S-extremal set systems and Gröbner bases
S-extremal set systems and Gröbner bases

Tamás Mészáros

Budapest University of Technology and Economics

Advisor: Lajos Rónyai

Budapest University of Technology and Economics
Contents

1 Introduction .................................................. 1

2 Some general statements ......................................... 3

3 An algebraic approach ........................................... 6

4 Testing extremality ............................................. 8
   4.1 Test #1 .................................................. 9
   4.2 Test #2 .................................................. 9

5 Set system operations ........................................... 10
   5.1 Bit flip ................................................... 11
   5.2 Translation ............................................... 11
   5.3 Sum ....................................................... 11

6 Downshifts ....................................................... 12

7 A graph-theoretical aspect ..................................... 16

8 Some remarks on the VC dimension ............................ 19

9 Conclusion and future work ..................................... 19
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. Our goal is to characterize $S$-extremal set systems, those which shatter exactly $|\mathcal{F}|$ sets. Several characterizations of these combinatorial structures are given using algebraic methods. These characterizations then lead to algorithms which determine whether a set system $\mathcal{F} \subseteq 2^{[n]}$ is extremal or not. The running time of one of the algorithms is $O(n^2|\mathcal{F}|)$, which is an improvement over the previous bound of G. Greco in [6].

This study is organized as follows. After the introduction in Section 1, in Section 2 we study some general statements from this field. At the end of Section 3 we present first of our main results, a new characterization of extremal set systems using Gröbner bases. After this, in Section 4 we apply the characterizations from the preceding section to develop efficient algorithms for testing extremality. In Section 5 and 6 we examine set system operations, mainly the downshift operation, and we present our results on their connection to extremal set systems. In Section 7 we investigate the problem from a graph-theoretical point of view. Finally in Section 8 we conclude with some remarks concerning the Vapnik-Chervonenkis dimension of a set system.

1 Introduction

The structures we want to characterize are defined in terms of local properties. In fact, to check whether these local properties are satisfied or not is a hard task in general. In this study we give characterizations using mainly algebraic methods.

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 $2^{[n]}$ and for the ring of polynomials in $n$ variables over $\mathbb{F}$ we will use the usual notation $\mathbb{F}[x_1,\ldots,x_n]$.

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

**Definition 1.1** 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 $\text{Sh}(\mathcal{F})$.

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

**Proposition 1.1** (See [3].) $|\text{Sh}(\mathcal{F})| \geq |\mathcal{F}|$

**Proof:** We will prove this statement by induction on $n$. For $n = 1$ the statement is trivial. Now suppose $n > 1$. We construct 2 new set systems:
\[ F_0 = \{ F : F \in \mathcal{F}; n \not\in F \}, \]
\[ F_1 = \{ F \setminus \{ n \} : F \in \mathcal{F}; n \in F \}. \]

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

\[ |Sh(\mathcal{F})| \geq |Sh(\mathcal{F}_0)| + |Sh(\mathcal{F}_1)| \geq |\mathcal{F}_0| + |\mathcal{F}_1| = |\mathcal{F}|. \]

Thus every set system \(\mathcal{F}\) shatters at least \(|\mathcal{F}|\) sets. This inequality was proved by various authors (Aharoni and Holzman; Pajor, 1985; Sauer, 1972; Shelah, 1972), and studied by many others. We are interested in the case of equality, when a set system shatters exactly \(|\mathcal{F}|\) sets:

**Definition 1.2** 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. When considering this definition one can make a useful observation in connection with Proposition 1.1. In its proof we have seen a decomposition of \(\mathcal{F}\) and a recursion-like inequality for the sizes of the families of shattered sets. From this inequality we can conclude the following corollary:

**Corollary 1.1.1** If \(\mathcal{F}\) is extremal then so is \(\mathcal{F}_0\) and \(\mathcal{F}_1\).

We can consider the elements of \(\mathcal{F}\) as characteristic vectors from \(\{0, 1\}^n \subseteq \mathbb{F}^n\). So if we consider a polynomial \(p(x) \in \mathbb{F}[x]\), it is correct to substitute the elements of \(\mathcal{F}\) into \(p(x)\). In this way we can define the ideal \(I(\mathcal{F}) \subseteq \mathbb{F}[x]\), the ideal of the polynomials vanishing on \(\mathcal{F}\).

