Some Combinatorial Applications of Gröbner Bases

Lajos Rónyai[1][2]* and Tamás Mészáros[3]

[1] Computer and Automation Research Institute, Hungarian Academy of Sciences
[2] Institute of Mathematics, Budapest University of Technology and Economics
lajos@ilab.sztaki.hu
[3] Department of Mathematics, Central European University
Meszaros_Tamas@ceu_budapest.edu

Abstract. Let $\mathbb{F}$ be a field, $V \subseteq \mathbb{F}^n$ be a (combinatorially interesting) finite set of points. Several important properties of V are reflected by the polynomial functions on V. To study these, one often considers $I(V)$, the vanishing ideal of V in the polynomial ring $\mathbb{F}[x_1, \ldots, x_n]$. Gröbner bases and standard monomials of $I(V)$ appear to be useful in this context, leading to structural results on V.
Here we survey some work of this type. At the end of the paper a new application of this kind is presented: an algebraic characterization of shattering-extremal families and a fast algorithm to recognize them.

Keywords: Gröbner basis, standard monomial, lexicographic order, vanishing ideal, Hilbert function, inclusion matrix, rank formula, combinatorial Nullstellensatz, S-extremal set family.

1 Introduction

Throughout the paper n will be a positive integer, and $[n]$ stands for the set $\{1, 2, \ldots, n\}$. The set of all subsets of $[n]$ is denoted by $2^{[n]}$. Subsets of $2^{[n]}$ are called *set families* or *set systems*. Let $\binom{[n]}{m}$ denote the family of all *m-subsets* of $[n]$ (subsets which have cardinality m), and $\binom{[n]}{\leq m}$ is the family of those subsets that have at most m elements. $\mathbb{N}$ denotes the set of the nonnegative integers, $\mathbb{Z}$ is the set of integers, $\mathbb{Q}$ is the field of rational numbers, and $\mathbb{F}_p$ is the field of p elements, where p is a prime.

Let $\mathbb{F}$ be a field. As usual, we denote by $\mathbb{F}[x_1, \ldots, x_n] = \mathbb{F}[\mathbf{x}]$ the ring of polynomials in variables $x_1, \ldots, x_n$ over $\mathbb{F}$. To shorten our notation, we write $f(\mathbf{x})$ for $f(x_1, \ldots, x_n)$. Vectors of length n are denoted by boldface letters, for example $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{F}^n$. If $\mathbf{w} \in \mathbb{N}^n$, we write $\mathbf{x}^{\mathbf{w}}$ for $x_1^{w_1} \ldots x_n^{w_n} \in \mathbb{F}[\mathbf{x}]$. For a subset $M \subseteq [n]$, the monomial x_M is $\prod_{i \in M} x_i$ (and $x_\emptyset = 1$).

Suppose that $V \subseteq \mathbb{F}^n$. Then the *vanishing ideal* $I(V)$ *of* V consists of the polynomials in $\mathbb{F}[\mathbf{x}]$, which, as functions, vanish on V. In our applications, we

* Research supported in part by OTKA grants NK72845, K 77476 and K 77778.

consider finite sets V, and use the Gröbner bases, and standard monomials of $I(V)$ (see the next subsection for the definitions) to prove claims on V.

Let $\mathbf{v}_F \in \{0,1\}^n$ denote the *characteristic vector of a set* $F \subseteq [n]$, that is, the ith coordinate of $\mathbf{v}_F$ is 1 iff $i \in F$. For a system of sets $\mathcal{F} \subseteq 2^{[n]}$, let us put $V_{\mathcal{F}}$ for the set of the characteristic vectors of elements of $\mathcal{F}$. By $I(\mathcal{F})$ we understand the vanishing ideal $I(V_{\mathcal{F}})$, as it will make no confusion.

In Sect. 2 we collected some basic facts about Gröbner bases and related notions, such as standard monomials, reduction and Hilbert functions. Section 3 is devoted to the complete uniform families and their extensions. Here we discuss results describing the Gröbner bases and standard monomials of the ideals $I(\mathcal{F})$, where $\mathcal{F}$ is a complete uniform family $\binom{[n]}{d}$ for some $0 \leq d \leq n$. We outline some combinatorial applications of these results, including an extension of Wilson's rank formula. Generalizations and extensions are also considered. Section 4 gives a brief explanation of the lex game method, which gives a powerful technique to determine lex standard monomials both in theory and practice. In Sect. 5 we consider ideals and Gröbner bases attached to more complex objects, such as partitions, permutations and graph colorings. The latter topic is particularly rich in results involving polynomial ideals. Section 6 briefly introduces a powerful algebraic technique of combinatorics, the combinatorial Nullstellensatz by Noga Alon, together with the resulting non-vanishing theorem. The last section gives an algebraic characterization of shattering-extremal set families. The characterization involves Gröbner bases and, together with the lex game method, it provides an efficient algorithm for recognizing shattering-extremal families.

Ideals I of $\mathbb{F}[\mathbf{x}]$ generated by monomials are perhaps the most important objects in algebraic combinatorics. Their study, initiated by Stanley, has led to some spectacular results, in particular, in the area of simplicial complexes and convex polytopes. Gröbner basis methods are also applicable there. In this paper we avoid the area of monomial ideals, as there are many excellent treatments of this subject. We refer the interested reader to the recent volume of Herzog and Hibi [30], and the sources cited therein.

2 Gröbner Bases, Standard Monomials and Hilbert Functions

We recall now some basic facts concerning Gröbner bases in polynomial rings over fields. For details we refer to [11], [12], [13], [14], and [2].

A total order $\prec$ on the monomials composed from variables $x_1, x_2, \ldots, x_n$ is a *term order*, if 1 is the minimal element of $\prec$, and $\prec$ is compatible with multiplication with monomials (if m_1, m_2, m_3 are monomials, $m_1 \prec m_2$, then $m_1 m_3 \prec m_2 m_3$). Two important term orders are the *lexicographic* (*lex* for short) and the *degree compatible lexicographic* (*deglex*) orders. We have $\mathbf{x}^{\mathbf{w}} \prec_{\text{lex}} \mathbf{x}^{\mathbf{u}}$ if and only if $w_i < u_i$ holds for the smallest index i such that $w_i \neq u_i$. As for deglex, we have that a monomial of smaller degree is smaller in deglex, and among monomials of the same degree lex decides the order. Also in general, $\prec$ is *degree compatible*, if $\deg(\mathbf{x}^{\mathbf{w}}) < \deg(\mathbf{x}^{\mathbf{u}})$ implies $\mathbf{x}^{\mathbf{w}} \prec \mathbf{x}^{\mathbf{u}}$.

The *leading monomial* $\mathrm{lm}(f)$ of a nonzero polynomial $f \in \mathbb{F}[\mathbf{x}]$ is the largest monomial (with respect to $\prec$) which appears with nonzero coefficient in f, when written as the usual linear combination of monomials. We denote the set of all leading monomials of polynomials of a given ideal $I \trianglelefteq \mathbb{F}[\mathbf{x}]$ by $\mathrm{Lm}(I) = \{\mathrm{lm}(f) : f \in I\}$, and we simply call them the *leading monomials of* I.

A monomial is called a *standard monomial* of I, if it is not a leading monomial of any $f \in I$. Let $\mathrm{Sm}(I)$ denote the set of standard monomials of I. Obviously, a divisor of a standard monomial is again in $\mathrm{Sm}(I)$.

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 $\mathrm{lm}(g)$ divides $\mathrm{lm}(f)$. It is not hard to verify that G is actually a basis of I, that is, G generates I as an ideal of $\mathbb{F}[\mathbf{x}]$. It is a fundamental fact that every nonzero ideal I of $\mathbb{F}[\mathbf{x}]$ has a Gröbner basis.

A Gröbner basis $G \subseteq I$ is *reduced*, if for all $g \in G$, the *leading coefficient* of g (i.e. the coefficient of $\mathrm{lm}(g)$) is 1, and $g \neq h \in G$ implies that no nonzero monomial in g is divisible by $\mathrm{lm}(h)$. For any fixed term order and any nonzero ideal of $\mathbb{F}[\mathbf{x}]$ there exists a unique reduced Gröbner basis. A Gröbner basis is *universal*, if it is a Gröbner basis for every term order $\prec$ on the monomials.

