the same coset, otherwise there would be $\mathbf{x}^{u_1}, \mathbf{x}^{u_2} \in Sm(I)$ for which $f(\mathbf{x}) = \mathbf{x}^{u_1} - \mathbf{x}^{u_2} \in I$, but none of these two monomials is a leading monomial of a polynomial in I. Similarly, we can see that the cosets represented by the elements of $Sm(I)$ are linearly independent.
Now let us take an arbitrary coset from the quotient ring represented by $f(\mathbf{x})$. There are two possibilities. If $lm(f(\mathbf{x}))$ is a standard monomial, then we can continue with $f(\mathbf{x}) - b_1 lm(f(\mathbf{x}))$, where b_1 is the coefficient of $lm(f(\mathbf{x}))$. If this is not the case, then there exists a polynomial $g_1(\mathbf{x}) \in I$ such that $lm(g_1(\mathbf{x})) = lm(f(\mathbf{x}))$. Now we can continue with $f_1(\mathbf{x}) = f(\mathbf{x}) - \frac{b_1}{c_1}g(\mathbf{x})$, where c_1 is the coefficient of $lm(f(\mathbf{x}))$ in $g_1(\mathbf{x})$. For this we have $lm(f_1(\mathbf{x})) \prec lm(f(\mathbf{x})$. Since there is no infinite downward chain of monomials starting with $lm(f(\mathbf{x}))$ according to the $\prec$ term order, this process terminates in finitely many steps with $f(\mathbf{x}) = s(\mathbf{x}) + g(\mathbf{x})$, where $s(\mathbf{x})$ contains just standard monomials and $g(\mathbf{x}) \in I$. So the coset represented by $f(\mathbf{x})$ is the sum of the cosets represented by the monomials in $s(\mathbf{x})$ in the quotient ring, that is, $Sm(I)$ generates $\mathbb{F}(\mathbf{x})/I$ as an $\mathbb{F}$ vector space. Together with linear independence, this means that $Sm(I)$ is a basis of it as well. ■

2.3 Gröbner bases

Definition 2.3 *Let I be an ideal of $\mathbb{F}[\mathbf{x}]$. For a fixed term order, a finite subset $G \subseteq I$ is a Gröbner basis of I if for every $f \in I$ there exists a $g \in G$*

such that $lm(g)$ *divides* $lm(f)$.

Gröbner bases are of great importance not only in connection with extremal set systems. They were introduced in 1965 by Austrian mathematician Bruno Buchberger in his Ph.D. thesis. He was motivated by questions from commutative algebra and algebraic geometry, but since then Gröbner bases have been applied in various fields of mathematics e.g. code theory, symbolic computation, automatic theorem proving, integer programming, statistics, partial differential equations and numerical computations. A good survey is provided by [7], [9] or in Hungarian by [8].

Now we discuss some useful facts about Gröbner bases.

Proposition 2.4 *A Gröbner basis of an ideal is a generating system of it as well.*

Proof: Let G be a Gröbner basis of the ideal I, and $0 \neq f(\mathbf{x}) \in I$ an arbitrary element. Since G is a Gröbner basis, there exists a polynomial $g_1(\mathbf{x}) \in G$ such that $lm(g_1(\mathbf{x}))|lm(f(\mathbf{x}))$. With $r_1(\mathbf{x}) = \frac{lm(f(\mathbf{x}))}{lm(g_1(\mathbf{x}))}$ and $f_1(\mathbf{x}) = f(\mathbf{x}) - b_1 r_1(\mathbf{x}) g_1(\mathbf{x})$, where b_1 is the coefficient of $lm(f(\mathbf{x}))$ in $f(\mathbf{x})$, we have $lm(f_1(\mathbf{x})) \prec lm(f(\mathbf{x}))$, and we can continue this with $f_1(\mathbf{x})$. Since there cannot be an infinite downward chain of monomials starting with $lm(f(\mathbf{x}))$ according to the $\prec$ term order, this process terminates in finitely many steps giving an expression

$$f(\mathbf{x}) = b_1 r_1(\mathbf{x}) g_1(\mathbf{x}) + b_2 r_2(\mathbf{x}) g_2(\mathbf{x}) + \cdots + b_m r_m(\mathbf{x}) g_m(\mathbf{x}),$$

where $g_i(\mathbf{x}) \in G$. So G is indeed a generating system of I. ■

The question arises, whether every nonzero ideal has a Gröbner basis. The answer is fortunately yes, but for this we need the notion of reduction.

Reduction

Let $f, g \in \mathbb{F}[\mathbf{x}]$, and suppose that there is one monomial $\mathbf{x}^{\mathbf{w}}$ in f with nonzero coefficient c_f that is divisible by $lm(g)$. Let the coefficient of $lm(g)$ in g be c_g and let

$$\widehat{f}(\mathbf{x}) = f(\mathbf{x}) - \frac{c_f \cdot \mathbf{x}^{\mathbf{w}}}{c_g \cdot lm(g)} g(\mathbf{x}).$$

Since the leading monomial of $\frac{\mathbf{x}^{\mathbf{w}}}{lm(g)} g(\mathbf{x})$ is $\mathbf{x}^{\mathbf{w}}$, in $\widehat{f}$ it is replaced by a monomial strictly less then $\mathbf{x}^{\mathbf{w}}$. This operation is called reduction.

If G is a finite set of polynomials and $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ is an arbitrary polynomial, we say that f is reduced with respect to G if there is no monomial in f with nonzero coefficient that is divisible by $lm(g)$ for some $g \in G$. Now take an arbitrary f and reduce it with the elements of G, every time replacing the greatest monomial with smaller ones, until we get a reduced polynomial with respect to G. Since there is no infinite downward chain of monomials starting with $lm(f)$, this process terminates in finitely many steps ending up with a factorization

$$f(\mathbf{x}) = \sum_{i=1}^{m} g_i(\mathbf{x}) h_i(\mathbf{x}) + \widehat{f}(\mathbf{x}),$$

where $G = \{g_1, \ldots, g_m\}$, $h_1, \ldots, h_m \in \mathbb{F}[\mathbf{x}]$, $\widehat{f}$ is reduced to G and $lm(g_i h_i) \preceq lm(f)$ for every index i. We say that $\widehat{f}$ is the reduced version of f if such polynomials $h_1, h_2, \ldots, h_m$ exist.

Example 2.1 *Let* $g_1(x_1, x_2) = x_1^2 x_2^2 + x_1$, $g_2(x_1, x_2) = x_1^2 x_2^2 + x_2$, $G = \{g_1, g_2\}$, $f(x_1, x_2) = x_1^2 x_2^2$ *and* $\prec$ *the standard lexicographic order. If we reduce* f *with* g_1 *we get* $-x_1$*, if with* g_2 *we get* $-x_2$*. It is easy to see that both,* $-x_1$ *and* $-x_2$ *are reduced with respect to* G*.*

This example shows that the reduced version of a polynomial $f(\mathbf{x})$ is not necessary unique with respect to a fixed set of polynomials G. However, if G is a Gröbner basis of the ideal I, this cannot happen. As a corollary of Proposition 2.3 one can prove the following:

Proposition 2.5 *If* G *is a Gröbner basis of the ideal* I*, then the reduced version of a polynomial* $f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}]$ *with respect to* G *is unique, and* $f \in I$ *if and only if the reduced version is* 0*. Specially* $\langle G \rangle = I$ *holds.*

Proof: See [7, p.32-34]. ■

Moreover, this statement can be reversed:

Proposition 2.6 *If $G \subseteq I$ is finite, and every polynomial $f \in I$ can be reduced to 0 by G, then G is a Gröbner basis of I.*
If $G \subseteq I$ is finite, and every polynomial $f \in \mathbb{F}[\mathbf{x}]$ has a unique reduced version with respect to G, then G is a Gröbner basis of I.

Proof: See [7, p.32-34]. ■

As a consequence of these propositions, one can conclude the following, very important statement:

Proposition 2.7 *Every nonzero ideal I has a Gröbner basis.*

Proof: See [7, p.34]. ■

Obviously the Gröbner basis of a nonzero ideal I is not unique. For example by adding finitely many polynomials from I to G, the resulting set of polynomials will be henceforward a Gröbner basis of I. For uniqueness we need some more notions.

Reduced Gröbner basis

Definition 2.4 *If G is a Gröbner basis of some nonzero ideal I, and every polynomial $g \in G$ has leading coefficient 1 and is reduced with respect to $G \backslash \{g\}$, then G is called a reduced Gröbner basis.*

Reformulating this, we get that a Gröbner basis G is reduced if and only if every polynomial $g \in G$ apart from $lm(g)$ consists only of standard monomials and has 1 as leading coefficient.

Proposition 2.8 *Every nonzero ideal I has a unique reduced Gröner basis with respect to a fixed term order.*

Proof: See [7, p.48]. ■

3 Standard monomials of vanishing ideals

In this section we introduce a special ideal of polynomials, and study its properties.

Let $V \subseteq \mathbb{F}^n$ be a finite set of vectors, and denote by $I(V)$ the set of polynomials vanishing on V, i.e.:

$$I(V) = \{f(\mathbf{x}) \in \mathbb{F}[\mathbf{x}] : f(\mathbf{y}) = 0 \ for\ all\ \mathbf{y} \in V\}.$$

It is easy to see that $I(V)$ is an ideal in $\mathbb{F}[\mathbf{x}]$.

Proposition 3.1 $dim_{\mathbb{F}}\mathbb{F}(\mathbf{x})/I(V) = |V|$

Proof: Since V is finite, by interpolation we get that $\mathbb{F}(\mathbf{x})/I(V)$ is isomorphic to the space of functions from V to $\mathbb{F}$. But the space of these functions has dimension $|V|$. ■

This, together with Proposition 2.3 imply that $|Sm(I(V))| = |V|$. This equality will be essential in some of our further results.

From now on we deal just with this case, thus with vanishing ideals.

3.1 Lex game

Now we present the lex game, introduced by Felszeghy, Ráth and Rónyai in [2]. This twosome game can be used to calculate the standard monomials of the vanishing ideal $I(V)$ for a fixed lexicographic term order.

Without loss of generality we can suppose that $\prec$ is the standard lexicographic term order, i.e. $x_1 \succ x_2 \succ \cdots \succ x_n$.

The game is defined for a fixed, nonempty, finite set $V \subseteq \mathbb{F}^n$ and a vector $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{N}^n$, and for these fixed parameters is denoted by $Lex(V; \mathbf{w})$.

At first, one of the players, Stan, thinks of an element $\mathbf{y} = (y_1, \ldots, y_n)$ of V. The task of the other player, Lea, is to find out one coordinate of $\mathbf{y}$. The game goes on as follows. First Lea tries to find out y_n by guessing at most w_n times. If she succeeds, she wins, if not, then y_n is revealed to her by Stan.

In the next turn Lea continues by guessing for y_{n-1} at most w_{n-1} times, etc.. The game ends if Lea finds out y_i for some index i, and so she wins, or if Stan reveals y_1 (in this case Stan is the winner). Both of the players know the parameters V and $\mathbf{w}$.

It is useful to extend the game for the case $V = \emptyset$. Later we will see that the reasonable choice in this case is to define Lea as the winner of the lex game $Lex(\emptyset; \mathbf{w})$ for all $\mathbf{w} \in \mathbb{N}^n$.