**Definition 1.3** The relation \(\prec\) is a term order on the monomials if it is a linear order with 1 as minimal element and is monotone with respect to multiplication.

As an example of a term order consider the lexicographic ordering of monomials. We say that \(x^w\) is smaller than or equal to \(x^u\) 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. By reordering the variables, we can get another lexicographic order.

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

**Proposition 1.2** (Dickson’s lemma, see [9].) 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 \(x^w | x^u\). Then \(\frac{x^w}{x^u}\) is a monomial as well, so \(1 \leq \frac{w}{u}\). And then when multiplying by \(x^w\) we get the desired inequality. For the proof of the second part see [9].
Now we introduce some notations related to ideals of polynomials. If $\prec$ is a fixed term order, and $f(x) = f_1(x) + f_2(x) + \cdots + f_n(x) \in \mathbb{F}[x]$ is an arbitrary nonzero polynomial, we denote by $\text{lm}(f(x))$ the leading monomial of $f(x)$ according to $\prec$. For an ideal of polynomials $I$ we denote by $\text{Lm}(I)$ the set of leading monomials of the polynomials in $I$. 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 $\text{Sm}(I)$. The standard monomials of an ideal will be henceforth of great importance. The other tool we use in the characterization of extremal set systems, as is visible from the title, is the idea of a Gröbner basis.

**Definition 1.4** Let $I$ be an ideal of $\mathbb{F}[x]$. 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 $\text{lm}(g)$ divides $\text{lm}(f)$.

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 [9], [8] or in Hungarian by [7].

2 Some general statements

In this section we discuss some general and well known properties of Gröbner bases and standard monomials, most of them can be found in [9].

**Proposition 2.1** 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 $0 \neq f(x) \in I$ an arbitrary element. Since $G$ is a Gröbner basis, there exists a polynomial $g_1(x) \in G$, such that $\text{lm}(g_1(x)) | \text{lm}(f(x))$. With $r_1(x) = \frac{\text{lm}(f(x))}{\text{lm}(g_1(x))}$ and $f_1(x) = f(x) - b_1r_1(x)g_1(x)$, where $b_1$ is the coefficient of $\text{lm}(f(x))$ in $f(x)$, we have $\text{lm}(f_1(x)) \prec \text{lm}(f(x))$, and we can continue this with $f_1(x)$. Since there cannot be an infinite downward chain of monomials starting with $\text{lm}(f(x))$ according to the $\prec$ term order, this process terminates in finitely many steps giving an expression

$$f(x) = b_1r_1(x)g_1(x) + b_2r_2(x)g_2(x) + \cdots + b_mr_m(x)g_m(x),$$

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

This is the reason why they are called bases. The question arises, whether every nonzero ideal has a Gröbner basis. The answer is fortunately yes, a proof
can be found in [9]. This suffices for us, since we only use simple properties that follow from the definition.

In the remainder of this section we investigate basic properties of standard monomials.

Definition 2.1 $\mathcal{F} \subseteq 2^{|n|}$ is a down-set if $H \subseteq F \in \mathcal{F}$ implies $H \in \mathcal{F}$, and similarly is an up-set if $H \supseteq F \in \mathcal{F}$ implies $H \in \mathcal{F}$.

Proposition 2.2 $\text{Sm}(I(\mathcal{F}))$ is a down-set and $\text{Lm}(I(\mathcal{F}))$ is an up-set.

Proof: If $x^u | x^v$ and $x^v \in \text{Lm}(I(\mathcal{F}))$, then there exists a polynomial $p(x)$ in $I(\mathcal{F})$ with $x^u$ as its leading monomial. Since $I(\mathcal{F})$ is an ideal, and $\frac{x^v}{x^u}$ a monomial, $q(x) = \frac{x^v}{x^u} p(x) \in I(\mathcal{F})$. However, because of the properties of a term order, the leading monomial of $q(x)$ is $x_u$, and so $x_u \in \text{Lm}(I(\mathcal{F}))$. 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 $\text{Sm}(I)$ is a basis of $\mathbb{F}(x)/I$ as an $\mathbb{F}$ vector space.