Suppose that $f \in \mathbb{F}[\mathbf{x}]$ contains a monomial $\mathbf{x}^{\mathbf{w}} \cdot \mathrm{lm}(g)$, where g is some other polynomial with leading coefficient c. Then we can *reduce* f *with* g (and obtain $\hat{f}$), that is, we can replace $\mathbf{x}^{\mathbf{w}} \cdot \mathrm{lm}(g)$ in f with $\mathbf{x}^{\mathbf{w}} \cdot \left(\mathrm{lm}(g) - \frac{1}{c}g\right)$. Clearly if $g \in I$, then f and $\hat{f}$ represent the same coset in $\mathbb{F}[\mathbf{x}]/I$. Also note that $\mathrm{lm}\left(\mathbf{x}^{\mathbf{w}} \cdot \left(\mathrm{lm}(g) - \frac{1}{c}g\right)\right) \prec \mathbf{x}^{\mathbf{w}} \cdot \mathrm{lm}(g)$. As $\prec$ is a well founded order, this guarantees that if we reduce f repeatedly with a set of polynomials G, then we end up with a *reduced* $\hat{f}$ in finitely many steps, that is a polynomial such that none of its monomials is divisible by any $\mathrm{lm}(g)$ $(g \in G)$.

If G is a Gröbner basis of an ideal I, then it can be shown that the reduction of any polynomial with G is unique. It follows that for a nonzero ideal I the set $\mathrm{Sm}(I)$ is a linear basis of the $\mathbb{F}$-vector space $\mathbb{F}[\mathbf{x}]/I$. If $I(V)$ is a vanishing ideal of a finite set V of points in $\mathbb{F}^n$, then $\mathbb{F}[\mathbf{x}]/I(V)$ can be interpreted as the space of functions $V \to \mathbb{F}$. An immediate consequence is that the number of standard monomials of $I(V)$ is $|V|$. In particular, for every family of sets we have $|\mathcal{F}| = |\mathrm{Sm}(I(\mathcal{F}))|$.

Another property of the standard monomials of $I(\mathcal{F})$ we will meet several times: for an arbitrary set family $\mathcal{F}$, one has $x_i^2 - x_i \in I(\mathcal{F})$, therefore all the elements of $\mathrm{Sm}(I(\mathcal{F}))$ are square-free monomials.

We write $\mathbb{F}[\mathbf{x}]_{\leq m}$ for the vector space of polynomials over $\mathbb{F}$ with degree at most m. Similarly, if $I \trianglelefteq \mathbb{F}[\mathbf{x}]$ is an ideal then $I_{\leq m} = I \cap \mathbb{F}[\mathbf{x}]_{\leq m}$ is the linear subspace of polynomials from I with degree at most m. The *Hilbert function* of the $\mathbb{F}$-algebra $\mathbb{F}[\mathbf{x}]/I$ is $H_I : \mathbb{N} \to \mathbb{N}$, where

$$H_I(m) = \dim_{\mathbb{F}}\left(\mathbb{F}[\mathbf{x}]_{\leq m}/I_{\leq m}\right).$$

Let $\prec$ be any degree compatible term order (deglex for instance). One can easily see that the set of standard monomials with respect to $\prec$ of degree at most m forms a linear basis of $\mathbb{F}[\mathbf{x}]_{\leq m}/I_{\leq m}$. Hence we can obtain $H_I(m)$ by determining the set $\mathrm{Sm}(I)$ with respect to any degree compatible term ordering.

In the combinatorial literature $H_{I(\mathcal{F})}(m)$ is usually given in terms of inclusion matrices. For two families $\mathcal{F},\mathcal{G}\subseteq 2^{[n]}$ the *inclusion matrix* $I(\mathcal{F},\mathcal{G})$ is a matrix of size $|\mathcal{F}|\times|\mathcal{G}|$, whose rows and columns are indexed by the elements of $\mathcal{F}$ and $\mathcal{G}$, respectively. The entry at position (F,G) is 1, if $G\subseteq F$, and 0 otherwise $(F\in\mathcal{F},G\in\mathcal{G})$. It is a simple matter to verify that the Hilbert function of $\mathcal{F}$ is given by

$$H_{I(\mathcal{F})}(m)=\dim_{\mathbb{F}}\left(\mathbb{F}\left[\mathbf{x}\right]_{\leq m}/I(\mathcal{F})_{\leq m}\right)=\mathrm{rank}_{\mathbb{F}}I\left(\mathcal{F},\binom{[n]}{\leq m}\right).$$

3 Complete Uniform Families, Applications and Extensions

3.1 Gröbner Bases and Standard Monomials for Complete Uniform Families

We start here with an explicit description of the (reduced) Gröbner bases for the ideals $I_{n,d}:=I(\mathcal{F})$, where $\mathcal{F}=\binom{[n]}{d}$ for some integer $0\leq d\leq n$. That is, we consider the vanishing ideal of the set of all 0,1-vectors in $\mathbb{F}^n$ whose Hamming weight is d.

Let t be an integer, $0<t\leq n/2$. We set $\mathcal{H}_t$ as the set of those subsets $\{s_1<s_2<\cdots<s_t\}$ of $[n]$ for which t is the smallest index j with $s_j<2j$.

We have $\mathcal{H}_1=\{\{1\}\}$, $\mathcal{H}_2=\{\{2,3\}\}$, and $\mathcal{H}_3=\{\{2,4,5\},\{3,4,5\}\}$. It is clear that if $\{s_1<\ldots<s_t\}\in\mathcal{H}_t$, then $s_t=2t-1$, and $s_{t-1}=2t-2$ if $t>1$.

For a subset $J\subseteq[n]$ and an integer $0\leq i\leq|J|$ we denote by $\sigma_{J,i}$ the i-th elementary symmetric polynomial of the variables x_j, $j\in J$:

$$\sigma_{J,i}:=\sum_{T\subseteq J,|T|=i}x_T\ \in\mathbb{F}[x_1,\ldots,x_n].$$

In particular, $\sigma_{J,0}=1$.

Now let $0<t\leq n/2$, $0\leq d\leq n$ and $H\in\mathcal{H}_t$. Set

$$H'=H\cup\{2t,2t+1,\ldots,n\}\subseteq[n].$$

We write

$$f_{H,d}=f_{H,d}(x_1,\ldots,x_n):=\sum_{k=0}^{t}(-1)^{t-k}\binom{d-k}{t-k}\sigma_{H',k}.$$

As an example, with $U=\{2,3,\ldots,n\}$ we have

$$f_{\{2,3\},d}=\sigma_{U,2}-(d-1)\sigma_{U,1}+\binom{d}{2}.$$

Gröbner bases of $I_{n,d}$ have been described in [26]:

Theorem 1. *Let $0 \leq d \leq n/2$ be integers, $\mathbb{F}$ a field, and $\prec$ be an arbitrary term order on the monomials of $\mathbb{F}[\mathbf{x}]$ for which $x_n \prec x_{n-1} \prec \ldots \prec x_1$. Then the following set G of polynomials is a Gröbner basis of the ideal $I_{n,d}$:*

$$G = \{x_1^2 - x_1, \ldots, x_n^2 - x_n\} \cup \{x_J : \ J \in \binom{[n]}{d+1}\} \cup$$

$$\cup \{f_{H,d} : \ H \in \mathcal{H}_t \text{ for some } 0 < t \leq d\} .$$

A similar description is valid for $I_{n,n-d}$ in the place of $I_{n,d}$. The standard monomials for the complete uniform families have also been obtained. The next theorem is valid for an arbitrary term order $\prec$ such that $x_n \prec x_{n-1} \prec \ldots \prec x_1$. For the lex order it was proved in [6], and later it was extended to general term orders in [26].