Now, for better understanding, we introduce some useful notations. For fixed $\beta \in \mathbb{F}$ let V_β be the elements of V ending in β, i.e.

$$V_\beta = \{(v_1, \ldots, v_{n-1}) \in \mathbb{F}^{n-1} : (v_1, \ldots, v_{n-1}, \beta) \in V\}.$$

It is clear that if in the lex game $Lex(V; (w_1, \ldots, w_{n-1}, w_n))$ Lea did not find out y_n, then the game continues just like if they were starting a lex game $Lex(V_{y_n}; (w_1, \ldots, w_{n-1}))$. Generally for $\beta_i, \beta_{i+1}, \ldots, \beta_n \in \mathbb{F}$ let

$$V_{\beta_n, \beta_{n-1}, \ldots, \beta_i} = \{(v_1, \ldots, v_{i-1}) \in \mathbb{F}^{i-1} : (v_1, \ldots, v_{i-1}, \beta_i, \ldots, \beta_n) \in V\},$$

and if Lea did not find out none of $y_n, y_{n-1}, \ldots, y_i$, then they continue by playing a lex game $Lex(V_{y_n, y_{n-1}, \ldots, y_i}; (w_1, \ldots, w_{i-1}))$.

Let $\{\alpha_1, \ldots, \alpha_k\} \subseteq \mathbb{F}$ the set of all field elements which occur in any of the elements in V. (Since V is finite, such a finite set exists.) Clearly $V \subseteq \{\alpha_1, \ldots, \alpha_k\}^n$ and we can suppose that all guesses of Lea are from the set $\{\alpha_1, \ldots, \alpha_k\}$. The complementary of V in $\{\alpha_1, \ldots, \alpha_k\}^n$ will be denoted by V^c.

Winning strategies

Now we will study, what kind of winning strategies may Stan and Lea have. At first sight it is just Lea, who is playing, however, with a small modification, one can introduce a winning strategy for Stan as well. For this we allow him a bit of cheating. More precisely, we suppose that he actually does not think of a fixed $\mathbf{y} \in V$, he just answers NO for all guesses of Lea until there remains an element $\mathbf{y} \in V$, which is consistent with all of his answers. In this sense we can talk about a winning strategy for Stan as well.

Proposition 3.2 *If $n > 1$, then Stan has a winning strategy in the lex game $Lex(V; (w_1, \ldots, w_{n-1}, w_n))$ if and only if there exist at least $w_n + 1$ elements $\beta \in \{\alpha_1, \ldots, \alpha_k\}$ such that he has a winning strategy for the lex game $Lex(V_\beta; (w_1, \ldots, w_{n-1}))$. Similarly, for $n > 1$, Lea has no winning strategy for the lex game $Lex(V; (w_1, \ldots, w_{n-1}, w_n))$ if and only if there exist at least $w_n + 1$ elements $\beta \in \{\alpha_1, \ldots, \alpha_k\}$ such that she has no winning strategy for the lex game $Lex(V_\beta; (w_1, \ldots, w_{n-1}))$.*

Proof: See [2]. ■

One can prove that Stan has a winning strategy if and only if Lea does not have one. This means that for all parameters either Lea or Stan has a winning strategy, i.e. the parameters of the game determine who wins the game, as long as both players are playing the best possible. Thus instead of talking about winning strategies, we can talk about the winner of the lex game $Lex(V; \mathbf{w})$.

Who wins the lex game?

In the following we will establish a connection between the lex game and the lexicographic standard and leading monomials.

Theorem 3.1 *Let $V \subseteq \mathbb{F}^n$ a finite set of points and $\mathbf{w} \in \mathbb{N}^n$. Lea wins the lex game $Lex(V; \mathbf{w})$ if and only if $\mathbf{x}^\mathbf{w} \in Lm(I(V))$.*

This theorem implies immediately another:

Theorem 3.2 *Let $V \subseteq \mathbb{F}^n$ a finite set of points and $\mathbf{w} \in \mathbb{N}^n$. Stan wins the lex game $Lex(V; \mathbf{w})$ if and only if $\mathbf{x}^\mathbf{w} \in Sm(I(V))$.*

The proof of them can be found in [2].

Properties of the lexicographic standard monomials of vanishing ideals

Now we present some important combinatoric properties of the lexicographic standard monomials of vanishing ideals. All of them can be found in [2].

If we analyze the lex game, we can notice that it is independent from $\mathbb{F}$ and from the berth of V in it. It is depending only on the equality of points in some coordinates. From this we can conclude the following statement:

Proposition 3.3 *Let $\widehat{\mathbb{F}}$ be an arbitrary field, and suppose that we are given a set of injective functions $\varphi_j : \{\alpha_1, \ldots, \alpha_k\} \to \widehat{\mathbb{F}}$, $j = 1, 2, \ldots, n$. Let $\widehat{V}$ be the image of V, i.e.*

$$\widehat{V} = \{(\varphi_1(v_1), \ldots, \varphi_n(v_n)) : (v_1, \ldots, v_n) \in V\}.$$

Now the lex standard monomials of V are the same in $\mathbb{F}[\mathbf{x}]$ as those of $\widehat{V}$ in $\widehat{\mathbb{F}}[\mathbf{x}]$. ■

Specially if $V \subseteq \{0,1\}^n$, then $Sm(I(V))$ is the same for all fields $\mathbb{F}$. Proposition 3.3 also means that without loss of generality we can suppose that for example $\mathbb{F} = \mathbb{R}$ and $V \subseteq \{0, 1, \ldots, k-1\}$.

Theorems 3.1 and 3.2 together with Proposition 3.2 imply another very important property of lexicographic standard monomials of vanishing ideals. We will refer to it as the recursive property.

Proposition 3.4

(i) For $n > 1$ we have $\mathbf{x}^{\mathbf{w}} \in Sm(I(V))$ if and only if there are at least $w_n + 1$ elements $\beta \in \{\alpha_1, \ldots, \alpha_k\}$ such that $x_1^{w_1} \ldots x_{n-1}^{w_{n-1}} \in Sm(I(V_\beta))$.

(ii) For $n > 1$ we have $\mathbf{x}^{\mathbf{w}} \in Lm(I(V))$ if and only if there are at least $k - w_n$ elements $\beta \in \{\alpha_1, \ldots, \alpha_k\}$ such that $x_1^{w_1} \ldots x_{n-1}^{w_{n-1}} \in Lm(I(V_\beta))$.

From this one can easily see that the degree of any variable in a lexicographic standard monomial can be at most $k-1$, i.e.

$$Sm(I(V)) \subseteq \{\mathbf{x}^{\mathbf{w}} : \mathbf{w} \in \{0, 1, \ldots, k-1\}^n\}.$$

Given the lexicographic standard and leading monomials of a vanishing ideal $I(V)$ one can compute them for $I(V^c)$ as well using the following proposition.

Proposition 3.5 *For every monomial* $\mathbf{x^w}$ *we have* $x_1^{w_1} \dots x_n^{w_n} \in Sm(I(V))$ *if and only if* $x_1^{k-1-w_1} \dots x_n^{k-1-w_n} \in Sm(I(V^c))$.

Proof: The proof is based on the lex game. See [2]. ■

A fast algorithm

Based on Proposition 3.4, we can construct a very fast algorithm for computing the standard monomials for a given set of points $V \subseteq \{\alpha_1, \dots, \alpha_k\}^n$ ($V \subseteq \{0, 1, \dots, k-1\}^n$). First let us build up a prefix tree, also called a trie, from the elements of V. (For the precise definitions see [19] or in Hungarian [20].) If we take a vertex v from the i^{th} level (a vertex at distance i from the root) and take the subtree rooted at v, then it will be the prefix tree of some set $V_{\beta_n,\beta_{n-1},\dots,\beta_{n+1-k}}$. Let us notice that on the $(n-1)^{th}$ level there will be one dimensional sets, corresponding to the case $n = 1$. For this case Proposition 3.4 gives that $x^l \in Sm(I(V))$ if and only if $l < |V|$. Thus we can easily compute the standard monomials for the sets on the $(n-1)th$ level, and using the recursive property, the standard monomials of the sets on higher levels. At the end we get $Sm(I(V))$.

Felszeghy, Ráth and Rónyai showed in [2] that with a more appropriate data structure one can do this in linear time.

Theorem 3.3 *Let* k *be the maximal degree of the trie and* $|V| = m$. *There is an algorithm, which computes* $Sm(I(V))$ *in* $O(nmk)$ *time. If we assume that there exists an ordering on the coordinate set of* V, *which can be tested in constant time then the algorithm makes* $O(nm \log k)$ *steps.*

Proof: See [2]. ■

Notice that in our case the size of the input is also $nm \log k$, so the algorithm is really linear.

4 Shattering

In this section we first introduce the central notion of our study, shattering.

Definition 4.1 *A set system $\mathcal{F} \subseteq 2^{[n]}$ shatters a given set $S \subseteq [n]$ if $2^S = \{F \cap S : F \in \mathcal{F}\}$. The family of subsets of $[n]$ shattered by $\mathcal{F}$ is denoted by $Sh(\mathcal{F})$.*

The size of $Sh(\mathcal{F})$ will play a key role in this paper. The following proposition gives a very surprising lower bound:

Proposition 4.1 *(See [1].)* $|Sh(\mathcal{F})| \geq |\mathcal{F}|$.

Proof: We will prove this statement by induction on n. For $n = 1$ the statement is trivial. Now suppose that $n > 1$. We construct 2 new set systems analogously to the previous section:

$$\mathcal{F}_0 = \{F : F \in \mathcal{F}; n \notin F\},$$
$$\mathcal{F}_1 = \{F \backslash \{n\} : F \in \mathcal{F}; n \in F\}.$$

Clearly $|\mathcal{F}| = |\mathcal{F}_0| + |\mathcal{F}_2|$, and by induction we have $|Sh(\mathcal{F}_0)| \geq |\mathcal{F}_0|$ and $|Sh(\mathcal{F}_1)| \geq |\mathcal{F}_1|$. It is obvious that $Sh(\mathcal{F}_0) \cup Sh(\mathcal{F}_1) \subseteq Sh(\mathcal{F})$. However, if $S \in Sh(\mathcal{F}_0) \cap Sh(\mathcal{F}_1)$, then according to the definition of $\mathcal{F}_0$ and $\mathcal{F}_1$ we have $S \cup \{n\} \in Sh(\mathcal{F})$. So altogether we have

$$|Sh(\mathcal{F})| \geq |Sh(\mathcal{F}_0)| + |Sh(\mathcal{F}_1)| \geq |\mathcal{F}_0| + |\mathcal{F}_1| = |\mathcal{F}|. \blacksquare$$

Thus every set system $\mathcal{F}$ shatters at least $|\mathcal{F}|$ sets. This inequality was proved by various authors (Aharoni and Holzman [10], Pajor [11], Sauer [12], Shelah [13]), and studied by many others. We are interested in the case of equality, when a set system shatters exactly $|\mathcal{F}|$ sets:

Definition 4.2 *The set system $\mathcal{F}$ is called S-extremal if $|Sh(\mathcal{F})| = |\mathcal{F}|$.*

From now on S-extremal set systems will be referred to as extremal. A good survey on extremal set theory is provided by [4]. When considering this

definitionm, one can make a useful observation in connection with Proposition 4.1. In its proof we have seen a decomposition of $\mathcal{F}$ and a recursion-like inequality for the sizes of the families of sets shattered by them. From this inequality we can conclude the following corollary:

Corollary 4.1.1 *If $\mathcal{F}$ is extremal, then so is $\mathcal{F}_0$ and $\mathcal{F}_1$.*

Our aim is to characterize somehow extremal set systems. Before getting started with this, we first present some interesting results in connection with shattering.

We start with an immediate consequence of Proposition 4.1, also known as Sauer's lemma, which has found applications in a variety of contexts, including applied probability.

Proposition 4.2 *(See [12], [13], [14].) Let $\mathcal{F}$ be a family of subsets of $[n]$ with no shattered set of size k. Then*

$$|\mathcal{F}| \leq \left(\begin{array}{c} n \\ k-1 \end{array} \right) + \left(\begin{array}{c} n \\ k-2 \end{array} \right) + \cdots + \left(\begin{array}{c} n \\ 0 \end{array} \right),$$

and this inequality is best possible.