Proof: Clearly there are no two elements of $\text{Sm}(I)$ belonging to the same coset, otherwise there would be $x^{u_1}, x^{u_2} \in \text{Sm}(I)$ for which $f(x) = x^{u_1} - x^{u_2} \in I$, 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 $\text{Sm}(I)$ are linearly independent.

Now let us take an arbitrary coset from the quotient ring represented by $f(x)$. There are two possibilities. If $\text{lm}(f(x))$ is a standard monomial, then we can continue with $f(x) - b_1 \text{lm}(f(x))$, where $b_1$ is the coefficient of $\text{lm}(f(x))$. If this is not the case, then there exists a polynomial $g_1(x) \in I$ such that $\text{lm}(g_1(x)) = \text{lm}(f(x))$. Now we can continue with $f_1(x) = f(x) - \frac{b_1}{1} g_1(x)$, where $c_1$ is the coefficient of $\text{lm}(f(x))$ in $g_1(x)$. For this we have $\text{lm}(f_1(x)) \prec \text{lm}(f(x))$. Since there is no infinite downward chain of monomials starting with $\text{lm}(f(x))$ according to the $\prec$ term order, this process terminates in finitely many steps with $f(x) = s(x) + g(x)$, where $s(x)$ contains just standard monomials and $g(x) \in I$. So the coset represented by $f(x)$ is the sum of the cosets represented by the monomials in $s(x)$ in the quotient ring, that is, $\text{Sm}(I)$ generates $\mathbb{F}(x)/I$ as an $\mathbb{F}$ vector space. Together with linear independence, this means that $\text{Sm}(I)$ is a basis of it as well. ■

Proposition 2.4 $\text{dim}_\mathbb{F}\mathbb{F}(x)/I(\mathcal{F}) = |\mathcal{F}|$

Proof: Let $F$ be the set of characteristic vectors from $\mathcal{F}$. $\mathbb{F}(x)/I(\mathcal{F})$ is isomorphic to the space of functions from $F$ to $\mathbb{F}$, since every such function is a polynomial. But the space of these functions has dimension $|F| = |\mathcal{F}|$. ■

From the last two statements a very important fact follows, namely that $|\text{Sm}(I(\mathcal{F}))| = |\mathcal{F}|$. This equality will be essential in some of our further results.
For a subset $H \subseteq [n]$ denote by $x_H$ the monomial $\prod_{i \in H} x_i$; in particular, $x_0 = 1$. For $f(x) \in \mathbb{F}(x)$ let us denote by $\overline{f(x)}$ the polynomial that we get from $f(x)$ by replacing each monomial $x_1^{\alpha_1}x_2^{\alpha_2}\ldots x_m^{\alpha_m}$ (where $\alpha_j > 0$ with $j = 1, \ldots, m$) by $x_1^{\alpha_2}x_2^{\alpha_2}\ldots x_m^{\alpha_m}$.

**Proposition 2.5** $f(x) \in I(\mathcal{F})$ if and only if $\overline{f(x)} \in I(\mathcal{F})$.

**Proof:** It is clear, that $x_i^2 - x_i \in I(\mathcal{F})$ for every index $i$. Let $f(x)$ be an arbitrary polynomial and $cx_1^{\alpha_1}x_2^{\alpha_2}\ldots x_m^{\alpha_m}$ one of its monomials with $\alpha_1 > 2$. Consider the polynomial

$$f^*(x) = f(x) - c(x_i^2 - x_i)x_1^{\alpha_1-1}x_2^{\alpha_2}\ldots x_m^{\alpha_m}.$$  

Since $x_i^2 - x_i \in I(\mathcal{F})$ we have $f(x) \in I(\mathcal{F})$ if and only if $f^*(x) \in I(\mathcal{F})$. Moreover, the degree of $x_i$ has decreased. By continuing this we get $f(x) = \overline{f(x)} + g(x)$, where $g(x) \in I(\mathcal{F})$, so $f(x) \in I(\mathcal{F})$ if and only if $\overline{f(x)} \in I(\mathcal{F})$. ■

So if we want to compute, for example, $Lm(I(\mathcal{F}))$ for the ideal $I(\mathcal{F})$, it suffices to deal with the polynomials where there is no monomial with variables of degree higher than 1. Now let us consider the set of standard monomials again, and investigate its structure.