Theorem 2. *Let $0 \leq d \leq n/2$ and denote by $\mathcal{M} = \mathcal{M}_d$ the set of all monomials x_G such that $G = \{s_1 < s_2 < \ldots < s_j\} \subset [n]$ for which $j \leq d$ and $s_i \geq 2i$ holds for every i, $1 \leq i \leq j$. Then $\mathcal{M}$ is the set of standard monomials for $I_{n,d}$ as well as for $I_{n,n-d}$ with respect to any term order $\prec$ as above.*

In particular, $|\mathcal{M}| = \binom{n}{d}$ and $\mathcal{M}$ is an $\mathbb{F}$ basis of the space of functions from $V_{\mathcal{F}}$ to $\mathbb{F}$. Also, Theorem 1 allows one to determine the reduced Gröbner bases of the ideals $I_{n,d}$. Here we note only the fact that a suitable *subset* of G turns out to be the reduced Gröbner basis of $I_{n,d}$ for $0 \leq d \leq \frac{n}{2}$.

3.2 Some Combinatorial Applications to q-uniform Families

Let p be a prime, k an integer, and $q = p^\alpha$, $\alpha \geq 1$. Put

$$\mathcal{F}(k,q) = \{K \subseteq [n] : \ |K| \equiv k \pmod{q}\} .$$

In [27] the following rank inequality is proved for the inclusion matrices of $\mathcal{F}(k,q)$:

Theorem 3. *Let p be a prime and k an integer. Let $q = p^\alpha > 1$. If $\ell \leq q-1$ and $2\ell \leq n$, then*

$$\mathit{rank}_{\mathbb{F}_p} \, I\left(\mathcal{F}(k,q), \binom{[n]}{\leq \ell}\right) \leq \binom{n}{\ell} .$$

This result is a generalization of a theorem of Frankl [19] covering the case $\alpha = 1$. Theorem 3 is a direct consequence of the next inclusion relation involving deglex standard monomials. In simple words, it states that the low degree standard monomials of $\mathcal{F}(k,q)$ are contained among the standard monomials of the *complete uniform* families.

Theorem 4. *Let p be a prime and $q = p^\alpha > 1$. Let $\prec$ be the deglex order on the monomials of $\mathbb{F}[\mathbf{x}]$ with $\mathbb{F} = \mathbb{F}_p$. Suppose further that $k, \ell \in \mathbb{N}$, for which $0 \leq k, \ell < q$, and $2\ell \leq n$. Then*

$$\mathrm{Sm}\,(I(\mathcal{F}(k,q))) \cap \mathbb{F}[\mathbf{x}]_{\leq \ell} \subseteq \mathcal{M}_\ell$$

hence

$$|\,\mathrm{Sm}\,(I(\mathcal{F}(k,q))) \cap \mathbb{F}\,[\mathbf{x}]_{\leq \ell}\,| \leq \binom{n}{\ell}\,.$$

The crucial point of the proof is the fact that we know quite explicitly a Gröbner basis for the complete uniform families. To be a bit more specific here, suppose that k' is an integer such that $0 \leq k' \leq n$ and $k \equiv k' \pmod q$. Using simple properties of binomial coefficients one can infer that $f_{H,k} \equiv f_{H,k'} \pmod p$, i.e., the coefficients of the two polynomials are the same modulo p. This fact, together with a Gröbner reduction argument leads to a proof of Theorem 4.

Babai and Frankl conjectured the following in [7], p. 115.

Theorem 5. *Let k be an integer and $q = p^\alpha$, $\alpha \geq 1$ a prime power. Suppose that $2(q-1) \leq n$. Assume that $\mathcal{F} = \{A_1, \ldots, A_m\}$ is a family of subsets of $[n]$ such that*

$$(a)\ |A_i| \equiv k \ (mod\ q)\ for\ i = 1, \ldots, m$$

$$(b)\ |A_i \cap A_j| \not\equiv k \ (mod\ q)\ for\ 1 \leq i, j \leq m,\ i \neq j\,.$$

Then

$$m \leq \binom{n}{q-1}\,.$$

We briefly sketch a proof from [27]: let $\mathbf{v}_i \in \mathbb{Z}^n$ denote the characteristic vector of A_i, and write

$$f_i(x_1, \ldots, x_n) = \binom{\mathbf{x} \cdot \mathbf{v}_i - k - 1}{q-1}\,.$$

This is a polynomial in n rational variables $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Q}^n$. By conditions (a) and (b) the integer $f_i(\mathbf{v}_j)$ is divisible by p iff $i \neq j$.

Let f_i' be the square-free reduction of f_i for $i = 1, \ldots, m$. Then $f_i' \in \mathbb{Z}[\mathbf{x}]$, because $f_i(\mathbf{v}) \in \mathbb{Z}$ for each $\mathbf{v} \in \{0,1\}^n$. Let $g_i \in \mathbb{F}_p[\mathbf{x}]$ is the reduction of f_i' modulo p, and $h_i \in \mathbb{F}_p[\mathbf{x}]$ be the reduction of g_i by a deglex Gröbner basis for the ideal $I(\mathcal{F}(k,q))$ over $\mathbb{F}_p$.

For $1 \leq i, j \leq m$ we have then

$$f_i(\mathbf{v}_j) = f_i'(\mathbf{v}_j) \equiv g_i(\mathbf{v}_j) \equiv h_i(\mathbf{v}_j) \pmod p\,.$$

These imply, that the polynomials h_i are linearly independent mod p. They have degree at most $q-1$ and they are and spanned by $\mathrm{Sm}\,(I(\mathcal{F}(k,q)))$. By Theorem 4 their number is at most $\binom{n}{q-1}$.

3.3 Wilson's Rank Formula

Consider the inclusion matrix $A = I\left(\binom{[n]}{d}, \binom{[n]}{m}\right)$, where $m \leq d \leq n - m$.

A famous theorem of Richard M. Wilson [46, Theorem 2] describes a diagonal form of A over $\mathbb{Z}$. As a corollary, he obtained the following rank formula:

Theorem 6. *Let p be a prime. Then*

$$rank_{\mathbb{F}_p}(A) = \sum_{\substack{0\le i\le m \\ p\nmid\binom{d-i}{m-i}}} \binom{n}{i} - \binom{n}{i-1}\,.$$

In [22] a simple proof is given which uses polynomial functions, and some basic notions related to Gröbner bases. The starting point is the observation that the rank of A is exactly the dimension of the linear space $\mathcal{P}_{d,m}$ over $\mathbb{F}_p$ of the functions from $V_{\binom{[n]}{d}}$ to $\mathbb{F}_p$ which are spanned by the monomials x_M with $|M| = m$.

The approach allows a considerable generalization of the rank formula. Here is a result of this kind from [22]:

Theorem 7. *Suppose that* $0 \le m_1 < m_2 \cdots < m_r \le d \le n - m_r$ *and let p be a prime. Consider the set family* $\mathcal{F} = \binom{[n]}{m_1} \cup \binom{[n]}{m_2} \cup \cdots \cup \binom{[n]}{m_r}$. *Then*

$$rank_{\mathbb{F}_p}\left(I\left(\binom{[n]}{d},\mathcal{F}\right)\right) = \sum_{\substack{0\le i\le m_r \\ p\nmid n_i}} \binom{n}{i} - \binom{n}{i-1},$$

where $n_i = \gcd\left(\binom{d-i}{m_1-i}, \binom{d-i}{m_2-i}, \ldots, \binom{d-i}{m_r-i}\right)$.

3.4 Generalizations of Uniform Families

Let n, k, ℓ be integers with $0 \le \ell - 1 \le k \le n$. The complete *$\ell$-wide* family is

$$\mathcal{F}^{k,\ell} = \{F \subseteq [n] : k - \ell < |F| \le k\}\,.$$

Theorem 1 was extended in [21] to complete ℓ-wide families. Gröbner bases and standard monomials are described there over an arbitrary ground field $\mathbb{F}$. As in the case $\ell = 1$, the bases are largely independent of the term order considered.