Proof: Clearly $Sh(\mathcal{F})$ is a down-set, and since there is no shattered set of size k, we have

$$Sh(\mathcal{F}) \subseteq [n]_{k-1} \cup [n]_{k-2} \cup \cdots \cup [n]_0,$$

thus

$$|Sh(\mathcal{F})| \leq \left(\begin{array}{c} n \\ k-1 \end{array} \right) + \left(\begin{array}{c} n \\ k-2 \end{array} \right) + \cdots + \left(\begin{array}{c} n \\ 0 \end{array} \right).$$

Combining this with Proposition 4.1 we get the desired inequality, where with $\mathcal{F} = [n]_{k-1} \cup [n]_{k-2} \cup \cdots \cup [n]_0$ equality is possible. ■

For uniform families Frankl and Pach in [18] prooved the following:

Proposition 4.3 *Let $\mathcal{F}$ be a uniform family of subsets of $[n]$, i.e. $\mathcal{F} \subseteq [n]_l$ for some l, with no shattered set of size k. Then*

$$|\mathcal{F}| \leq \begin{pmatrix} n \\ k-1 \end{pmatrix}.$$

Proof: See [18]. ■

To finish with, we present a version of shattering, introduced by Anstee, Rónyai and Sali in [1] that does always result in equality in a version of Proposition 4.1. We define the concept of order shattered in an inductive way.

Definition 4.3 *We say that the set $S = \{s_1, s_2, \ldots, s_d\} \subseteq [n]$ is order shattered by a given family $\mathcal{F} \subseteq 2^{[n]}$ if the following holds: in the case $S = \emptyset$ the family $\mathcal{F}$ has to contain a set; when $|S| > 0$ and $s_1 < s_2 < \cdots < s_d$, then there are 2^d sets in $\mathcal{F}$ that can be divided into two families $\mathcal{F}_0$ and $\mathcal{F}_1$ such that $s_d \notin F_0$ for all $F_0 \in \mathcal{F}_0$, $s_d \in F_1$ for all $F_1 \in \mathcal{F}_1$, and both $\mathcal{F}_0, \mathcal{F}_1$ order shatter the set $S \backslash \{s_d\}$, furthermore $T \cap F_0 = T \cap F_1$ holds for $T = \{s_d + 1, s_d + 2, \ldots, n\}$ and for all $F_0 \in \mathcal{F}_0$, $F_1 \in \mathcal{F}_1$.*

Let $osh(\mathcal{F})$ be the family of sets order shattered by $\mathcal{F}$. It is easy to see that $osh(\mathcal{F})$ is a down-set for every $\mathcal{F} \subseteq 2^{[n]}$. For the size of $osh(\mathcal{F})$ Anstee, Rónyai and Sali proved in [1] the following:

Proposition 4.4 *Let $\mathcal{F}$ be a family of subsets of $[n]$. Then*

$$|osh(\mathcal{F})| = |\mathcal{F}|.$$

Proof: See [1]. ■

For further definitions and the algebraic interpretation of $osh(\mathcal{F})$ see [1] and as an example of its applications see [16].

4.1 Algebraic approach

There are many algebraic methods that play an important role in combinatorics. For such methods see [21]. When studying extremal set systems it turned out that standard monomials can be of great help. In the following sections we deal with the standard monomials of vanishing ideals for the case $k = 2$, thus we are given a set $V \subseteq \{0,1\}^n$, and we are interested in $Sm(I(V))$. If we consider the elements of V as characteristic vectors of subsets of $[n]$, then V can be viewed as a set system of $2^{[n]}$.

For a subset $H \subseteq [n]$ denote by x_H the monomial $\prod_{i \in H} x_i$; in particular, $x_\emptyset = 1$. Using this notation, a set system $V \subseteq 2^{[n]}$ can be viewed as a set of monomials.

According to these correspondences, from now on, it will depend on the context, whether we are considering a set of vectors, sets or monomials, and all corresponding definitions are naturally extended to all three cases.

Previously we have already discussed the properties of standard monomials in the general case. Now let us recall some of them for the case of vanishing ideals and $k = 2$. For $\mathcal{F} \subseteq \{0,1\}^n$ ($\mathcal{F} \subseteq 2^{[n]}$) we have:

- $|Sm(I(\mathcal{F}))| = |\mathcal{F}|$
- $Sm(I(\mathcal{F})) \subseteq \{\mathbf{x}^{\mathbf{w}} : \mathbf{w} \in \{0,1\}^n\} = \{x_H, H \subseteq [n]\}$ ($Sm(I(\mathcal{F})) \subseteq 2^{[n]}$)
- $Sm(I(\mathcal{F})) = Sm(I(\mathcal{F}_0)) \cup Sm(I(\mathcal{F}_1)) \cup \{D \cup \{n\} : D \in Sm(I(\mathcal{F}_0)) \cap Sm(I(\mathcal{F}_1))\}$ (recursive property - Proposition 3.4)
- $Sm(I(\mathcal{F}))$ can be computed in $O(n|\mathcal{F}|)$ time (Theorem 3.3)

Now we present some statements as a preparation to our main results. We investigate the connection between the standard monomials and the family of shattered sets. For extremal families this connection clearly must be independent in some sense of the term order, since the shattered sets themselves do not depend on the term order either.

Proposition 4.5 *If $x_H \in Sm(I(\mathcal{F}))$ for some term order, then $H \in Sh(\mathcal{F})$.*

Proof: Suppose that H is not shattered by $\mathcal{F}$. This means that there exists a $G \subseteq H$ for which there is no $F \in \mathcal{F}$ such that $G = H \cap F$. Consider the polynomial $f(\mathbf{x}) = x_G(\prod_{j \in H \setminus G}(x_j - 1))$. Denote the characteristic vector of the set F by $\mathbf{v}_F$. Now $f(\mathbf{v}_F) \neq 0$ only if $H \cap F = G$. According to our assumption, there is no such set $F \in \mathcal{F}$, so $f(\mathbf{x})$ vanishes on $\mathcal{F}$, and so it is in $I(\mathcal{F})$. This implies that $x_H \in Lm(I(\mathcal{F}))$ for all term orders, and so we got a contradiction. ■

This means that $Sm(I(\mathcal{F})) \subseteq Sh(\mathcal{F})$ for every term order. We now investigate the other direction:

Proposition 4.6 *If $H \in Sh(\mathcal{F})$, then there is a lexicographic term order for which we have $x_H \in Sm(I(\mathcal{F}))$.*

Proof: We prove that a lexicographic order where the variables of x_H are the smallest satisfies the condition. Suppose the contrary, namely that $x_H \in Lm(I(\mathcal{F}))$ for this term order. Then there is a polynomial $f(\mathbf{x})$ vanishing on $\mathcal{F}$ with leading monomial x_H. Since the variables in x_H are the smallest according to this term order, there cannot appear any other variable in $f(x)$. So $f(\mathbf{x})$ has the form $\sum_{G \subseteq H} \alpha_G x_G$. Take a subset $G_0 \subseteq H$ which appears with a nonzero coefficient in $f(\mathbf{x})$, and is minimal. $\mathcal{F}$ shatters H, so there exists a set $F_0 \in \mathcal{F}$ such that $G_0 = F_0 \cap H$. For this we have $x_{G_0}(\mathbf{v}_{F_0}) = 1$, and since G_0 was minimal, $x_G(\mathbf{v}_{F_0}) = 0$ for every other set G. So

$$\sum_{G \subseteq H} \alpha_G x_G(\mathbf{v}_{F_0}) = \alpha_{G_0} \neq 0.$$

But on the other hand, since $f(\mathbf{x}) \in I(\mathcal{F})$, $f(\mathbf{v}_{F_0}) = 0$. This contradiction proves the statement. ■

Combining the last two results we have

$$\mathcal{S}h(\mathcal{F}) = \bigcup_{lex\ orders} Sm(I(\mathcal{F})).$$

Even though $Sm(\mathcal{F})$ can be computed fast for every term order according to Theorem 3.3, this formula does not give an efficient way for computing $Sh(\mathcal{F})$, because the number of lexicographic term orders is $n!$. However for extremal set systems we obtain the following very important corollary:

Corollary 4.6.1 *$\mathcal{F}$ is extremal if and only if $Sm(I(\mathcal{F}))$ is the same for all lexicographic term orders.*

Proof: Suppose that $\mathcal{F}$ is extremal, i.e. $|\mathcal{F}| = |Sh(\mathcal{F})|$. Since $\mathcal{S}h(\mathcal{F}) = \bigcup_{lex\ orders} Sm(I(\mathcal{F}))$ and for every term order $|Sm(I(\mathcal{F}))| = |\mathcal{F}|$, there cannot be two lexicographic term orders for which the set of standard monomials differ, otherwise the first equality could not hold. The other direction can be proved in a similar way. ■

From this corollary we can make another useful observation. Suppose that $\mathcal{F} \subseteq 2^{[n]}$ is a down-set. From the definition of shattering one can easily see that $\mathcal{F}$ shatters all of its elements and no other set, that is $Sh(\mathcal{F}) = \mathcal{F}$ and so $\mathcal{F}$ is extremal. From this one can conclude that in this case $Sm(I(\mathcal{F})) = \mathcal{F}$ holds as well. In particular, for a down-set $\mathcal{F}$, the standard monomials, the family of sets shattered by $\mathcal{F}$ and $\mathcal{F}$ coincide.

4.2 Result on shattering

For a pair of sets $G \subseteq H \subseteq [n]$ define the following polynomial

$$f_{H,G} = (\prod_{j \in G} x_j)(\prod_{i \in H \setminus G} (x_i - 1)).$$

Proposition 4.7 *If $S \notin Sh(\mathcal{F})$, then there exists a set $H \subseteq S$ such that $f_{S,H}(\mathbf{v}_F) = 0$, $\forall\ F \in \mathcal{F}$, i.e. $f_{S,H} \in I(\mathcal{F})$.*

Proof: The statement was already proved in Proposition 4.5. ■

Proposition 4.8 *If the set S in the previous proposition is minimal (in the sense that all proper subsets S' of S are in $Sh(\mathcal{F})$) and $\mathcal{F}$ is extremal, then the corresponding H is unique.*

Proof: Suppose that there are two different sets $H_1, H_2 \subseteq S$ for which $f_{S,H_i} \in I(\mathcal{F})$ for $i = 1, 2$. Then $g = f_{S,H_1} - f_{S,H_2} \in I(\mathcal{F})$. Let us fix a term order. For this term order $lm(g) = x_{S'}$ with a set $S' \subsetneq S$. $\mathcal{F}$ is extremal, so $Sm(I(\mathcal{F})) = Sh(\mathcal{F})$. But $x_{S'}$ is not a standard monomial and therefore it is not shattered by $\mathcal{F}$. This contradicts with the minimality of S, hence the corresponding H is unique. ■

When reversing this statement one can get another characterization of extremal set systems:

Proposition 4.9 *If for all but $|\mathcal{F}|$ sets $S \subseteq [n]$ there exists a set $H \subseteq S$ for which $f_{S,H} \in I(\mathcal{F})$, then $\mathcal{F}$ is extremal.*

Proof: $f_{S,H} \in I(\mathcal{F})$ and $lm(f_{S,H}) = x_S$ holds for all term orders. So for all but $|\mathcal{F}|$ sets $S \in Lm(I(\mathcal{F}))$ for all term orders. Fix a term order, and consider the set X of standard monomials with respect to this term order. Then X must be $Sm(I(\mathcal{F}))$ for all term orders, from which it follows by Corollary 4.6.1 that $\mathcal{F}$ is extremal. ■

Now we have made all necessary preparations to present our first result together with its proof. We have characterized extremal set system using Gröbner bases. To our knowledge, this is the first time that Gröbner bases are used for characterizing extremal set systems.