**Proposition 2.6** $Sm(I(\mathcal{F})) \subseteq \{x_H, H \subseteq [n]\}$

**Proof:** It is easy to see that if $\mathcal{H} \subseteq \mathcal{F} \subseteq 2^{[n]}$, then $Sm(I(\mathcal{H})) \subseteq Sm(I(\mathcal{F}))$. So it suffices to show that for $\mathcal{F} = 2^{[n]}$ we have $Sm(I(\mathcal{F})) = \{x_H, H \subseteq [n]\}$. This follows immediately from the fact that $I(\mathcal{F}) = (x_i^2 - x_i, i = 1, \ldots, n)$.

To prove the last equality it suffices to show that a polynomial of the form $0 \neq f(x) = \sum_{H \subseteq [n]} \alpha_H x_H$ cannot vanish on $\{0,1\}^n$. Every $\{0,1\}^n \rightarrow \mathbb{F}$ function can be written in this form. The dimension of the space of all these functions is $2^n$, but the number of monomials of the form $x_H$ is $2^n$ as well, so consequently they must be linearly independent. In particular, no such $f$ can vanish on $\{0,1\}^n$. ■

With the correspondence established in Proposition 2.6, the collection of monomials $Sm(I(\mathcal{F}))$ can be viewed either as a downward closed set system or as a collection of vectors. Instead of the monomial $x_H$ we can consider the set $H$, or equivalently, its characteristic vector. From now on it will depend on the context whether we consider the elements of $Sm(I(\mathcal{F}))$ to be monomials, sets or vectors.

At the end of this section we give a recursive method for the construction of $Sm(I(\mathcal{F}))$ for a fixed lexicographic term order. Let $\mathcal{F} \subseteq 2^{[n]}$ be an arbitrary set system. Without loss of generality we can suppose that $x_1 > x_1 > \cdots > x_n$. Let $\mathcal{F}_0$ and $\mathcal{F}_1$ be set systems constructed from $\mathcal{F}$ as in Proposition 1.1. We claim that

5
\[ Sm(I(F)) = Sm(I(F_0)) \cup Sm(I(F_1)) \cup \{ D \cup \{ n \} : D \in Sm(I(F_0)) \cap Sm(I(F_1)) \} \]

The detailed proof of this statement can be found in [2]. For the case \(|F| = 1\), the set of standard monomials can be computed easily, because it is simply \(\{1\}\), and starting from this one can easily compute \(Sm(I(F))\) using the recursion above.

Another method to obtain \(Sm(I(F))\) from \(F\) using set system operations will be given in Section 6. A further very efficient algorithm for this task was given in [5].

### 3 An algebraic approach

In the first part of this section 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 3.1** If \(x_H \in Sm(I(F))\) for some term order, then \(H \in Sh(F)\).

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

This means that \(Sm(I(F)) \subseteq Sh(F)\) for every \(i \in [n]\). We now investigate the other direction:

**Proposition 3.2** If \(H \in Sh(F)\), then there is a lexicographic term order for which we have \(x_H \in Sm(I(F))\).

**Proof:** We prove that a lexicographic order where the variables of \(x_H\) are the smallest satisfies the condition. Suppose the contrary, that \(x_H \in Lm(I(F))\) for this term order. Then there is a polynomial \(f(x)\) vanishing on \(F\) 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 \(f(x)\) 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 \(f(x)\), and is minimal. \(F\) shatters \(H\), so there exists a set \(F_0 \in F\) such that \(G_0 = F_0 \cap H\). For this we have \(x_{G_0}(v_{F_0}) = 1\), and since \(G_0\) was minimal, \(x_{G}(v_{F_0}) = 0\) for every other set \(G\). So

\[
\sum_{G \subseteq H} \alpha_G x_G(v_{F_0}) = \alpha_{G_0} \neq 0.
\]

But on the other hand, since \(f(x) \in I(F), f(v_{F_0}) = 0\). This contradiction proves the statement. ■
Combining the last two results we have
\[ Sh(F) = \bigcup_{\text{lex orders}} Sm(I(F)). \]

Even though \( Sm(F) \) can be computed fast for every term order [5], this formula does not give an efficient way for computing \( Sh(F) \), because the number of lexicographic orders is \( n! \). But for extremal set systems we obtain the following very important corollary:

**Corollary 3.2.1** \( F \) is extremal if and only if \( Sm(I(F)) \) is the same for all lexicographic term orders.

**Proof:** Suppose that \( F \) is extremal, i.e. \( |F| = |Sh(F)| \). Since \( Sh(F) = \bigcup_{\text{lex orders}} Sm(I(F)) \) and for every term order \( |Sm(I(F))| = |F| \), 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. \( \blacksquare \)

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

For a pair of sets \( G \subseteq H \subseteq [n] \) define the following polynomial
\[ f_{H,G} = \left( \prod_{j \in G} x_j \right) \left( \prod_{i \in H \setminus G} (x_i - 1) \right). \]

**Proposition 3.3** If \( S \notin Sh(F) \), then there exists a set \( H \subseteq S \), such that \( f_{S,H}(v_F) = 0 \), \( \forall F \in F \), i.e. \( f_{S,H} \in I(F) \).

**Proof:** The statement was already proved in Proposition 3.1. \( \blacksquare \)

**Proposition 3.4** If the set \( S \) in the previous proposition is minimal (in the sense that all proper subsets \( S' \) of \( S \) are in \( Sh(F) \)) and \( F \) is extremal, then the corresponding \( H \) is unique.

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

**Proposition 3.5** If for all but \(|F|\) sets \(S \subseteq [n]\) there exists a set \(H \subseteq S\) that \(f_{S,H} \in I(F)\), then \(F\) is extremal.

**Proof:** \(f_{S,H} \in I(F)\) and \(lm(f_{S,H}) = x_S\) holds for all term orders. So for all but \(|F|\) sets \(S \in Lm(I(F))\) 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 \(Sm(I(F))\) for all term orders, from which it follows by Corollary 3.2.1 that \(F\) 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 3.1** \(F \subseteq 2^{[n]}\) 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 \(I(F)\) for all term orders.

**Proof:** For the first part, suppose that \(F\) is extremal. Consider all minimal sets \(S \subseteq [n], S \notin Sh(F)\) with the corresponding unique polynomials \(f_{S,H}\). Denote the set of these sets by \(S\) 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 \(I(F)\). In order to show this we have to prove, that for all monomials \(m \in Lm(I(F))\), there is a monomial in \(\{x_S, S \in S\} \cup \{x_i^2, i \in [n]\}\) that divides \(m\). If there is a variable in \(m\) with degree higher then 1, then this is trivial. Since \(F\) is extremal, we have \(Sm(I(F)) = Sh(F)\), and this, together with the minimality of the sets in \(S\), 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 \(S\). Since the leading monomial of \(f_{S,H}\) is \(x_S\) for all term orders, \(Lm(G) = \{x_S, S \in S\} \cup \{x_i^2, i \in [n]\}\). This fact, together with the properties of a Gröbner basis, imply that \(Sm(G) = \{x_F, F \subseteq [n], \exists S \in S\) such that \(S \subseteq F\}\) for all term orders. So \(Sm(I(F))\) is the same for all term orders, which means by Corollary 3.2.1 that \(F\) is extremal. ■

### 4 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 3.1. The second one is simple as well, moreover it has a very good running time.
4.1 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 [5]). Suppose, that $\mathcal{F}$ is extremal. In this case according to Corollary 3.2.1 $Sm(I(\mathcal{F}))$ is the same for all term orders, so for $\prec$ in particular, we have $Sm(I(\mathcal{F})) = Sh(\mathcal{F})$. From Theorem 3.1 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 $S$. For every $S \in S$ 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 polynomials 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 3.1. There are many ways to verify whether a system of polynomials is a Gröbner basis or not. For such methods see [7]. I have not analyzed the time requirement of this method yet, however, it does not seem to be sufficiently efficient.