These results have been extended even further in [16]. Let q be a power of a prime p, and let n, d, ℓ be integers such that $1 \le n$, $1 \le \ell < q$. Consider the modulo q complete ℓ-wide family:

$$\mathcal{G} = \{F \subseteq [n] \,:\, \exists\, f \in \mathbb{Z} \text{ s. t. } d \le f < d + \ell \text{ and } |F| \equiv f \pmod{q}\}\ .$$

In [16] a Gröbner basis of the vanishing ideal $I(\mathcal{G})$ has been computed over fields of characteristic p. As before, it turns out that this set of polynomials is a Gröbner basis for all term orderings $\prec$, for which the order of the variables is $x_n \prec x_{n-1} \prec \cdots \prec x_1$. The standard monomials and the Hilbert function of $I(\mathcal{G})$ were also obtained. In this work the lex game method (see Sect. 4) was substantially used. As corollaries, several combinatorial applications follow. One of them is described next. It is a generalization of Theorem 5.

Let L be a subset of integers and $\mathcal{F}$ be a system of sets. Then $\mathcal{F}$ is *modulo q L-avoiding* if $G \in \mathcal{F}$ and $f \in L$ implies $|G| \not\equiv f \pmod{q}$. We call $\mathcal{F}$ *L-intersecting* if for any two distinct sets $G_1, G_2 \in \mathcal{F}$ the congruence $|G_1 \cap G_2| \equiv f \pmod{q}$ holds for some $f \in L$. A set $L \subseteq \{0, \dots, q-1\}$ is called a *modulo q interval* if it is either an interval of integers or a union of two intervals L_1 and L_2, such that $0 \in L_1$ and $q-1 \in L_2$.

Theorem 8. *Let q be a power of a prime, L be a modulo q interval and $\mathcal{F} \subseteq 2^{[n]}$ be a modulo q L-avoiding, L-intersecting family of sets. If $|L| \leq n - q + 2$, then*

$$|\mathcal{F}| \leq \sum_{k=|L|}^{q-1} \binom{n}{k} .$$

More generally, one may consider arbitrary fully symmetric set families. Let $D \subseteq [n]$ be arbitrary, and put

$$\mathcal{F}_D := \{Z \subseteq [n] \,:\, |Z| \in D\} .$$

Thus, $\mathcal{F}_D$ consists of all subsets of $[n]$ whose size is in D. It would be quite interesting to describe Gröbner bases and related structures for general set families of the form $\mathcal{F}_D$. Only some preliminary results are available, the most important of them being a beautiful theorem of Bernasconi and Egidi from [9]. It provides the deglex Hilbert function $h_{I(\mathcal{F}_D)}(m)$ of $I(\mathcal{F}_D)$ over $\mathbb{Q}$.

Theorem 9. *Let $0 \leq m \leq n$, and suppose that*

$$D = \{l_1, \dots, l_s\} \cup \{m_1, \dots, m_t\} ,$$

where $l_j \leq m$ and $m < m_1 < m_2 < \dots < m_t$. Assume also, that

$$\{0, 1, \dots, m\} \setminus \{l_1, \dots, l_s\} = \{n_1, n_2, \dots, n_{m+1-s}\} ,$$

with $n_1 > n_2 > \dots > n_{m+1-s}$ and $u = \min\{t, m+1-s\}$. Then we have

$$h_{I(\mathcal{F}_D)}(m) = \sum_{j=1}^{s} \binom{n}{l_j} + \sum_{j=1}^{u} \min\{\binom{n}{m_j}, \binom{n}{n_j}\} .$$

A combinatorial description of the deglex standard monomials for $I(\mathcal{F}_D)$ over $\mathbb{Q}$ was obtained in [42] in the case when D has the following property: for each integer i, at most one of i and $n-i$ is in D. This characterization uses generalized ballot sequences. It would be of interest to extend this to more general sets D. Multivalued generalizations of uniform families are considered in [29].

4 The Lex Game and Applications

Based on [17], we outline a combinatorial approach to the lexicographic standard monomials of the vanishing ideal of a finite sets of points $V \subseteq \mathbb{F}^n$. This technique can be applied to compute the lex standard monomials of sets of combinatorial interest. The idea has been extended to general zero dimensional ideals in [18].

In this section, we use the lexicographic ordering. As before, let $\mathbb{F}$ be a field, $V \subseteq \mathbb{F}^n$ a finite set, and $\mathbf{w} = (w_1, \ldots, w_n) \in \mathbb{N}^n$ an n dimensional vector of natural numbers. With these data as parameters, we define the Lex Game $\mathrm{Lex}(V; \mathbf{w})$, which is played by two persons, Lea and Stan.
Both Lea and Stan know V and $\mathbf{w}$. Their moves are:

1 Lea chooses w_n elements of $\mathbb{F}$.
Stan picks a value $y_n \in \mathbb{F}$, different from Lea's choices.
2 Lea now chooses w_{n-1} elements of $\mathbb{F}$.
Stan picks a $y_{n-1} \in \mathbb{F}$, different from Lea's (last w_{n-1}) choices.
... (The game proceeds in this way until the first coordinate.)
n Lea chooses w_1 elements of $\mathbb{F}$.
Stan finally picks a $y_1 \in \mathbb{F}$, different from Lea's (last w_1) choices.

The winner is Stan, if during the game he could select a vector $\mathbf{y} = (y_1, \ldots, y_n)$ such that $\mathbf{y} \in V$, otherwise Lea wins the game. (If in any step there is no proper choice y_i for Stan, then Lea wins also.)

Example. Let $n = 5$, and $\alpha, \beta \in \mathbb{F}$ be different elements. Let V be the set of all α-β sequences in $\mathbb{F}^5$ in which the number of the α coordinates is 2 or 3. We claim that Lea can win with the question vector $\mathbf{w} = (11100)$, but for $\mathbf{w} = (00110)$ Stan has a winning strategy.

First consider $\mathbf{w} = (11100)$. To have $\mathbf{y} \in V$, Stan is forced to select values from $\{\alpha, \beta\}$. If Stan gives only β for the last 2 coordinates, then Lea will choose α in the first three, therefore $\mathbf{y}$ cannot contain any α coordinates. However if Stan gives at least one α for the last 2 coordinates, then Lea, by keeping on choosing β, can prevent $\mathbf{y}$ to have at least two β coordinates.

For $\mathbf{w} = (00110)$ Stan's winning strategy is to pick $y_5 = \beta$, and choose from $\{\alpha, \beta\}$ (for the 4th and 3rd coordinates). If he selected so far α twice, then he can win by setting the first two coordinates to β. Otherwise he wins with the moves $y_1 = y_2 = \alpha$.

The game allows a nice characterization of the lexicographic leading monomials and standard monomials for V:

Theorem 10. *Let $V \subseteq \mathbb{F}^n$ be a finite set and $\mathbf{w} \in \mathbb{N}^n$. Stan wins $\mathrm{Lex}(V; \mathbf{w})$ if and only if $\mathbf{x}^{\mathbf{w}} \in \mathrm{Sm}\,(I(V))$. Equivalently, Lea wins the game if and only if $\mathbf{x}^{\mathbf{w}}$ is a leading monomial for $I(V)$.*

The theorem leads to a fast combinatorial algorithm to list those vectors $\mathbf{w} \in \mathbb{N}^n$ for which $\mathbf{x}^{\mathbf{w}} \in \mathrm{Sm}\,(I(V))$. The method uses constant times $|V|\,nk$ comparisons of field elements in the worst case, where k is the maximum number of different elements which appear in a fixed coordinate of points of V; see [17]. In particular, if $V \subseteq \{0, 1\}^n$ then $k \leq 2$ and hence we have a linear time algorithm.