Theorem 4.1 *$\mathcal{F} \subseteq 2^{[n]}$ is extremal if and only if there are polynomials of the form $f_{S,H}$, which together with $\{x_i^2 - x_i, i \in [n]\}$ form a Gröbner basis of $I(\mathcal{F})$ for all term orders.*

Proof: For the first part, suppose that $\mathcal{F}$ is extremal. Consider all minimal sets $S \subseteq [n]$, $S \notin Sh(\mathcal{F})$ with the corresponding unique polynomials $f_{S,H}$. Denote the set of these sets by $\mathcal{S}$ and fix a term order. We prove that these polynomials, together with $\{x_i^2 - x_i, i \in [n]\}$, form a Gröbner basis of $I(\mathcal{F})$. In order to show this we have to prove that for all monomials $m \in Lm(I(\mathcal{F}))$, there is a monomial in $\{x_S, S \in \mathcal{S}\} \cup \{x_i^2, i \in [n]\}$ that divides m. If there

is a variable in m with degree higher then 1, then this is trivial. Since $\mathcal{F}$ is extremal, we have $Sm(I(\mathcal{F})) = Sh(\mathcal{F})$, and this, together with the minimality of the sets in $\mathcal{S}$, proves the statement in the case when m is of the form x_M.
For the other direction, suppose that there is a common Gröbner basis G for all term orders of the desired form. Denote the collection of the sets S in the polynomials of the form $f_{S,H}$ in G by $\mathcal{S}$. Since the leading monomial of $f_{S,H}$ is x_S for all term orders, $Lm(G) = \{x_S, S \in \mathcal{S}\} \cup \{x_i^2, i \in [n]\}$. This fact, together with the properties of a Gröbner basis, imply that $Sm(\mathcal{F}) = \{x_F, F \subseteq [n], \nexists S \in \mathcal{S} \text{ such that } S \subseteq F\}$ for all term orders. So $Sm(I(\mathcal{F}))$ is the same for all term orders, which means by Corollary 4.6.1 that $\mathcal{F}$ is extremal. ■

4.3 Testing extremality

The importance of any good characterization, in addition to its mathematical beauty, is the possibility of an efficient algorithm. Along this line of thinking we propose two algorithms for deciding the extremality of a set system. To our best knowledge neither of them have been presented so far. The first one is a straightforward implementation of Theorem 4.1. The second one is simple as well, moreover it has a very good running time.

Test #1

Let $\mathcal{F}$ be a set system, and let us fix a lexicographic term order $\prec$. For a lexicographic term order $Sm(\mathcal{F})$ can be computed fast (see [2]). Suppose that $\mathcal{F}$ is extremal. In this case, according to Corollary 4.6.1, $Sm(I(\mathcal{F}))$ is the same for all term orders, so for $\prec$ in particular, we have $Sm(I(\mathcal{F})) = Sh(\mathcal{F})$. From Theorem 4.1 we know that there is a Gröbner basis of a special form, and we can construct it from $Sh(\mathcal{F})$. Take the minimal sets, S for which x_S is not in $Sm(I(\mathcal{F}))$, and denote their set by $\mathcal{S}$. For every $S \in \mathcal{S}$ there must be a (unique) set $H \subseteq S$ such that $f_{H,S} \in I(\mathcal{F})$. Now these polynomials, together with $\{x_i^2 - x_i : i \in [n]\}$, form a Gröbner basis of $I(\mathcal{F})$. According to this, the test runs as follows:

- Compute $Sm(I(\mathcal{F}))$ for an arbitrary lexicographic term order, e.g. standard lex.
- Compute the set family S.
- Construct the corresponding $f_{H,S}$ polynomials.
- Verify if these polynomialsm, together with the polynomials $\{x_i^2 - x_i\}$, form a Gröbner basis of the ideal $I(\mathcal{F})$.

$\mathcal{F}$ is extremal if and only if we get a Gröbner basis with this process. This is straightforward from Proposition 4.1. There are many ways to verify whether a system of polynomials is a Gröbner basis or not. For such methods see [8]. I have not analyzed the time requirement of this method yet, however, it does not seem to be sufficiently efficient.

Test #2

According to the lex game [2] we know that for a fixed lexicographic term order $Sm(I(\mathcal{F}))$ can be computed essentially in linear time. (Note that the size of the input is nm, where m is the size of $\mathcal{F}$.) This forms the base of another extremality test. $\mathcal{F}$ is extremal if and only if for every lex term order $Sm(I(\mathcal{F}))$ is the same. Our aim is to find a family of lexicographic term orders with the property that if $\mathcal{F}$ is not extremal, then we can find two term orders in this family for which the standard monomials differ. This can be done with a family of size n:

Theorem 4.2 *Take n orders of the variables such that for every index i there is one in which x_i is the greatest element, and take the corresponding lexicographic term orders. If $\mathcal{F}$ is not extremal, then among these we can find two term orders for which the standard monomials of $I(\mathcal{F})$ differ.*

Proof: Let us fix one of the above mentioned term orders. $\mathcal{F}$ is not extremal, hence there is a set $H \in \mathcal{F}$ shattered by $\mathcal{F}$ for which x_H is not a standard monomial but a leading one. $Sm(I(\mathcal{F}))$ is a basis of the vector space $\mathbb{F}[\mathbf{x}]/I(\mathcal{F})$, and since all functions from $\mathcal{F}$ to $\mathbb{F}$ are polynomials, every

leading monomial can be written uniquely as the sum of standard monomials, as a function on $\mathcal{F}$. This holds for x_H as well:

$$x_H = \sum \alpha_G x_G,$$

as functions on $\mathcal{F}$. Suppose that for all sets G in the above sum we have $G \subseteq H$. Take a minimal G_0 with a nonzero coefficient. Since H is shattered by $\mathcal{F}$, there is an $F \in \mathcal{F}$ such that $G_0 = F \cap H$. For this $x_{G_0}(\mathbf{v}_F) = 1$. From the minimality of G_0 we have that $x_{G'}(\mathbf{v}_F) = 0$ for every other G'. So

$$\sum \alpha_G x_G(\mathbf{v}_F) = \alpha_{G_0}.$$

On the other hand $x_H(\mathbf{v}_F) = 0$, since $H \cap F = G_0$, but $H \neq G$ because x_H is a leading monomial, and x_G is a standard monomial, and this is a contradiction. Therefore in the above sum there is a set G with nonzero coefficient such that $G \backslash H \neq \emptyset$. Now let us fix an index $i \in G \backslash H$. For the term order where x_i is the greatest variable, x_H cannot be the leading monomial of the polynomial $x_H - \sum \alpha_G x_G$. Then the leading monomial is another $x_{G'}$, which, for the original term order was a standard monomial. So we have found two term orders for which the standard monomials differ. ■

According to this theorem it is enough to calculate the standard monomials e.g. for a lexicographic term order and it's cyclic permutations, and to check, whether they differ or not. Since the standard monomials can be calculated in $O(n|\mathcal{F}|)$ time for one lexicographic term order, and we need n term orders, the total running time of the algorithm is $O(n^2|\mathcal{F}|)$.

Corollary 4.2.1 *Given a set family $\mathcal{F} \subseteq 2^{[n]}$, $|\mathcal{F}| = m$ by a list of characteristic vectors, we can decide in $O(n^2 m)$ time whether $\mathcal{F}$ is extremal or not.*

This improves the algorithm given in [6] by G. Greco, where the time bound is $O(nm^3)$.

Open question 1 *Can extremality be tested in linear time (i.e. in $O(nm)$)?*

5 Set system operations

$Sh(\mathcal{F})$ is a down-set, and it is evident that if $\mathcal{F}$ is a down-set, then $Sh(\mathcal{F}) = \mathcal{F}$, i.e. $\mathcal{F}$ is extremal. So the question presents itself: can every extremal set system be obtained from down-sets by some natural operations? For this we have studied different set system operations.

Bit flip

For vectors $(v_1, \ldots, v_n) \in \{0,1\}^n$ we denote by φ_i the ith bit flip:

$$\varphi_i(v_1, \ldots, v_{i-1}, v_i, v_{i+1}, \ldots, v_n) := (v_1, \ldots, v_{i-1}, 1 - v_i, v_{i+1}, \ldots, v_n).$$

It is easy to verify that extremality is invariant with respect to this operation. However this operation does not have the desired property, not every extremal set system can be obtained from down-sets using only this operation:

Example 5.1 *For the set system* $\mathcal{F} = \{\emptyset, \{1\}, \{2\}, \{1,3\}\}$

$$Sh(\mathcal{F}) = \{\emptyset, \{1\}, \{2\}, \{3\}\},$$

hence it is extremal. However, this cannot be obtained by bit flips from a down-set. We can verify this by applying bit flips to it in different order. We cannot transform it to a down-set, so it cannot be obtained from a down-set.

Translation

Let $\mathbf{v}$ be a fixed $0-1$ vector of length n. Then in the translation by $\mathbf{v}$ we add up to all characteristic vectors $\mathbf{v}$ *modulo* 2. This corresponds to the compositions of several bit flips, hence this operation also preserves extremality. We denote the translation by the vector $\mathbf{v}$ by $\varphi_{\mathbf{v}}$.

Sum

The set system considered in Example 5.1 cannot be obtained from a down-set using the previous operations. To fix this problem, we introduce a new operation, the sum of two set systems.

Definition 5.1 *Let $\mathcal{F}_1$ and $\mathcal{F}_2$ be two set systems with disjoint supports (there is no $i \in [n]$ for which $\exists\ F_1 \in \mathcal{F}_1$ and $F_2 \in \mathcal{F}_2$ such that $i \in F_1 \cap F_2$) and $\emptyset \in \mathcal{F}_1 \cap \mathcal{F}_2$. We define the sum of these set systems as*

$$\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2 := \mathcal{F}_1 \cup \mathcal{F}_2.$$

Proposition 5.1 *$\mathcal{F}$ is extremal if and only if $\mathcal{F}_1$ and $\mathcal{F}_2$ are both extremal.*

Proof: $\mathcal{F}_1$ and $\mathcal{F}_2$ have disjoint supports, hence $Sh(\mathcal{F}) = Sh(\mathcal{F}_1) \cup Sh(\mathcal{F}_2)$ and $Sh(\mathcal{F}_1) \cap Sh(\mathcal{F}_2) = \{\emptyset\}$, which means that

$$|Sh(\mathcal{F})| = |Sh(\mathcal{F}_1)| + |Sh(\mathcal{F}_2)| - 1. \tag{1}$$

On the other hand $\mathcal{F}_1 \cap \mathcal{F}_2 = \{\emptyset\}$, hence

$$|\mathcal{F}| = |\mathcal{F}_1| + |\mathcal{F}_2| - 1. \tag{2}$$

The statement follows directly from (1) and (2). ■

Let us go back to Example 5.1. This extremal family can be obtained from down-sets as follows:

$$\mathcal{F} = \{\emptyset, \{2\}\} + \varphi_1(\{\emptyset, \{1\}, \{3\}\})$$

Open question 2 *Can every extremal family $\mathcal{F}$ be obtained from downward families by translations and sums?*

6 Downshifts

In this section we study one of the most frequently used set operations in this field. Let us denote by D_i the downshift by the element i. If $\mathcal{F} \subseteq 2^{[n]}$ then:

Definition 6.1 $D_i(\mathcal{F}) := \{F \in \mathcal{F} | i \notin F\} \cup \{F \in \mathcal{F} | i \in F, F\backslash\{i\} \in \mathcal{F}\} \cup \{F\backslash\{i\} | F \in \mathcal{F}, i \in F, F\backslash\{i\} \notin \mathcal{F}\}$.

From the definition it is clear that $|\mathcal{F}| = |D_i(\mathcal{F})|$ and if $\mathcal{F}$ is a down-set, then a downshift has no effect on $\mathcal{F}$. When studying the structure of $D_i(\mathcal{F})$ we can observe that $\mathbf{v} \in D_i(\mathcal{F})$ if and only if there are $v_i + 1$ vectors in $\mathcal{F}$ equal to $\mathbf{v}$ in all but the i^{th} coordinate.