4.2 Test #2

According to the lex game [5] 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 4.1** 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 $\mathbb{F}[x]/I(\mathcal{F})$,
and since all functions from $F$ to $F$ are polynomials, every leading monomial can be written uniquely as the sum of standard monomials, as a function on $F$. This holds for $x_H$ as well:

$$x_H = \sum x_G,$$

as functions on $F$. 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 $F$, there is an $F \in F$ such that $G_0 = F \cap H$. For this $x_{G_0}(v_F) = 1$. From the minimality of $G_0$ we have that $x_{G'}(v_F) = 0$ for every other $G'$. So

$$\sum \alpha_G x_G(v_F) = \alpha_{G_0}.$$  

On the other hand $x_H(v_F) = 0$, since $H \cap F = G_0$, but $H \neq G$ because $x_H$ is a leading monomial, and $x_{G_0}$ is a standard monomial, and this is a contradiction. Therefore in the above sum there is a set $G$ with nonzero coefficient, such that $G \setminus H \neq \emptyset$. Now let us fix an index $i \in G \setminus H$. 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. $lacksquare$

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 $O(nm)$ time for one lexicographic term order, and we need $n$ term orders, the total running time of the algorithm is $O(n^2m)$.

**Corollary 4.1.1** Given a set family $F \subseteq 2^{[n]}$, $|F| = m$ by a list of characteristic vectors, we can decide in $O(n^2m)$ time if $F$ 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

$Sh(F)$ is a down-set, and it is evident that if $F$ is a down-set, then $Sh(F) = F$, i.e. $F$ 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.


5.1 Bit flip

For vectors \((v_1, \ldots, v_n) \in \{0, 1\}^n\) we denote by \(\varphi_i\) the \(i\)th bit flip:
\[
\varphi_i(v_1, \ldots, v_i, v_{i+1}, \ldots, v_n) := (v_1, \ldots, v_{i-1}, 1 - v_i, v_{i+1}, \ldots, v_n).
\]
It is easy to verify that extremality is invariant with respect to this operation. But 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 \(F = \{\emptyset, \{1\}, \{2\}, \{1, 3\}\}\)
\[
\Sh(F) = \{\emptyset, \{1\}, \{2\}, \{3\}\},
\]
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.

5.2 Translation

Let \(v\) be a fixed \(0-1\) vector of length \(n\). Then in the translation by \(v\) we add to all characteristic vectors \(v \mod 2\). This corresponds to the compositions of several bit flips, hence this operation also preserves extremality. We denote the translation by the vector \(v\) by \(\varphi_v\).

5.3 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 \(F_1\) and \(F_2\) be two set systems with disjoint supports (there is no \(i \in [n]\) for which \(\exists F_1 \in F_1\) and \(F_2 \in F_2\) such that \(i \in F_1 \cap F_2\)) and \(\emptyset \in F_1 \cap F_2\). We define the sum of these set systems as
\[
F = F_1 + F_2 := F_1 \cup F_2.
\]

Proposition 5.1 \(F\) is extremal if and only if \(F_1\) and \(F_2\) are both extremal.

Proof: \(F_1\) and \(F_2\) have disjoint supports, hence \(\Sh(F) = \Sh(F_1) \cup \Sh(F_2)\) and \(\Sh(F_1) \cap \Sh(F_2) = \{\emptyset\}\), which means that
\[
|\Sh(F)| = |\Sh(F_1)| + |\Sh(F_2)| - 1. \tag{1}
\]
On the other hand \(F_1 \cap F_2 = \{\emptyset\}\), hence
\[
|F| = |F_1| + |F_2| - 1. \tag{2}
\]
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:
\[ F = \{ \emptyset, \{2\} \} + \varphi_1(\{\emptyset, \{1\}, \{3\}\}) \]

**Open question 2** Can every extremal family \( F \) 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 \( F \subseteq 2^{[n]} \) then:

**Definition 6.1**

\[
D_i(F) := \{ F \in F | i \notin F \} \cup \{ F \in F | i \in F, F \setminus \{i\} \in F \} \cup \{ F \setminus \{i\} | F \in F, i \in F, F \setminus \{i\} \notin F \}.
\]

If in this definition we consider \( F \) as a set of bit vectors, one can say that \( v \in D_i(F) \) if and only if there are \( v_i + 1 \) elements \( s \in \{0, 1\} \) such that \((v_1, \ldots, v_{i-1}, s, v_{i+1}, \ldots, v_n) \in F \). From the definition it is clear that \(|F| = |D_i(F)|\) and if \( F \) is a down-set, then a downshift has no effect on \( F \). Moreover, it is easily seen that downshifts and bit flips commute:

\[
\varphi_i(D_j(F)) = D_j(\varphi_i(F))
\]