The problem of computing lexicographic standard monomials for finite sets has had a long history starting with the seminal paper by Buchberger and Möller

[40]. Their algorithm, as well as the subsequent methods of Marinari, Möller and Mora [36] and Abbott, Bigatti, Kreuzer and Robbiano [1] give also a Gröbner basis of $I(V)$. For the arithmetic complexity of these methods we have the bound $O(n^2m^3)$ when V is a subset of $\mathbb{F}^n$ and $|V| = m$ (see Sect. 3 in [15] for a related discussion). The Lex Game provides only the standard monomials, but in return it appears to lead to a much faster algorithm (see [17] for the details). In general we have the bound $O(nm^2)$. In some important special cases, such as the case of small finite ground fields which appear naturally in coding applications, we have a linear bound $O(nm)$ on the time cost of the algorithm.

5 Partitions and Colorings

5.1 Permutations, Trees and Partitions

Let $\alpha_1, \ldots, \alpha_n$ be n different elements of $\mathbb{F}$ and put

$$\Pi_n := \Pi_n(\alpha_1, \ldots, \alpha_n) := \{(\alpha_{\pi(1)}, \ldots, \alpha_{\pi(n)}) : \ \pi \in S_n\}\,.$$

Π_n is the set of all permutations of the α_i, considered as a subset of $\mathbb{F}^n$.

We recall the definition of the complete symmetric polynomials. Let i be a nonnegative integer and write

$$h_i(x_1, \ldots, x_n) = \sum_{a_1+\cdots+a_n=i} x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}\,.$$

Thus, $h_i \in \mathbb{F}[x_1, \ldots, x_n]$ is the sum of all monomials of total degree i. For $0 \le i \le n$ we write σ_i for the i-th elementary symmetric polynomial:

$$\sigma_i(x_1, \ldots, x_n) = \sum_{S \subset [n],\ |S|=i} x_S\,.$$

For $1 \le k \le n$ we introduce the polynomials $f_k \in \mathbb{F}\,[\mathbf{x}]$ as follows:

$$f_k = \sum_{i=0}^{k} (-1)^i h_{k-i}(x_k, x_{k+1}, \ldots, x_n)\sigma_i(\alpha_1, \ldots, \alpha_n)\,.$$

We remark, that $f_k \in \mathbb{F}[x_k, x_{k+1}, \ldots, x_n]$. Moreover, $\deg f_k = k$ and the leading monomial of f_k is x_k^k with respect to any term order $\prec$ for which $x_1 \succ x_2 \succ \ldots \succ x_n$. In [25] the following was proved:

Theorem 11. *Let $\mathbb{F}$ be a field and let $\prec$ be an arbitrary term order on the monomials of $\mathbb{F}[x_1, \ldots, x_n]$ such that $x_n \prec \ldots \prec x_1$. Then the reduced Gröbner basis of $I(\Pi_n)$ is $\{f_1, f_2, \ldots, f_n\}$. Moreover the set of standard monomials is*

$$\{x_1^{\alpha_1} \ldots x_n^{\alpha_n} : \ 0 \le \alpha_i \le i-1, \ \text{for } 1 \le i \le n\}\,.$$

We remark, that [25] gives also the reduced Gröbner basis of the set Y_m of characteristic vectors of oriented trees with vertex set $[m]$. Here we have $Y_m \subseteq \mathbb{F}^n$ with $n = m(m-1)$ and the coordinates are indexed by the edges of the complete digraph KD_m. The term order $\prec$ involved is a lexicographic order. It would be interesting to understand a Gröbner basis of Y_m with respect to a deglex (or other degree compatible) order.

Recall, that a sequence $\lambda = (\lambda_1, \ldots, \lambda_k)$ of positive integers is a *partition* of n, if $\lambda_1 + \lambda_2 + \cdots + \lambda_k = n$ and $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k > 0$. Let $\mathbb{F}$ be a field, and $\alpha_0, \ldots, \alpha_{k-1}$ be k distinct elements of $\mathbb{F}$. Let $\lambda = (\lambda_1, \ldots, \lambda_k)$ be a partition of n and V_λ be the set of all vectors $\mathbf{v} = (v_1, \ldots, v_n) \in \mathbb{F}^n$ such that

$$|\{j \in [n] : \; v_j = \alpha_i\}| = \lambda_{i+1}$$

for $0 \leq i \leq k-1$.

In their study of the q-Kostka polynomials, Garsia and Procesi have described the deglex standard monomials of $I(V_\lambda)$ (Proposition 3.2 in [23]). They worked over $\mathbb{Q}$, but their argument is valid over an arbitrary field. The associated graded ring $\mathrm{gr}\mathbb{F}[\mathbf{x}]/I(V_\lambda)$ is also described there.

In [28] it is shown that the lexicographic standard monomials of $I(V_\lambda)$ are the same as the deglex standard monomials over an arbitrary $\mathbb{F}$. In the proof a new description of the orthogonal complement $(S^\lambda)^\perp$ (with respect to the James scalar product) of the Specht module S^λ is given. As applications, a basis of $(S^\lambda)^\perp$ is exhibited, and a combinatorial description of the Hilbert function of V_λ is provided. This approach provides a new, simpler proof of the Garsia-Procesi theorem on the deglex standard monomials. An interesting feature of the results is that both in the lex and deglex cases the standard monomials are independent of the specific choice of $\alpha_0, \ldots, \alpha_{k-1}$, or the field $\mathbb{F}$ itself.

These results partially extend the special cases we treated here earlier: the complete uniform set families, i.e., $\lambda = (n-d, d)$, see Theorem 2, and the permutations (the case $\lambda = (1^n)$), see Theorem 11. For general λ it seems to be difficult to give explicit Gröbner bases of $I(V_\lambda)$.

5.2 Graph Colorings

The algebraic study of graph colorings also employs fruitfully some Gröbner basis techniques. Here we briefly discuss some of these. Let G be a simple undirected graph on the vertex set $V = [n]$ and with edge set E. Let k be a fixed positive integer, and $\mathbb{F}$ be a field which contains k distinct k-th roots of unity. The set of those k-th roots of unity will be denoted by C_k. The *graph polynomial* $f_G \in \mathbb{F}[\mathbf{x}]$ is the polynomial

$$f_G := \prod_{(i,j) \in E, \; i<j} (x_i - x_j) \, .$$

Recall that a *k-coloring* of G is a map μ from $V(G)$ to C_k such that $\mu(i) \neq \mu(j)$, whenever $(i,j) \in E$. Moreover, a k coloring can be viewed as an element of $C_k^n \subseteq \mathbb{F}^n$. Let $\mathcal{K}$ be the set of graphs whose vertex set is $[n]$, which consist of a

$k+1$-clique and $n-k-1$ isolated vertices. We introduce some important ideals from $\mathbb{F}[\mathbf{x}]$:

$$J_{n,k} := \langle f_H : \ H \in \mathcal{K} \rangle$$

is the ideal generated by the graph polynomials of $k+1$-cliques on $[n]$. Put

$$I_{n,k} := \langle x_i^k - 1 : \ i \in V \rangle .$$

It is easy to show that $I_{n,k}$ is actually $I(C_k^n)$, in fact, $\{x_1^k - 1, \dots x_n^k - 1\}$ is a universal Gröbner basis of $I_{n,k}$. Finally set

$$I_{G,k} := I_{n,k} + \langle x_i^{k-1} + x_i^{k-2} x_j + \cdots + x_j^{k-1} : \ (i,j) \in E \rangle .$$

It is a simple matter to verify, that $I_{G,k}$ is the ideal of the k-colorings of G: $\mu \in \mathbb{F}^n$ is a common zero for all polynomials from $I_{G,k}$ iff μ is a valid k-coloring of G.

The fact that G is *not* k colorable admits an algebraic characterization:

Theorem 12. *The next statements are equivalent:*
(1) G is not k-colorable.
(2) The constant polynomial 1 belongs to $I_{G,k}$.
(3) The graph polynomial f_G belongs to $I_{n,k}$.
(4) The graph polynomial f_G belongs to $J_{n,k}$.