It is easily seen that downshifts and bit flips commute:

$$\varphi_i(D_j(\mathcal{F})) = D_j(\varphi_i(\mathcal{F}))$$

holds for any set family $\mathcal{F}$ and $1 \leq i, j \leq n$. (Here D_j is assumed to act on the set of characteristic vectors $\{\mathbf{v}_F : F \in \mathcal{F}\} \subseteq \{0,1\}^n$.) Now we look at some properties of the downshift operation.

Proposition 6.1 *(See [3].) For every $i \in [n]$ we have $Sh(D_i(\mathcal{F})) \subseteq Sh(\mathcal{F})$.*

Proof: Let $S \in Sh(D_i(\mathcal{F}))$. If $i \notin S$, then clearly we have $S \in Sh(\mathcal{F})$. Now suppose that $i \in S$. Since S is shattered by $D_i(\mathcal{F})$, for every set $H \subseteq S$ there is a set $F_H \in D_i(\mathcal{F})$ such that $F_H \cap S = H$. For any set $H \subseteq S$ let us define the set G_H to be F_H if $i \in H$ and $F_{H\cup\{i\}}\backslash\{i\}$ if $i \notin H$. With this definition we have $G_H \in \mathcal{F}$ (since if $i \in G \in D_i(\mathcal{F})$ then $G, G\backslash\{i\} \in \mathcal{F}$ must hold) and $G_H \cap S = H$ for every set $H \subseteq S$. The family $\{G_H : H \subseteq S\}$ shows that $S \in Sh(\mathcal{F})$. ■

Corollary 6.1.1 *If $\mathcal{F} \subseteq 2^{[n]}$ is extremal, then $D_i(\mathcal{F})$ is extremal for every $i \in [n]$.*

Proof: From the extremality of $\mathcal{F}$ and from the previous proposition, for every $i \in [n]$ we have:

$$|\mathcal{F}| = |Sh(\mathcal{F})| \geq |Sh(D_i(\mathcal{F}))| \geq |D_i(\mathcal{F})|.$$

Since $|\mathcal{F}| = |D_i(\mathcal{F})|$, there must be equality everywhere, giving that $D_i(\mathcal{F})$ is extremal. ■

Let $\mathcal{F} \subseteq 2^{[n]}$. For the indices $i_1, i_2, \ldots, i_l$ we introduce the following notation:

$$D_{i_1,i_2,\ldots,i_l}(\mathcal{F}) := D_{i_1}(D_{i_2}(\ldots(D_{i_l}(\mathcal{F}))))$$

When applying several different downshifts the question arises whether the order of the downshifts is relevant. In general, the order of the downshifts has an effect on the result. For an example consider the set system $\mathcal{F} = \{\emptyset, \{1,2\}\}$. For this we have $D_{1,2}(\mathcal{F}) = \{\emptyset, \{1\}\}$ and $D_{2,1}(\mathcal{F}) = \{\emptyset, \{2\}\}$, thus $D_{1,2}(\mathcal{F}) \neq D_{2,1}(\mathcal{F})$. However, for extremal families we have the following result:

Proposition 6.2 *If $\mathcal{F} \subseteq 2^{[n]}$ is extremal, then different downshifts commute, i.e. $D_{i,j}(\mathcal{F}) = D_{j,i}(\mathcal{F})$.*

Proof: Without loss of generality we can suppose that $i, j = 1, 2$. For $H \subseteq \{3, 4, \ldots, n\}$ we denote by $\mathcal{F}(H)$ the family

$$\mathcal{F}(H) = \{F \in \mathcal{F} : F \cap \{3, 4, \ldots, n\} = H\}$$

From the definitions it is immediate that for a family $\mathcal{F} \subseteq 2^{[n]}$, $H \subseteq \{3, 4, \ldots, n\}$ and $i \in \{1, 2\}$ we have

$$D_i(\mathcal{F}(H)) = D_i(\mathcal{F})(H).$$

Moreover

$$D_i(\mathcal{F}) = \bigcup_{H \subseteq \{3,4,\ldots,n\}} D_i(\mathcal{F}(H)),$$

and for $\{i, j\} = \{1, 2\}$ we have

$$D_i(D_j(\mathcal{F})) = D_i(\bigcup_{H \subseteq \{3,4,\ldots,n\}} D_j(\mathcal{F}(H))) = D_i(\bigcup_{H \subseteq \{3,4,\ldots,n\}} D_j(\mathcal{F})(H)) =$$
$$= \bigcup_{H \subseteq \{3,4,\ldots,n\}} D_i(D_j(\mathcal{F})(H)) = \bigcup_{H \subseteq \{3,4,\ldots,n\}} D_i(D_j(\mathcal{F}(H))).$$

Thus it suffices to verify the claim $D_{1,2}(\mathcal{F}) = D_{2,1}(\mathcal{F})$ for families $\mathcal{F}$ of the form $\mathcal{G}(H)$ where $\mathcal{G}$ is an extremal set system and $H \subseteq \{3, 4, \ldots, n\}$. But this reduces the problem to extremal families $\mathcal{F} \subseteq 2^{[2]}$.
Note that for a vector $\mathbf{v} \in \{0,1\}^n$ and $\mathcal{F}_1, \mathcal{F}_2 \subseteq 2^{[n]}$

$$\varphi_{\mathbf{v}}(\mathcal{F}_1) = \varphi_{\mathbf{v}}(\mathcal{F}_2) \Longleftrightarrow \mathcal{F}_1 = \mathcal{F}_2$$

If $\mathcal{F}$ is not empty, then by a composition $\varphi_{\mathbf{v}}$ of some bit flips we can achieve that $\emptyset \in \mathcal{F}$. Also,

$$D_{1,2}(\mathcal{F}) = D_{2,1}(\mathcal{F}) \Longleftrightarrow \varphi_{\mathbf{v}}(D_{1,2}(\mathcal{F})) = \varphi_{\mathbf{v}}(D_{2,1}(\mathcal{F})) \Longleftrightarrow$$
$$D_{1,2}(\varphi_{\mathbf{v}}(\mathcal{F})) = D_{2,1}(\varphi_{\mathbf{v}}(\mathcal{F})).$$

Thus to verify $D_{1,2}(\mathcal{F}) = D_{2,1}(\mathcal{F})$, we can assume that $\emptyset \in \mathcal{F}$. If $\mathcal{F}$ is a down-set, we are done, since $D_i(\mathcal{F}) = \mathcal{F}$. Now if $\mathcal{F} \subseteq 2^{[n]}$, $\emptyset \in \mathcal{F}$, $\mathcal{F}$ is extremal and $\mathcal{F}$ is not a down-set, then $\{1,2\} \in \mathcal{F}$ and we have

$$\mathcal{F} = \{\emptyset, \{1\}, \{1,2\}\}$$

or

$$\mathcal{F} = \{\emptyset, \{2\}, \{1,2\}\}.$$

By symmetry, it suffices to do the calculation for the first case. Then $D_1(\mathcal{F}) = \mathcal{F}^* = \{\emptyset, \{1\}, \{2\}\}$, $D_2(\mathcal{F}) = \mathcal{F}$ and $D_2(\mathcal{F}^*) = \mathcal{F}^*$. Thus

$$D_{1,2}(\mathcal{F}) = D_1(\mathcal{F}) = \mathcal{F}^* = D_2(\mathcal{F}^*) = D_{2,1}(\mathcal{F}). \blacksquare$$

We have already seen that down-sets have very good properties. Now we weaken the definition of a down-set.

Definition 6.2 *$\mathcal{F} \subseteq 2^{[n]}$ is an i-down-set if for every $F \in \mathcal{F}$ with $i \in F$ we have $F \backslash \{i\} \in \mathcal{F}$.*

From the definition we can make some trivial observations. For every $\mathcal{F} \subseteq 2^{[n]}$, $D_i(\mathcal{F})$ is an i-down-set and furthermore if $\mathcal{F} \subseteq 2^{[n]}$ is an i-down-set, then $D_i(\mathcal{F}) = \mathcal{F}$.

Proposition 6.3 *If $\mathcal{F} \subseteq 2^{[n]}$ is an i-down-set, then $D_j(\mathcal{F})$ is an i-down-set as well, for every $j \neq i$.*

Proof: This statement will be proved in Section 10 in a general case. (Proposition 10.2) ■

We have the following important consequence:

Proposition 6.4 *Let $\mathcal{F} \subseteq 2^{[n]}$. Then n different downshifts applied to $\mathcal{F}$ in an arbitrary order result in a down-set.*

Proof: According to the previous statements, after applying a downshift with i we get an i-down-set and this property remains invariant under further downshifts. Therefore after n different downshifts we get a set system $\widetilde{\mathcal{F}}$ which is an i-down-set and so $D_i(\widetilde{\mathcal{F}}) = \widetilde{\mathcal{F}}$ for all $i \in [n]$. This last property is equivalent to the fact that $\widetilde{\mathcal{F}}$ is a down-set. ■

The preceding statement occurs in [3], and can be proved by induction on n as well. In Proposition 6.1 we have seen that for a family $\mathcal{F} \subseteq 2^{[n]}$ we have $Sh(D_i(\mathcal{F})) \subseteq Sh(\mathcal{F})$ for all $i \in [n]$. Thus

$$Sh(D_{i_1,\ldots,i_n}(\mathcal{F})) \subseteq Sh(\mathcal{F})$$

if all the i_k-s are different. On the other hand Proposition 6.4 says that $D_{i_1,\ldots,i_n}(\mathcal{F})$ is a down-set, hence $Sh(D_{i_1,\ldots,i_n}(\mathcal{F})) = D_{i_1,\ldots,i_n}(\mathcal{F})$. This also means that

$$|Sh(D_{i_1,\ldots,i_n}(\mathcal{F}))| = |D_{i_1,\ldots,i_n}(\mathcal{F})| = |\mathcal{F}|$$

But if $\mathcal{F}$ was not extremal then $|Sh(\mathcal{F})| > |\mathcal{F}|$, so with the downshifts the size of Sh decreases from $|Sh(\mathcal{F})|$ to $|\mathcal{F}|$. (So one of the downshifts will result an extremal set system.) However, if $\mathcal{F}$ is extremal then, by Proposition 6.1 and Proposition 4.1, the size of Sh cannot be decreased by downshifts. This fact gives the idea of the following definition:

Definition 6.3 *$\mathcal{F} \subseteq 2^{[n]}$ is weakly extremal if for every $i \in [n]$ we have*

$$Sh(D_i(\mathcal{F})) = Sh(\mathcal{F}).$$

In Section **??** we discuss the connection between extremality and weak extremality.

In Section 3 we have already discussed a method for constructing $Sm(\mathcal{F})$ for a set system $\mathcal{F} \subseteq 2^{[n]}$ using the recursive property. Now we give another method using the downshift operation.

Theorem 6.1 *Let $\mathcal{F} \subseteq 2^{[n]}$ and $\prec$ be a lexicographic term order for which $x_{i_1} \succ x_{i_2} \succ \cdots \succ x_{i_n}$. If we apply the downshifts $D_{i_1}, D_{i_2}, \ldots, D_{i_n}$ to $\mathcal{F}$ in this order, then we have $D_{i_n, i_{n-1}, \ldots, i_1}(\mathcal{F}) = Sm(I(\mathcal{F}))$.*

Proof: We will prove a more general form of this statement in Section 10. (Proposition 10.1) ■

7 A graph-theoretical aspect

In this section we make some observations related to [6], and develop some extensions of the ideas presented there. As before, the elements of a set system $\mathcal{F} \subseteq 2^{[n]}$ can be regarded as characteristic vectors, i.e. as $0-1$ vectors of length n, thus $\mathcal{F} \subseteq \{0,1\}^n$.