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

**Proposition 6.1** (See [4].) For every \( i \in [n] \) we have \( \text{Sh}(D_i(F)) \subseteq \text{Sh}(F) \).

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

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

**Proof:** From the extremality of \( F \) and from the previous proposition, for every \( i \in [n] \) we have:

\[
|F| = |\text{Sh}(F)| \geq |\text{Sh}(D_i(F))| \geq |D_i(F)|.
\]

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

Let \( F \subseteq 2^{[n]} \). For the indices \( i_1, i_2, \ldots, i_l \) we introduce the following notation:
\[ D_{i_1,i_2,...,i_l}(F) := D_{i_l}(D_{i_{l-1}}(\ldots(D_{i_1}(F)))) \]

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

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 \( v \in \{0, 1\}^n \) and \( \mathcal{F}_1, \mathcal{F}_2 \subseteq 2^{[n]} \)
\[ \varphi_v(\mathcal{F}_1) = \varphi_v(\mathcal{F}_2) \iff \mathcal{F}_1 = \mathcal{F}_2 \]

If \( \mathcal{F} \) is not empty, then by a composition \( \varphi_v \) of some bit flips we can achieve that \( \emptyset \in \mathcal{F} \). Also,
\[ D_{1,2}(\mathcal{F}) = D_{2,1}(\mathcal{F}) \iff \varphi_v(D_{1,2}(\mathcal{F})) = \varphi_v(D_{2,1}(\mathcal{F})) \iff \]
\[ D_{1,2}(\varphi_v(\mathcal{F})) = D_{2,1}(\varphi_v(\mathcal{F})). \]

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
\[ \mathcal{F} = \{\emptyset, \{1\}, \{1, 2\}\} \]

13
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

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

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 \( F \setminus \{i\} \in \mathcal{F} \).

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:** 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 \( v \) of the from \((v_2, v_3, \ldots, v_n)\) from \( D_2(\mathcal{F}) \), then \((0, v_2, v_3, \ldots, v_n) \in D_2(\mathcal{F}) \) must hold. Since \( v \in D_2(\mathcal{F}) \), there are \( v_2+1 \) elements \( s \in \{0, 1\} \) such that \((1, s, v_3, \ldots, v_n) \in \mathcal{F} \). Since \( \mathcal{F} \) is an 1-down-set, we must have \((0, s, v_3, \ldots, v_n) \in \mathcal{F} \) for the \( v_2 + 1 \) elements \( s \), what implies that \((0, v_2, v_3, \ldots, v_n) \in D_2(\mathcal{F}) \). ■

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 \( \mathcal{F}_0 \) which is an \( i \)-down-set and so \( D_i(\mathcal{F}_0) = \mathcal{F}_0 \) for all \( i \in [n] \). This last property is equivalent to the fact that \( \mathcal{F}_0 \) is a down-set. ■

The preceding statement occurs in [4], 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
Proposition 1.1, the size of $Sh$ of $F \subseteq \text{Definition 6.3}$ gives the idea of the following definition: extremal set system.) However, if $F$ is extremal then, by Proposition 6.1 and Proposition 1.1, the size of $Sh$ cannot be decreased by downshifts. This fact gives the idea of the following definition:

**Definition 6.3** $F \subseteq 2^n$ is weakly extremal if for every $i \in [n]$ we have $Sh(D_i(F)) = Sh(F)$.

In Section 7 we discuss the connection between extremality and weak extremality.

At the very end of Section 2 we already discussed a recursive method for constructing the $Sm(F)$ for a set system $F \subseteq 2^n$. Now we give another method using the downshift operation.

**Theorem 6.1** Let $F \subseteq 2^n$ 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 $F$ in this order, then we have $D_{i_n,i_{n-1},\ldots,i_1}(F) = Sm(I(F))$.