The equivalence of (1) and (2) is due to Bayer [8], (1) $\Leftrightarrow$ (3) is from Alon and Tarsi [5], this was reproved by Gröbner basis techniques by de Loera [34] and Mnuk [39]. The equivalence (1) $\Leftrightarrow$ (4) is due to Kleitman and Lovász [35]. The following beautiful theorem is due to de Loera [34]:

Theorem 13. *The set of polynomials $\{f_H : \ H \in \mathcal{K}\}$ is a universal Gröbner basis of the ideal $J_{n,k}$.*

Let μ be a k-coloring of G, $\ell \leq k$ be the number of colors actually used by μ. The class $cl(i)$ is the set of vertices with the same color as i. Let

$$m_1 < m_2 < \cdots < m_\ell = n$$

be the maximal elements (coordinates) of the color classes.

For a set U of indices let h_U^d denote the complete symmetric polynomial of degree d in the variables whose indices are in U.

We define the polynomials g_i as follows:

$$g_i = \begin{cases} x_i^k - 1 & \text{if } i = m_\ell, \\ h_{\{m_j,\dots,m_\ell\}}^{k-\ell+j} & \text{if } i = m_j \text{ for some } j \neq \ell, \\ x_i - x_{\max cl(i)} & \text{otherwise.} \end{cases}$$

Let $A_\mu := \langle g_1, g_2, \dots, g_n \rangle$ be the ideal generated by the polynomials g_i. Hillar and Windfeldt [31] obtained the following:

Theorem 14. *We have*

$$I_{G,k} = \cap_\mu A_\mu \,,$$

where μ runs over the k-colorings of G.

In the course of the proof they established, that for term orders $\prec$ with $x_1 \succ x_2 \succ \cdots \succ x_n$ the set $\{g_1, g_2, \ldots, g_n\}$ is actually a Gröbner basis of A_μ. By using these facts, they developed an algebraic characterization of unique k-colorability:

Theorem 15. *Let μ be a k-coloring of G that uses all k colors, $g_1, g_2, \ldots, g_n$ be the corresponding basis of A_μ. The following are equivalent.*
(1) G is uniquely k-colorable.
(2) The polynomials $g_1, g_2, \ldots, g_n$ generate $I_{G,k}$.
(3) The polynomials $g_1, g_2, \ldots, g_n$ are in $I_{G,k}$.
(4) The graph polynomial f_G is in the ideal $I_{n,k} : \langle g_1, g_2, \ldots, g_n \rangle$.
(5) $dim_{\mathbb{F}} \mathbb{F}\,[\mathbf{x}]\,/I_{G,k} = k!$

Condition (5) leads easily to an algebraic algorithm for testing the unique k-colorability of G. The left hand side is the number of the standard monomials for $I_{G,k}$ with respect to an arbitrary term order, hence (5) can be checked by standard techniques for computing Gröbner bases. See Sect. 6 in [31] for more details and data on computational experiments.

6 Alon's Combinatorial Nullstellensatz

Alon's Combinatorial Nullstellensatz, and in particular the resulting non-vanishing criterion from [4] is one of the most powerful algebraic tools in combinatorics, with dozens of important applications.

Let $\mathbb{F}$ be a field and $S_1, \ldots, S_n$ be nonempty, finite subsets of $\mathbb{F}$, $|S_i| = t_i$. Put $S = S_1 \times \cdots \times S_n$ and define $g_i(x_i) = \prod_{s \in S_i}(x_i - s)$.

Theorem 16. *(Theorem 1.1 from [4].) Let $f = f(\mathbf{x})$ be a polynomial from $\mathbb{F}[\mathbf{x}]$ that vanishes over all the common zeros of $g_1, \ldots, g_n$ (that is, if $f(\mathbf{s}) = 0$ for all $s \in S$). Then there exist polynomials $h_1, \ldots, h_n \in \mathbb{F}[\mathbf{x}]$ satisfying $deg(h_i) \leq deg(f) - deg(g_i)$ so that*

$$f = \sum_{i=1}^{n} h_i g_i \,.$$

Moreover if $f, g_1, \ldots, g_n$ lie in $R[\mathbf{x}]$ for some subring R of $\mathbb{F}$, then there are polynomials $h_i \in R[\mathbf{x}]$ as above.

The Combinatorial Nullstellensatz can be reformulated in terms of Gröbner bases. It states that $\{g_1, g_2, \ldots, g_n\}$ is a universal Gröbner basis of the ideal $I(S)$. The most important corollary of Theorem 16 is a non-vanishing criterion:

Theorem 17. *(Theorem 1.2 from [4].) Let $f = f(\mathbf{x})$ be a polynomial in $\mathbb{F}[\mathbf{x}]$. Suppose the degree of f is $\sum_{i=1}^n d_i$, where $d_i < t_i$ for all i and the coefficient of $\prod_{i=1}^n x_i^{d_i}$ in f is nonzero. Then there is a point $\mathbf{s} \in S$ such that $f(\mathbf{s}) \neq 0$.*

The theorem has numerous applications in combinatorial number theory, graph theory and combinatorics. To demonstrate the amazing versatility of Theorem 17, we give here an argument form [4] to prove the Cauchy and Davenport inequality from additive number theory. The inequality states that if p is a prime and A, B are two nonempty subsets of $\mathbb{Z}_p$, then

$$|A+B| \geq \min\{p, |A|+|B|-1\}\,.$$

We remark first, that the case $|A|+|B|>p$ is easy. We may then assume that $|A|+|B|\leq p$. Assume for contradiction that there is a subset $C \subset \mathbb{F}_p$ such that $C \supseteq A+B$, and $|C| = |A|+|B|-2$. Put

$$f(x,y) = \prod_{c\in C}(x+y-c) \in \mathbb{F}_p[x,y]\,.$$

Clearly f is identically zero on $A\times B$. Now set $n=2$, $S_1=A$, $S_2=B$, $t_1=|A|$ and $t_2=|B|$. We have $\deg f = t_1-1+t_2-1$; the coefficient of $x^{t_1-1}y^{t_2-1}$ is $\binom{t_1-1+t_2-1}{t_1-1}$, which is not 0 in $\mathbb{F}_p$, as $t_1-1+t_2-1<p$. By Theorem 17 f can not be identically zero on $A\times B$. This is a contradiction proving the Cauchy Davenport inequality.

There are natural ways to generalize Theorems 16 and 17. One is to prove a variant of the Nullstellensatz over rings instead of fields, an other is to consider the non-vanishing problem for multisets and not merely sets. Extensions along these lines are considered in [32], [33], and [38].

7 Gröbner Bases and S-extremal Set Systems

Gröbner basis methods may be useful when studying extremal problems of combinatorics (see Frankl [20] for a survey of extremal questions on set families). We give here a new application of this kind.

We say that a set system $\mathcal{F}\subseteq 2^{[n]}$ *shatters* a given set $S\subseteq[n]$ if

$$2^S = \{F\cap S:\ F\in\mathcal{F}\}.$$

The family of subsets of $[n]$ shattered by $\mathcal{F}$ is denoted by $\mathrm{Sh}(\mathcal{F})$. The notion of shattering occurs in various fields of mathematics, such as combinatorics, statistics, computer science, and logic. As an example, one can mention the Vapnik-Chervonenkis dimension of a set system $\mathcal{F}$, i.e. the size of the largest S shattered by $\mathcal{F}$. Sauer [43], Shelah [44] and Vapnik and Chervonenkis [45] proved that if $\mathcal{F}$ is a family of subsets of $[n]$ with no shattered set of size k (i.e. $\mathrm{VC}-\dim\mathcal{F}<k$), then

$$|\mathcal{F}| \leq \binom{n}{k-1}+\binom{n}{k-2}+\cdots+\binom{n}{0}\,,$$

and this inequality is best possible. The result is known as Sauer's lemma and has found applications in a variety of contexts, including applied probability.

It was proved by various authors (Aharoni and Holzman [3], Pajor [41], Sauer [43], Shelah [44]) that for every set system $\mathcal{F} \subseteq 2^{[n]}$ we have that $|\mathrm{Sh}(\mathcal{F})| \geq |\mathcal{F}|$. Accordingly, we define a set system $\mathcal{F}$ to be *S-extremal*, if $|\mathrm{Sh}(\mathcal{F})| = |\mathcal{F}|$. We refer to Bollobás and Radcliffe [10] for some basic properties of S-extremal set families, where they mention the lack of a good structural description of these families.