The n-cube is a graph $Q_n = (\{0,1\}^n, E_n)$ where E_n is the set of pairs $\{F,G\}$ from $\{0,1\}^n$ such that F and G differ in just one component. If we denote by $d(F,G)$ the number of coordinates in which the two vectors differ (i.e., the Hamming distance of the pair), then

$$E_n = \{\{F,G\} : d(F,G) = 1\}.$$

We note that the number of edges of the shortest path connecting any pair of vectors in Q_n coincides with the Hamming distance of the pair. If $\mathcal{F} \subseteq 2^{[n]}$ and $F, G \in \mathcal{F}$, let $d_{\mathcal{F}}(F,G)$ be the number of edges in the shortest path between F and G in the subgraph induced by $\mathcal{F}$ in Q_n if a path exists, $d_{\mathcal{F}}(F,G) = \infty$ otherwise. In this section we alternate between the vector view and the set view of the elements of $\mathcal{F}$. The following two notations are from [6].

Definition 7.1 *$\mathcal{F} \subseteq \{0,1\}^n$ is isometrically embedded in Q_n if for any pair of different elements F and G in $\mathcal{F}$*

$$d_{\mathcal{F}}(F,G) = d(F,G).$$

Definition 7.2 *$\mathcal{F} \subseteq \{0,1\}^n$ is strongly isometrically embedded in Q_n if $D_{i_1,i_2,\ldots,i_m}(\mathrm{F})$ is isometrically embedded for every m and $i_1, i_2, \ldots, i_m \in [n]$.*

The next two propositions discuss the connection between these notions and extremality. They were already proposed by G.Greco, but the proofs below seem to be simpler than those in [6].

Proposition 7.1 *If $\mathcal{F} \subseteq \{0,1\}^n$ is extremal, then $\mathcal{F}$ is isometrically embedded in Q_n.*

Proof: Suppose the contrary, namely that $\mathcal{F}$ is not isometrically embedded in Q_n. Then there exist sets $A, B \in \mathcal{F}$ such that $d(A,B) = k < d_{\mathcal{F}}(A,B)$. Suppose that k is minimal. Clearly $k \geq 2$. The Hamming distance is invariant under bit flips, and using bit flips one can achieve that $A = \emptyset$ and $|B| = k$.
We prove that there is no set C for which $C \in \mathcal{F}$ and $C \subset B$. Otherwise

$$d(A,C) + d(C,B) = k < d_{\mathcal{F}}(A,B) \leq d_{\mathcal{F}}(A,C) + d_{\mathcal{F}}(C,B).$$

We have either $d(A,C) < d_{\mathcal{F}}(A,C)$ or $d(C,B) < d_{\mathcal{F}}(C,B)$. This is a contradiction since $d(A,C) < k$, $d(C,B) < k$ and k was minimal.
If $\mathcal{F}$ is extremal, then applying Corollary 4.1.1 to the elements of $[n]\backslash B$ we get that $\mathcal{H} = \{F \in \mathcal{F} : F \subseteq B\}$ is extremal as well. But in this case $\mathcal{H} = \{\emptyset, B\}$ and $Sh(\mathcal{H}) = \{\emptyset, \{b_1\}, \ldots, \{b_k\}\}$. So $|Sh(\mathcal{H})| = k + 1 \geq 3 > 2 = |\mathcal{H}|$, that is $\mathcal{H}$ cannot be extremal. From this contradiction we have that $\mathcal{F}$ is isometrically embedded in Q_n. ■

Corollary 7.1.1 *If $\mathcal{F} \subseteq \{0,1\}^n$ is extremal, then $\mathcal{F}$ is strongly isometrically embedded in Q_n.*

Proof: If $\mathcal{F}$ is extremal, then so is $D_{i_1,i_2,\ldots,i_m}(\mathrm{F})$ for every m and $i_1, i_2, \ldots, i_m \in [n]$, and according to the previous proposition all of them are isometrically embedded in Q_n. ■

The main result of [6] is that the converse of the last statement also holds. In the following we give a novel characterization of isometrically embedded families. The main result is stated in Theorem 7.1, which simplifies much of the results in [6]. We need some preparations first.

A chunk of the system $\mathcal{F} \subseteq \{0,1\}^n$ is the subsystem that we get by fixing the bits in some positions, i.e. for fixed $i_1, i_2, \ldots, i_m \in [n]$ positions and fixed $\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m$ bits, the chunk $\mathcal{C}$ defined by them is

$$\mathcal{C} = \{F \in \mathcal{F} : F(i_1) = \varepsilon_1, \ldots, F(i_m) = \varepsilon_m\}$$

This notion can be found in [3] already. As a consequence of Proposition 6.4, it follows that when downshifting, the family of shattered sets can shrink if $\mathcal{F}$ is not extremal.

Proposition 7.2 *Let $\mathcal{F} \subseteq \{0,1\}^n$. If for some $i \in [n]$ there exists a set S such that $S \in Sh(\mathcal{F})$ and $S \notin Sh(D_i(\mathcal{F}))$ (i.e. if the downshift with i reduces $Sh(\mathcal{F})$), then $\mathcal{F}$ is not isometrically embedded in Q_n.*

Proof: Since S is not shattered by $D_i(\mathcal{F})$, there is a set $H \subseteq S$ for which there is no set $D \in D_i(\mathcal{F})$ such that $S \cap D = H$. But S is shattered by $\mathcal{F}$, so there is a set $F \in \mathcal{F}$ such that $S \cap F = H$. From the previous observation $F \notin D_i(\mathcal{F})$, so it is downshifted to $F \backslash \{i\}$. This is possible only if $F \backslash \{i\} \notin \mathcal{F}$, and this holds for any set F for which $S \cap F = H$.
On the other hand, since S is shattered by $\mathcal{F}$, there must be a set F' such that $S \cap F' = H \backslash \{i\}$. But $d_{\mathcal{F}}(F', F)$ must be greater than $d(F', F)$ and so $\mathcal{F}$ cannot be isometrically embedded in Q_n, because any shortest path in Q_n between F' and F goes through a set of the form $G \backslash \{i\}$ where $S \cap G = H$ holds for G. But no such set is contained in $\mathcal{F}$. ■

We have a partial converse to the above proposition. At the same time it can be regarded as a characterization for isometrically embedded set systems.

Theorem 7.1 *$\mathcal{F} \subseteq \{0,1\}^n$ is isometrically embedded in Q_n if and only if for every chunk $\mathcal{C}$ and every $i \in [n]$, $Sh(\mathcal{C}) = Sh(D_i(\mathcal{C}))$.*

Proof: For the first direction apply Proposition 7.2. We get that if $\mathcal{F}$ is isometrically embedded in Q_n then there is no set S with the described properties. Since for every $i \in [n]$, $Sh(D_i(\mathcal{F})) \subseteq Sh(\mathcal{F})$, it follows that for every $i \in [n]$, $Sh(D_i(\mathcal{F})) = Sh(\mathcal{F})$. But it is clear that if $\mathcal{F}$ is isometrically embedded in Q_n then so is every chunk of it. So we have $Sh(\mathcal{C}) = Sh(D_i(\mathcal{C}))$ for every $i \in [n]$ and for every chunk $\mathcal{C}$ as well.
For the other direction recall the proof of Proposition 7.1 and as we did there, suppose that $\mathcal{F}$ is not isometrically embedded in Q_n. If we follow that proof, we obtain a chunk $\mathcal{H}$ of $\mathcal{F}$ such that $\mathcal{H} = \{\emptyset, B\}$ with $|B| \geq 2$.

For this chunk we have $Sh(\mathcal{H}) = \{\emptyset, \{b_1\}, \ldots, \{b_k\}\}$. But if we apply a downshift to $\mathcal{H}$ by an index $i \in B$, then $D_i(\mathcal{H}) = \{\emptyset, B\backslash\{i\}\}$ and therefore $Sh(D_i(\mathcal{H})) = Sh(\mathcal{H})\backslash\{\{i\}\}$. So we have found a chunk $\mathcal{H}$ such that $Sh(D_i(\mathcal{H})) \neq Sh(\mathcal{H})$, but this is a contradiction, so $\mathcal{F}$ must be isometrically embedded. ■

We have already mentioned that extremality is equivalent to the fact that $\mathcal{F}$ is strongly isometrically embedded in Q_n, thus the extremality of $\mathcal{F}$ does not follow from the fact that $\mathcal{F}$ is isometrically embedded in Q_n. We will demonstrate this on some examples.

Example 7.1 *$\mathcal{F} = [n]_{k-1} \cup [n]_k$ for $2 \leq k \leq \lceil \frac{n}{2} \rceil$ is isometrically embedded in Q_n, but is not extremal.*

Proof: Clearly $\mathcal{F}$ is isometrically embedded in Q_n. To see whether $\mathcal{F}$ is extremal or not, compute $Sh(\mathcal{F})$. From the condition $k \leq \lceil \frac{n}{2} \rceil$ we get that every set of size at most k is shattered by $\mathcal{F}$ so

$$Sh(\mathcal{F}) = [n]_0 \cup [n]_1 \cup \cdots \cup [n]_k.$$

That means that for $2 \leq k$ we have $\mathcal{F} \subsetneq Sh(\mathcal{F})$, giving that $\mathcal{F}$ is not extremal. ■

The previous example can be generalized:

Example 7.2 *Let $0 < a < b < n$ be integers and $\mathcal{F} = [n]_a \cup [n]_{a+1} \cup \cdots \cup [n]_b$. We claim that $\mathcal{F}$ is isometrically embedded in Q_n but not extremal.*

Proof: As far as extremality is concerned, we are allowed to perform flips. After possibly flipping at all coordinates, we can assume that $a \leq n - b$. With this assumption it is immediate that

$$Sh(\mathcal{F}) = [n]_0 \cup [n]_1 \cup \cdots \cup [n]_b,$$

therefore $\mathcal{F}$ is not extremal. It is straightforward to see that $\mathcal{F}$ is isometrically embedded in Q_n. ■

Let us recall the definition of weak extremality. In all of the above mentioned examples $\mathcal{F}$ is isometrically embedded in Q_n, so according to Proposition 7.1, $Sh(\mathcal{C}) = Sh(D_i(\mathcal{C}))$ for every chunk $\mathcal{C}$ and every $i \in [n]$. Thus for $\mathcal{F}$ itself $Sh(\mathcal{F}) = Sh(D_i(\mathcal{F}))$, which means that $\mathcal{F}$ is weakly extremal. But in none of the examples is $\mathcal{F}$ extremal. So all of these example show that weak extremality is really weaker then extremality.

8 Some remarks on the VC dimension

The Vapnik-Chervonenkis (VC) dimension is a widely known and used notion in mathematics, with applications among others in machine learning [25]-[27], probability theory [24] and combinatorics [23].

Definition 8.1 *The Vapnik-Chervonenkis dimension of a set system* $\mathcal{F} \subseteq 2^{[n]}$*, denoted by* $VC-dim(\mathcal{F})$*, is the maximum cardinality of a set shattered by* $\mathcal{F}$*.*

As an example of its application, consider [22]. This says that the problem of computing the VC-dimension is in $SAT_{log^2 n}$, the class of algorithmic problems which are polynomial-time reducible to the satisfiability problem of a boolean formula of length J with $O(log^2 J)$ variables, and hard in $SAT^{CNF}_{log^2 n}$ (as $SAT_{log^2 n}$, only with inputs in conjunctive normal form). This section is about the problem of computing $VC-dim(\mathcal{F})$ for a set system $\mathcal{F} \subseteq 2^{[n]}$.