It turns out, that shattered sets are in close connection with the standard monomials of the ideal $I(\mathcal{F})$ for different term orders. To make this connection explicit, we first have to define a special family of term orders. At the beginning we have already defined the lex term order. By reordering the variables one can define another lex term order, so from now on we will talk about lex term orders based on some permutation of the variables $x_1, x_2, \ldots, x_n$. There are $n!$ possible lexicographic orders on n variables.

For a pair of sets $G \subseteq H \subseteq [n]$ we define the polynomial $f_{H,G}$ as

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

Lemma 1. *a) If $x_H \in \mathrm{Sm}\,(I(\mathcal{F}))$ for some term order, then $H \in \mathrm{Sh}(\mathcal{F})$.*
b) If $H \in \mathrm{Sh}(\mathcal{F})$, then there is a lex term order for which $x_H \in \mathrm{Sm}\,(I(\mathcal{F}))$.

Proof. a) Let $x_H \in \mathrm{Sm}\,(I(\mathcal{F}))$, and 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 $G = H \cap F$. Now $f_{H,G}(\mathbf{v}_F) \neq 0$ only if $H \cap F = G$. According to our assumption, there is no such set $F \in \mathcal{F}$, so $f_{H,G}(\mathbf{x}) \in I(\mathcal{F})$. This implies that $x_H \in \mathrm{Lm}(I(\mathcal{F}))$ for all term orders, since $\mathrm{lm}(f_{H,G}) = x_H$ for all term orders, giving a contradiction.

b) We prove that a lex order, where the variables of x_H are the smallest ones, satisfies the claim. Suppose the contrary, that $x_H \in \mathrm{Lm}(I(\mathcal{F}))$ 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 we may assume that $f(\mathbf{x})$ 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 w.r.t. this property. $\mathcal{F}$ 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 in the sum. So $f(\mathbf{v}_{F_0}) = \alpha_{G_0} \neq 0$, which contradicts $f \in I(\mathcal{F})$). This contradiction proves the statement. □

Combining the two parts of Lemma 1, we obtain that $\mathrm{Sm}\,(I(\mathcal{F})) \subseteq \mathrm{Sh}(\mathcal{F})$ for every term order, and

$$\mathrm{Sh}(\mathcal{F}) = \bigcup_{\text{term orders}} \mathrm{Sm}\,(I(\mathcal{F}))\,. \tag{1}$$

(Here, by identifying a squarefree monomial x_H with the set of indices $H \subseteq [n]$, we view $\mathrm{Sm}\,(I(\mathcal{F}))$ as a set family over $[n]$.) Note that on the right hand side it is sufficient to take the union over the lex term orders only. Using the lex game method, one can efficiently compute $\mathrm{Sm}\,(I(\mathcal{F}))$ for any lex term order. However, as the number of lex orders is $n!$, (1) does not immediately provide an efficient way to calculate $\mathrm{Sh}(\mathcal{F})$. Nevertheless Lemma 1 implies at once a simple algebraic characterization of S-extremal set systems:

Theorem 18. *$\mathcal{F}$ is S-extremal if and only if $\mathrm{Sm}\,(I(\mathcal{F}))$ is the same for all lex term orders.*

Theorem 18 leads to an algebraic characterization of S-extremal set systems, involving the Gröbner bases of $I(\mathcal{F})$.

Theorem 19. *$\mathcal{F} \subseteq 2^{[n]}$ is S-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. Suppose first, that $\mathcal{F}$ is S-extremal. Consider all minimal sets $S \subseteq [n]$ for which $S \notin \mathrm{Sh}(\mathcal{F})$, with a corresponding polynomial $f_{S,H}$. Here $H \subseteq S$ is a set which is not of the form $S \cap F$ for any $F \in \mathcal{F}$. Denote the set of these sets S by $\mathcal{S}$ and fix an arbitrary 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 $m \in \mathrm{Lm}(I(\mathcal{F}))$, there is a monomial in $\{x_S, S \in \mathcal{S}\} \cup \{x_i^2, i \in [n]\}$ that divides m. If m is not square-free, then this is trivial. Now suppose m is square-free, say $m = x_F$ for a subset $F \subseteq [n]$. $\mathcal{F}$ is extremal, thus we have $|\mathrm{Sh}(\mathcal{F})| = |\mathcal{F}| = |\,\mathrm{Sm}\,(I(\mathcal{F}))\,|$ and hence $\mathrm{Sm}\,(I(\mathcal{F})) = \mathrm{Sh}(\mathcal{F})$. We have then $F \notin \mathrm{Sh}(\mathcal{F})$, as m is a leading monomial. Then there is an $S \in \mathcal{S}$ with $S \subseteq F$. This proves that our basis is a Gröbner basis.
For the opposite direction, suppose that there is a common Gröbner basis G for all term orders of the desired form. G is a Gröbner basis of $I(\mathcal{F})$, so $\mathrm{Lm}(G) = \{x_S, S \in \mathcal{S}\} \cup \{x_i^2, i \in [n]\}$ determines $\mathrm{Lm}(I(\mathcal{F}))$ and so $\mathrm{Sm}\,(I(\mathcal{F}))$. This clearly implies that $\mathrm{Sm}\,(I(\mathcal{F}))$ is the same for all term orders, since G is a common Gröbner basis for all term orders. □

We remark that in the theorem the phrase *all term orders* may be replaced by *a term order*. To see this, please note that the standard monomials of $\mathcal{F}$ are then precisely the monomials x_F where there is no polynomial $f_{S,H}$ in the basis with $S \subseteq F$. This is independent of the term order considered.

In addtion to this characterization, Theorem 18 leads also to an efficient algorithm for testing the S-extremality. The test is based on the theorem below.

Theorem 20. *Take n orderings of the variables such that for every index i there is one in which x_i is the greatest element, and take the corresponding lex 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 lex orders. Suppose that $\mathcal{F}$ is not S-extremal. Then 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. $\mathrm{Sm}\,(I(\mathcal{F}))$ is a basis of the linear space $\mathbb{F}[\mathbf{x}]/I(\mathcal{F})$, and since all functions from $V_{\mathcal{F}}$ to $\mathbb{F}$ are polynomials, every leading monomial can be written uniquely as an $\mathbb{F}$-linear combination of standard monomials, as a function on $V_{\mathcal{F}}$. This holds for x_H as well. As functions on $V_{\mathcal{F}}$ we have

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

Suppose that for all sets G in the 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 $G' \subseteq H$, giving that

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

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, giving 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. □

In view of the preceding theorem, it is enough to calculate the standard monomials e.g. for a lexicographic term order and its cyclic permutations, and to check, whether they differ or not. The standard monomials can be calculated in time $O(n|\mathcal{F}|)$ for one lexicographic term order, see [17]. We have n term orders, therefore the total running time of the algorithm is $O(n^2|\mathcal{F}|)$.

Theorem 21. *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 [24] by Greco, where the time bound is $O(nm^3)$. But it is still open whether it is possible to test S-extremality in linear time (i.e. in time $O(nm)$).

We note that the results discussed here can be generalized to a multivalued setting, see [37]. The starting point is Theorem 18. We define a set $V \subseteq \mathbb{F}^n$ to be *S-extremal*, if $\mathrm{Sm}\,(I(V))$ is independent of the term order, i.e. it stays the same for all term orders.

References

1. Abbott, J., Bigatti, A., Kreuzer, M., Robbiano, L.: Computing Ideals of Points. J. Symbolic Comput. 30, 341–356 (2000)
2. Adams, W. W., Loustaunau, P.: An Introduction to Gröbner bases, Graduate Studies in Mathematics, Vol. 3, American Mathematical Society, Providence (1994)

3. Aharoni, R., Holzman, R.: Personal communication, cited in [24].
4. Alon, N.: Combinatorial Nullstellensatz. Combinatorics, Probability and Computing. 8, 7–29 (1999)
5. Alon, N., Tarsi, M.: Colorings and Orientation of Graphs. Combinatorica. 12, 125–134 (1992)
6. Anstee, R.P., Rónyai, L., Sali, A.: Shattering News. Graphs and Combinatorics. 18, 59–73 (2002)
7. Babai, L., Frankl, P.: Linear Algebra Methods in Combinatorics. Prel. vers. (1992)
8. Bayer, D.: The Division Algorithm and the Hilbert Scheme. PhD. Thesis. Harvard University (1982)
9. Bernasconi, A., Egidi, L.: Hilbert Function and Complexity Lower Bounds for Symmetric Boolean Functions. Information and Computation. 153, 1–25 (1999)
10. Bollobás, B., Radcliffe, A.J.: Defect Sauer Results. Journal of Combinatorial Theory, Series A. 72, 189–208 (1995)
11. Buchberger, B.: Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal. Doctoral thesis, University of Innsbruck, 1965. *English Translation:* An Algorithm for Finding the Basis Elements in the Residue Class Ring Modulo a Zero Dimensional Polynomial Ideal. Journal of Symbolic Computation, Special Issue on Logic, Mathematics, and Computer Science: Interactions. 41, 475-511 (2006)
12. Buchberger, B.: Ein algorithmisches Kriterium fur die Lösbarkeit eines algebraischen Gleichungssystems. Aequationes Mathematicae. 4, 374-383 (1970) *English translation:* An Algorithmic Criterion for the Solvability of Algebraic Systems of Equations. In: Buchberger, B., Winkler, F. (eds.) Gröbner Bases and Applications, London Mathematical Society Lecture Note Series, vol. 251, pp. 535 -545. Cambridge University Press (1998)
13. Buchberger, B.: Gröbner-Bases: An Algorithmic Method in Polynomial Ideal Theory. In: Bose, N.K. (ed.) Multidimensional Systems Theory - Progress, Directions and Open Problems in Multidimensional Systems Theory, pp. 184-232. Reidel Publishing Company, Dordrecht - Boston - Lancaster (1985)
14. Cox, D., Little, J., O'Shea, D.: Ideals, Varieties, and Algorithms. Springer-Verlag, Berlin, Heidelberg (1992)
15. Farr, J. B., Gao, S.: Computing Gröbner Bases for Vanishing Ideals of Finite Sets of Points. In: Fossorier, P.C.M, Imai, H., Lin, S., Poli, A. (eds.) Applied Algebra, Algebraic Algorithms and Error-correcting Codes, LNCS, vol. 3857, pp. 118–127. Springer, Berlin (2006)
16. Felszeghy, B., Hegedűs, G., Rónyai, L.: Algebraic Properties of Modulo q complete ℓ-wide Families. Combinatorics, Probability and Computing. 18, 309–333 (2009)
17. Felszeghy, B., Ráth, B., Rónyai, L.: The lex game and some applications. J. Symbolic Computation. 41, 663–681 (2006)
18. Felszeghy, B., Rónyai, L.: On the lexicographic standard monomials of zero dimensional ideals. In: Proc. 10th Rhine Workshop on Computer Algebra (RWCA), pp. 95-105 (2006)
19. Frankl, P.: Intersection Theorems and mod p Rank of Inclusion Matrices. Journal of Combinatorial Theory, Series A. 54, 85–94 (1990)
20. Frankl, P.: Extremal set systems. In: Graham, R.L., Grötschel, M., Lovász, L. (eds.) Handbook of combinatorics, vol. 2., pp. 1293 - 1329. MIT Press, Cambridge (1996)
21. Friedl, K., Hegedűs, G., Rónyai, L.: Gröbner Bases for Complete ℓ-wide Families. Publ. Math. Debrecen. 70, 271–290 (2007)

22. Friedl, K., Rónyai, L.: Order Shattering and Wilson's Theorem. Discrete Mathematics. 270, 127–136 (2003)
23. Garsia, A.M., Procesi, C.: On Certain Graded S_n-modules and the q-Kostka Polynomials. Advances in Mathematics. 94, 82–138 (1992)
24. Greco G.: Embeddings and Trace of Finite Sets. Information Processing Letters. 67, 199–203 (1998)
25. Hegedűs, G., Nagy, A., Rónyai, L.: Gröbner bases for permutations and oriented trees. Annales Univ. Sci. Budapest, Sectio Computatorica. 23, 137–148 (2004)
26. Hegedűs, G., Rónyai, L.: Gröbner Bases for Complete Uniform Families. Journal of Algebraic Combinatorics. 17, 171–180 (2003)
27. Hegedűs, G., Rónyai, L.: Standard Monomials for q-uniform Families and a Conjecture of Babai and Frankl. Central European Journal of Mathematics. 1, 198–207 (2003)
28. Hegedűs, G., Rónyai, L.: Standard Monomials for Partitions. Acta Mathematica Hungarica. 111, 193–212 (2006)
29. Hegedűs, G., Rónyai, L.: Multivalued Generalizations of the Frankl–Pach Theorem. To appear, Journal of Algebra and its Applications. http://arxiv.org/pdf/1008.4660
30. Herzog, J., Hibi, T.: Monomial Ideals, GTM vol. 260., Springer-Verlag, London, Dordrecht, Heidelberg, New York (2010)
31. Hillar, C.J., Windfeldt, T.: Algebraic Characterization of Uniquely Vertex Colorable Graphs. Journal of Combinatorial Theory, Series B. 98, 400–414 (2007)
32. Kós, G., Rónyai, L.: Alon's Nullstellensatz for multisets. http://arxiv.org/pdf/1008.2901
33. Kós, G., Mészáros, T., Rónyai, L.: Some Extensions of Alon's Nullstellensatz. http://arxiv.org/abs/1103.4768
34. de Loera, J.A.: Gröbner Bases and Graph Colorings. Beiträge zur Algebra und Geometrie. 36, 89–96 (1995)
35. Lovász, L.: Stable sets and Polynomials. Discrete Mathematics. 124, 137–153 (1994)
36. Marinari, M.G., Möller, H.M., Mora, T.: Gröbner Bases of Ideals Defined by Functionals with an Application to Ideals of Projective Points. Appl. Algebra Engrg. Comm. Comput. 4, 103–145 (1993)
37. Mészáros, T.: S-extremal Set Systems and Gröbner Bases. MSc Thesis, BME, Budapest (2010) http://www.math.bme.hu/~slovi/thesiswork.pdf
38. Michałek, M.: A Short Proof of Combinatorial Nullstellensatz. American Mathematical Monthly. 117, 821–823 (2010)
39. Mnuk, M.: On an Algebraic Description of Colorability of Planar Graphs. In: Nakagawa, K. (ed.) Logic, Mathematics and Computer Science: Interactions, Proc. of the Symposium in Honor of Bruno Buchberger's 60th Birthday, pp. 177-186. (2002)
40. Möller, H. M.; Buchberger, B.: The Construction of Multivariate Polynomials with Preassigned Zeros. In: Calmet, J. (ed.) Computer algebra, EUROCAM 1982. LNCS, vol. 144, pp. 24–31, Springer, Berlin-New York (1982)
41. Pajor, A.: Sous-espaces l_1^n des espaces de Banach, Travaux en Cours. Hermann, Paris (1985)
42. Pintér, D., Rónyai, L.: Standard Monomials of some Symmetric Sets. Acta Universitatis Apulensis. Math. Inform. No. 10, 331–344 (2005)
43. Sauer, N.: On the Density of Families of Sets. Journal of Combinatorial Theory, Series A. 13, 145–147 (1972)
44. Shelah, S.: A Combinatorial Problem: Stability and Order for Models and Theories in Infinitary Language. Pacific Journal of Mathematics. 41, 247–261 (1972)

45. Vapnik, V.N., Chervonenkis, A.Ya.: On the Uniform Convergence of Relative Frequencies of Events to their Probabilities. Theory of Probability and its Applications. 16, 264–280 (1971)
46. Wilson, R.M.: A Diagonal Form for the Incidence Matrices of t-subsets vs. k-subsets. European Journal of Combinatorics. 11, 609–615 (1990)