Proposition 8.1 *(See [22].) For any set system* $\mathcal{F} \subseteq 2^{[n]}$*,* $VC-dim(\mathcal{F}) \leq \log|\mathcal{F}|$*. (Here* log *stands for the logarithm with base* 2*.)*

Proof: If the set $S \subseteq [n]$ is shattered by $\mathcal{F}$, then $2^S = \{F \cap S : F \in \mathcal{F}\}$. This can only hold if there are at least $2^{|S|}$ sets in $\mathcal{F}$. Thus for any set S shattered by $\mathcal{F}$, we have $|S| \leq \log|\mathcal{F}|$. ■

By this proposition, the simple algorithm for computing the VC-dimension of a set system $\mathcal{F}$ which enumerates all possible sets to be shattered, shall terminate in $mn^{O(log(m))}$ time, where m is the size of $\mathcal{F}$ (see [15]). We give another algorithm with the same time bound. First let us recall that

$$Sh(\mathcal{F}) = \bigcup_{lex\ orders} Sm(I(\mathcal{F})).$$

From the proof of Proposition 4.6 we know that if we take a set S from $Sh(\mathcal{F})$ then for the term order where the variables from x_S are the smallest we have $S \in Sm(\mathcal{F})$ (i.e. $x_S \in Sm(I(\mathcal{F}))$). Thus in the above sum it

suffices to sum up over a family of lexicographic orders where for every possible set S there is a suitable term order. According to Proposition 8.1, the number of possible sets S, i.e. the number of sets with size at most $\log m$, is $O(n^{\log m})$. Hence to get $Sh(F)$, and consequently $VC-dim(\mathcal{F})$, it is enough to compute $Sm(I(\mathcal{F}))$ for $O(n^{\log m})$ term orders. The computation of the standard monomials for a particular term order can be done in $O(nm)$ time, so altogether the time bound that we get is $O(mn^{\log m})$.

9 Generalization of shattering

In the previous sections we have presented our results concerning shattered sets, extremal set systems and standard monomials of vanishing ideals. Some of them can be generalized to the case of an arbitrary $k > 0$.

There is a usual way of generalizing the notion of shattering (See e.g. [22].):

Definition 9.1 *Let $\mathcal{F}$ be a class of $[n] \to \{0, 1, ..., k-1\}$ functions. We say that $\mathcal{F}$ shatters a set $S \subseteq [n]$ if for every function $\mathbf{g} : S \to \{0, 1, ..., k-1\}$ there exists a function $\mathbf{f} \in \mathcal{F}$ such that $\mathbf{f}|_S = \mathbf{g}$.*

We can look at the elements of $\mathcal{F}$ as vectors from $\{0, 1, \ldots, k-1\}^n$, thus both $I(\mathcal{F})$ and $Sm(I(\mathcal{F}))$ are well defined. When considering this definition we can observe that our basic inequality, Proposition 4.1 can not be generalized, because the size of $\mathcal{F}$ can be much greater than that of $Sh(\mathcal{F})$. However it is easy to see that the same recursive property holds for shattered sets like earlier, namely that

$$(\bigcup_{\beta} Sh(\mathcal{F}_\beta)) \bigcup \{S \cup \{n\} \ : \ S \in \bigcap_{\beta} Sh(\mathcal{F}_\beta)\} \subseteq Sh(\mathcal{F}).$$

Since Proposition 4.1 does not hold in the general case, we do not define extremality in the usual way, but using Corollary 4.6.1.

Definition 9.2 *$\mathcal{F} \subseteq \{0, 1, \ldots, k-1\}^n$ is said to be extremal if $Sm(I(\mathcal{F}))$ is the same for all lexicographic term orders.*

For a subset $H \subseteq [n]$ we write x_H for the monomial $\prod_{i \in H} x_i^{k-1}$. The polynomial $f_{S,g}$ is defined for a subset $S \subseteq [n]$ and function (vector) $\mathbf{g} : S \to \{0, 1, \ldots, k-1\}$ as

$$f_{S,g} = \prod_{i \in S} \frac{x_i(x_i - 1) \ldots (x_i - (k-1))}{x_i - g_i}.$$

With this notation now we look at the analogues of Propositions 4.7, 4.5 and 4.6. Their proofs are similar to the case $k = 2$.

Proposition 9.1 *If $S \notin Sh(\mathcal{F})$, then there exists a function (vector) $\mathbf{g}$: $S \to \{0, 1, \ldots, k-1\}$ such that $f_{S,g}(\mathbf{v}) = 0$, $\forall\ \mathbf{v} \in \mathcal{F}$, i.e. $f_{S,g} \in I(\mathcal{F})$.*

Proof: Since S is not shattered by $\mathcal{F}$, there is a function $\mathbf{g} : S \to \{0, 1, ..., k-1\}$ for which there is no function $\mathbf{v} \in \mathcal{F}$ such that $\mathbf{v}|_S = \mathbf{g}$. Consider the corresponding $f_{S,g}$ polynomial. For this we have $f_{S,g}(\mathbf{v}) \neq 0$ if and only if $vev|_S = \mathbf{g}$. According to our assumption there is no such $\mathbf{v} \in \mathcal{F}$, that is $f_{S,g}(\mathbf{v}) = 0\ \forall\ \mathbf{v} \in \mathcal{F}$, i.e. $f_{S,g} \in I(\mathcal{F})$. ■

Proposition 9.2 *If $x_H \in Sm(I(\mathcal{F}))$ for some term order, then $H \in Sh(\mathcal{F})$.*

Proof: Suppose the contrary, namely that $H \notin Sh(\mathcal{F})$. According to Proposition 9.1 there is a polynomial $f_{S,g} \in I(\mathcal{F})$. The leading monomial of this polynomial is x_S for all term orders, thus $x_H \in Lm(I(\mathcal{F}))$ which is a contradiction. ■

This means that $(Sm(I(\mathcal{F})) \cap \{x_H, H \subseteq [n]\}) \subseteq Sh(\mathcal{F})$ for every lexicographic term order.

Proposition 9.3 *If $H \in Sh(\mathcal{F})$, then there is a lexicographic term order for which we have $x_H \in Sm(I(\mathcal{F}))$.*

Proof: We begin our proof just like in the case $k = 2$. Take the lexicographic order where the variables of x_H are the smallest. To prove that $x_H \in Sm(I(\mathcal{F}))$ we will use the recursive property and the prefix tree constructed at the end of subsection 3.1. From the fact that $\mathcal{F}$ shatters H we can conclude that the first $|H|$ levels of this trie are complete, i.e. each vertex above the $|H|^{th}$ level has exactly k children. This, together with the recursive property imply that $x_H \in Sm(I(\mathcal{F}))$. ■

These two statements together imply that

$$Sh(\mathcal{F}) = \bigcup_{lex\ orders} (Sm(I(\mathcal{F})) \cap \{x_H, H \subseteq [n]\}).$$

In [17] Alon proved the following interesting statement:

Proposition 9.4 *For every $\mathcal{F} \subseteq \{0,1,\ldots,k-1\}^n$ there exists an up-set $\mathcal{G} \subseteq \{0,1,\ldots,k-1\}^n$ such that*
i) $|\mathcal{G}| = |\mathcal{F}|$, *and*
ii) $|\{g|_I \ : \ g \in \mathcal{G}\}| \leq |\{f|_I \ : \ f \in \mathcal{F}\}|$ *for all* $I \subseteq [n]$.

Proof: See [17]. ■

An immediate corollary of this proposition is that for every $\mathcal{F} \subseteq \{0,1,\ldots,k-1\}^n$ there exists an up-set $\mathcal{G} \subseteq \{0,1,\ldots,k-1\}^n$ such that $|\mathcal{G}| = |\mathcal{F}|$ and $Sh(\mathcal{G}) \subseteq Sh(\mathcal{F})$.

Before getting started with the generalization of downshifts in the next section, we first mention the generalization of bit flips. Let $\chi : \{0,1,\ldots,k-1\} \to \{0,1,\ldots,k-1\}$ be a fixed bijective function. For a vector $\mathbf{v} \in \{0,1,\ldots,k-1\}^n$ we denote by φ_i^χ the ith bit flip:

$$\varphi_i^\chi(\mathbf{v}) = (v_1,\ldots,v_{i-1},\chi(v_i),v_{i+1},\ldots,v_n).$$

Using Proposition 3.3 it is clear that extremality is henceforward invariant to this operation.

10 Generalization of downshifts

Downshifts can be generalized as well (See e.g. [5].), but for this we need first some new notations. We write the standard basis of $\{0,1,\ldots,k-1\}^n$ as $\mathbf{e_1},\ldots,\mathbf{e_n}$. For $J\subseteq[n]$, the span of $\{\mathbf{e_j}:j\in J\}$ in $\{0,1,\ldots,k-1\}^n$ is denoted by $\{0,1,\ldots,k-1\}^J$. The complement of J in $[n]$ is written $\widehat{J}$, specially for the complement of $\{j\}$ we write $\widehat{j}$. The J-section of $\mathcal{F}\subseteq\{0,1,\ldots,k-1\}^n$ for $\mathbf{z}\in\{0,1,\ldots,k-1\}^{\widehat{J}}$ is the set

$$\mathcal{F}_J(\mathbf{z})=\{\mathbf{f}|_J:\mathbf{f}\in\mathcal{F}\ and\ \mathbf{f}|_{\widehat{J}}=\mathbf{z}\},$$

that is $\mathcal{F}_J(\mathbf{z})$ is the collection of vectors in $\mathcal{F}$ that are equal with $\mathbf{z}$ outside J.

The downshift by the element i, denoted by D_i, is defined using i-sections:

Definition 10.1 *For $\mathcal{F}\subseteq\{0,1,\ldots,k-1\}^n$ $D_i(\mathcal{F})$ is the set for which*

$$(D_i(\mathcal{F}))_i(\mathbf{z})=\{0,1,\ldots,|\mathcal{F}_i(\mathbf{z})|-1\}$$

for every $\mathbf{z}\in\{0,1,\ldots,k-1\}^{\widehat{i}}$.

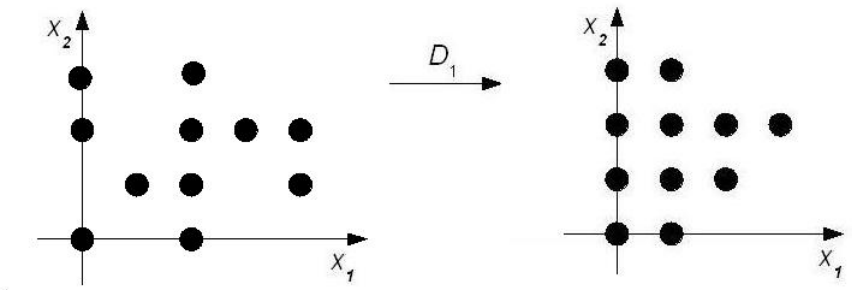

Figure 1: The action of D_1

In other words, in each i-section $D_i(\mathcal{F})$ has the same number of points as $\mathcal{F}$, but they are downshifted to the right of the plane $x_i=0$. Clearly $|D_i(\mathcal{F})|=|\mathcal{F}|$ and $\mathbf{v}\in D_i(\mathcal{F})$ if and only if there are v_i+1 elements

$s_0, \ldots, s_{v_i} \in \{0, 1, \ldots, k-1\}$ such that $(v_1, \ldots, v_{i-1}, s_j, v_{i+1,\ldots,v_n}) \in \mathcal{F}$ for $j = 0, 1 \ldots, v_i$. Figure 10 shows a two-dimensional example.

In the following we present the properties of downshifts in the general case.

Proposition 10.1 *For every $i \in [n]$ we have $Sh(D_i(\mathcal{F})) \subseteq Sh(\mathcal{F})$.*

Proof: Let $S \in Sh(D_i(\mathcal{F}))$. If $i \notin S$, then clearly $S \in Sh(\mathcal{F})$ holds as well. Now suppose that $i \in S$. Let us build up the prefix trees of $\mathcal{F}$ and $D_i(\mathcal{F})$ with the first $|S|$ levels corresponding to the coordinates in S. Call them respectively T_1 and T_2. Since $D_i(\mathcal{F})$ shatters S, in T_2 the first $|S|$ levels must be complete, i.e. every vertex on a level smaller than $|S|$ has exactly k children. According to the definition of D_i, this can happen only if the first $|S|$ levels are complete in T_1 already, thus S is already shattered by $\mathcal{F}$. This finishes our proof. ■

Definition 10.2 *$\mathcal{F} \subseteq \{0, 1, \ldots, k-1\}^n$ is an i-down-set if for every $\mathbf{f} \in \mathcal{F}$ we have $(f_1, \ldots, f_{i-1}, a, f_{i+1}, \ldots, f_n) \in \mathcal{F}$ whenever $a < f_i$.*

Clearly, in the general case, i-down sets have the same basic properties, i.e. if $\mathcal{F}$ is an i-down-set, then $D_i(\mathcal{F}) = \mathcal{F}$ and $D_i(\mathcal{F})$ is always an i-down-set. This also means that the operator D_i is idempotent.

Proposition 10.2 *If $\mathcal{F} \subseteq \{0, 1, \ldots, k-1\}^n$ is an i-down-set, then $D_j(\mathcal{F})$ is an i-down-set as well, for every $j \neq i$.*

Proof: Without loss of generality we can suppose that $i, j = 1, 2$. We have to prove that $D_2(\mathcal{F})$ is a 1-down-set, thus if we take an arbitrary vector $\mathbf{v}$ from $D_2(\mathcal{F})$, then $(c, v_2, v_3, \ldots, v_n) \in D_2(\mathcal{F})$ must hold for all $c < v_1$. Since $\mathbf{v} \in D_2(\mathcal{F})$, there are $v_2 + 1$ elements $s_0, \ldots, s_{v_2} \in \{0, 1, \ldots, k-1\}$ such that $(v_1, s_j, v_3, \ldots, v_n) \in \mathcal{F}$ for $j = 0, 1 \ldots, v_2$. Since $\mathcal{F}$ is an 1-down-set, if we take $c < v_1$, then $(c, s_j, v_3, \ldots, v_n) \in \mathcal{F}$ for $j = 0, 1 \ldots, v_2$, what implies that $(c, v_2, v_3, \ldots, v_n) \in D_2(\mathcal{F})$. ■

Proposition 10.3 *Let $\mathcal{F} \subseteq \{0, 1, \ldots, k-1\}^n$. Then n different downshifts applied to $\mathcal{F}$ in an arbitrary order result in a down-set.*

This is an immediate consequence of the last statement just like in the case $k = 2$.

Theorem 10.1 *Let $\mathcal{F} \subseteq \{0, 1, \ldots, k-1\}^n$ and $\prec$ be a lexicographic term order for which $x_{i_1} \succ x_{i_2} \succ \cdots \succ x_{i_n}$. If we apply the downshifts $D_{i_1}, D_{i_2}, \ldots, D_{i_n}$ to $\mathcal{F}$ in this order, then we have $Sm(I(\mathcal{F})) = \{\mathbf{x}^{\mathbf{w}} \ : \ \mathbf{w} \in D_{i_n, i_{n-1}, \ldots, i_1}(\mathcal{F})\}$. ($Sm(I(\mathcal{F}))$" = "$D_{i_n, i_{n-1}, \ldots, i_1}(\mathcal{F})$)*

Proof: Without loss of generality we can suppose that $\prec$ is based on the natural lex order, thus $i_k = k$. We apply induction on n. For $n = 1$, $x^l \in Sm(I(\mathcal{F}))$ and $l \in D_1(\mathcal{F})$ hold at the same time, namely when $l < |\mathcal{F}|$, thus for $n = 1$ the statement holds. Now let us suppose that the statement is true for all values smaller than n, and consider the case of $n > 1$. A downshift D_i for $i \neq n$ acts on the subsystems of $\mathcal{F}$ in which the n^{th} coordinates are equal, i.e.

$$D_{n-1,\ldots,1}(\mathcal{F}) = \bigcup_{\beta=0}^{k-1} \{(\mathbf{w}, \beta) \ : \ \mathbf{w} \in D_{n-1,\ldots,1}(\mathcal{F}_\beta)\}$$

From our induction hypothesis $Sm(I(\mathcal{F}_\beta))$" = "$D_{n-1,\ldots,1}(\mathcal{F}_\beta)$. However when constructing $Sm(I(\mathcal{F}))$ from the $Sm(I(\mathcal{F}_\beta))$-s there is the same rule for the exponent vectors as for the elements of $D_{n,n-1,\ldots,1}(\mathcal{F})$ when constructing it from the $D_{n-1,\ldots,1}(\mathcal{F}_\beta)$-s. $\mathbf{x}^{(\mathbf{w},l)} \in Sm(I(\mathcal{F}))$ exactly in the same case when $(\mathbf{w}, l) \in D_{n,n-1,\ldots,1}(\mathcal{F})$, namely when there are at least $l + 1$ β-s such that $\mathbf{x}^{\mathbf{w}} \in Sm(I(\mathcal{F}_\beta))$ ($\mathbf{w} \in D_{n-1,\ldots,1}(\mathcal{F}_\beta)$). This finishes the proof. ■

As a corollary of Theorem 10.1 we investigate now how the downshift operation affects the standard monomials.

Corollary 10.3.1 *Let $\mathcal{F} \subseteq \{0, 1, \ldots, k-1\}^n$ and $\prec$ be a lexicographic term order for which $x_{i_1} \succ x_{i_2} \succ \cdots \succ x_{i_n}$.. Suppose that $\mathcal{F}$ is a i_j-down-set for $1 \leq j < k$. With this we have*

$$Sm(I(D_{i_k}(\mathcal{F}))) = Sm(I(\mathcal{F})).$$

Proof: As previously, without loss of generality we can suppose that $i_j = j$. According to Theorem 10.1

$$Sm(I(\mathcal{F})) = D_{n,n-1,\ldots,1}(\mathcal{F})$$

and

$$Sm(I(D_k(\mathcal{F}))) = D_{n,n-1,\ldots,1}(D_k(\mathcal{F})).$$

According to our assumption $\mathcal{F}$ is a j-down-set for $1 \leq j < k$, however, using Proposition 10.2 we get that the same holds for $D_k(\mathcal{F})$ as well. Since D_i has no effect on an i-down-set, we conclude that

$$Sm(I(\mathcal{F})) = D_{n,n-1,\ldots,k}(\mathcal{F})$$

and

$$Sm(I(D_k(\mathcal{F}))) = D_{n,n-1,\ldots,k}(D_k(\mathcal{F})).$$

The fact that D_k is idempotent finishes the proof. ■

11 Conclusion and Future work

The aim of this study was to demonstrate that algebraic methods can be very useful when studying combinatorial objects. We have presented basic definitions and statements concerning Gröbner bases, standard monomials, shattering and set system operations. We achieved several results in all of these areas. To start with, in Section 4 we gave a new characterization for extremal set systems using Gröbner bases. At the end of the same section we proposed an efficient, $O(n^2m)$, algorithm for testing extremality using the standard monomials of a set system. We also analyzed the connection between the effect of set system operations and extremality, and proposed a new method for constructing the family of standard monomials using the downshift operation. In Section 7 we discussed the graph theoretical consequences of extremality, characterized the event when downshifting makes $Sh(\mathcal{F})$ shrink, and made some remarks on the work of Greco in [6]. Then in Section 8 we proposed a new algorithm for solving the problem of computing the Vapnik-Chervonenkis dimension of a set system $\mathcal{F} \subseteq 2^{[n]}$. Section 9 and 10 were given over to the generalizations of some of our results.

There are many ways to follow up this study. These include finding (slightly) faster (possibly linear time) algorithms for testing extremality, improving the $O(mn^{\log m})$ time bound of the algorithm for computing the VC-dimension of a set system or extend more of our results to the general case.

Acknowledgments

I thank my advisor Lajos Rónyai for proposing this problem to work on, and for all the guidance and support throughout. I also thank Miklós Rácz for the help given in connection with language issues.

References

[1] R.P. Anstee, L. Rónyai, A. Sali, Shattering News, Graphs and Combinatorics, Vol.18, 59-73 (2002)

[2] B. Felszeghy, B. Ráth, L. Rónyai, The lex game and some applications, Journal of Symbolic Computation, Vol. 41, 663-681 (2006)

[3] B. Bollobás, A.J. Radcliffe, Defect Sauer Results, Journal of Combinatorial Theory Series A, Vol. 72, 189-208 (1995)

[4] P. Frankl, Extremal set systems, Handbook of combinatorics (vol. 2), MIT Press, Cambridge, MA, 1996

[5] B. Bollobás, I. Leader, A.J. Radcliffe, Reverse Kleitman Inequalities, Proceedings of the London Mathematical Society, Vol. s3-58, 153-168 (1989)

[6] G. Greco, Embeddings and trace of finite sets, Information Processing Letters, Vol. 67, 199-203 (1998)

[7] W. W. Adams, P. Loustaunau, An Introduction to Gröbner bases, Graduate Studies in Mathematics, Vol. 3, American Mathematical Society (1994)

[8] B. Felszeghy, Bevezetés a Gröbner-bázisok elméletébe, http://www.math.bme.hu/ fbalint/publ/grobnerjegyzet.pdf

[9] B. Buchberger, H. M. Möller, Gröbner Bases and Applications, London Mathematical Society Series, Vol. 251 (1998), Proc of the international conference "33 Years of Gröbner Bases"

[10] R. Aharoni, R. Holzman, Personal communication

[11] A. Pajor, Sous-spaces 1: des Espaces de Banach, Travaux en Cours, Hermann, Paris, (1985)

[12] N. Sauer, On the Density of Families of Sets, Journal of Combinatorial Theory, Series A, Vol. 13, 145-147 (1972)

[13] S. Shelah, A Combinatorial Problem: Stability and Order for Models and Theories in Infinitary Language, Pacific Journal of Mathematics, Vol. 41, 247-261 (1972)

[14] V. N. Vapnik, A. Ya. Chervonenkis, On the Uniform Convergence of Relative Frequencies of Events to their Probabilities, Theory of Probability and its Applications, Vol. 16, 264-280 (1971)

[15] N.Linial, Y. Mansour, R.L. Rivest, Results on learnability and the Vapnik-Chervonenkis dimension, Proceedings of the 29th IEEE Symposium on Foundations of Computer Science (White Plains, N.Y., Oct.). IEEE, New York, 1988, pp. 120-129.

[16] K. Friedl, L. Rónyai, Order shattering and Wilson's theorem, Discrete Mathematics, Vol. 270, 127-136 (2003)

[17] N. Alon, On the density of sets of vectors, Discrete Mathematics, Vol. 46, 199-202 (1983)

[18] P. Frankl, J. Pach, On disjointly representable sets, Combinatorica, Vol. 4, 39-45 (1994)

[19] T. H. Cormen, C. E. Leiserson, R. L. Rivest, C. Stein, Introduction to Algorithms, Second edition, The MIT Press (2001)

[20] L. Rónyai, G. Ivanyos, R. Szabó, Algoritmusok, Typotex könyvkiadó (1999)

[21] L. Babai, P. Frankl, Linear Algebra Methods in Combinatorics, Department of Computer Science, University of Chicago, Preliminary version 2 (1992)

[22] A. Shinohara, Complexity of computing Vapnik-Chervonekis dimension and some generalized dimensions, Theoretical Computer Science, Vol. 137, 129-144 (1995)

[23] R. P. Anstee, A. Sali, Sperner families of bounded VC-dimension, Discrete Mathematics, Vol. 175, 13-21 (1997)

[24] M. Talagrand, Vapnik-Chervonenkis type conditions and Donsker classes of functions, Annals of Probability, Vol. 31, 1565-1582 (2003)

[25] S. Floyd, M. K. Warmuth, Sample compression, earnability and the Vapnik-Chervonenkis dimension, Machine Learning, Vol. 21, 269-304 (2005)

[26] A. Blumer, A. Ehrenfeucht, D. Haussler, M. K. Warmuth, Learnability and the Vapnik-Chervonenkis dimension, Journal of the ACM, Vol. 36, 929-965 (1989)

[27] B. K. Natarajan, Machine Learning: a theoretical approach, Morgan Kaufmann (1991)