**Proof:** Without loss of generality we can suppose that $\prec$ is based on the natural lex order, thus $i_k = k$. We prove the statement by induction on $n$. For $n = 1$ the statement can be easily verified for all possible set systems $F$. Now let us suppose that the statement is true for all values smaller than $n$, and consider the case of $n > 1$. Let

$$F_0 = \{F : F \in F; n \notin F\},$$
$$F_1 = \{F \setminus \{n\} : F \in F; n \in F\}.$$

Now apply $D_{n-1,n-2,\ldots,1}$, and let

$$H = D_{n-1,n-2,\ldots,1}(F),$$

and

$$H_0 = \{H : H \in H; n \notin H\},$$
$$H_1 = \{H \setminus \{n\} : H \in H; n \in H\}.$$

Clearly for $i \in \{0, 1\}$, $H_i = D_{n-1,n-2,\ldots,1}(F_i)$. Now apply $D_n$. According to Proposition 6.4, $H_i$ is a down-set for $i = 0, 1$. It is easy to see that

$$D_n(H) = H_0 \cup H_1 \cup \{H \cup \{n\} : H \in H_0 \cup H_1\}.$$

At the very end of Section 2 we gave a recursive method for constructing $Sm(I(F))$ from $F_0$ and $F_1$,

$$Sm(I(F)) = Sm(I(F_0)) \cup Sm(I(F_1)) \cup \{D \cup \{n\} : D \in Sm(I(F_0)) \cap Sm(I(F_1))\}.$$ 

But since $F_i \subseteq 2^{[n-1]}$, according to the induction hypothesis $Sm(I(F_i)) = H_i$. Thus

$$Sm(I(F)) = H_0 \cup H_1 \cup \{H \cup \{n\} : H \in H_0 \cup H_1\} = D_n(H) = D_{n,n-1,\ldots,1}(F).$$
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 \( 0 \rightarrow 1 \) vectors of length \( n \), so \( \mathcal{F} \subseteq \{0,1\}^n \).

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 \( D_{i_1,i_2,...,i_m}(F) \) is isometrically embedded for every \( m \) and \( i_1,i_2,...,i_m \in [n] \).

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 \( C \subseteq B \). 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 1.1.1 to the elements of \([n] \setminus B\) 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 $D_{i_1,i_2,\ldots,i_m}(\mathcal{F})$ 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 [4] 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 $S \in \text{Sh}(\mathcal{F})$ and $S \notin \text{Sh}(D_i(\mathcal{F}))$ (i.e. if the downshift with $i$ reduces $\text{Sh}(\mathcal{F})$), 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 $F \setminus \{i\}$. This is possible only if $F \setminus \{i\} \notin \mathcal{F}$, 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 $S \cap F' = H \setminus \{i\}$. 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 $G \setminus \{i\}$ 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 $i \in [n]$, $\text{Sh}(\mathcal{C}) = \text{Sh}(D_i(\mathcal{C}))$.

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 $i \in [n]$, $\text{Sh}(D_i(\mathcal{F})) \subseteq \text{Sh}(\mathcal{F})$, it follows that for every $i \in [n]$, $\text{Sh}(D_i(\mathcal{F})) = \text{Sh}(\mathcal{F})$. But it is clear that if $\mathcal{F}$ is isometrically embedded in $Q_n$. 17
then so is every chunk of it. So we have \( \text{Sh}(C) = \text{Sh}(D_i(C)) \) for every \( i \in [n] \) and for every chunk \( C \) as well. For the other direction recall the proof of Proposition 7.1 and as we did there, suppose that \( F \) is not isometrically embedded in \( Q_n \). If we follow that proof, we obtain a chunk \( H \) of \( F \), such that \( H = \{ \emptyset, B \} \) with \( |B| \geq 2 \). For this chunk we have \( \text{Sh}(H) = \{ \emptyset, \{ b_1 \}, \ldots, \{ b_k \} \} \). But if we apply a downshift to \( H \) by an index \( i \in B \), then \( D_i(H) = \{ \emptyset, B \setminus \{ i \} \} \) and therefore \( \text{Sh}(D_i(H)) = \text{Sh}(H) \setminus \{ \{ i \} \} \). So we have found a chunk \( H \) such that \( \text{Sh}(D_i(H)) \neq \text{Sh}(H) \), but this is a contradiction, so \( F \) must be isometrically embedded. ■

We have already mentioned that extremality is equivalent to the fact that \( F \) is strongly isometrically embedded in \( Q_n \), thus the extremality of \( F \) does not follow from the fact that \( F \) is isometrically embedded in \( Q_n \). We will demonstrate this on some examples. Let us denote by \( F_k \) the set of all subsets of size \( k \) (and similarly the set of vectors \( v \in \{0,1\}^n \) with \( k \) ones).

**Example 7.1** \( F = F_{k-1} \cup F_k \) for \( 2 \leq k \leq \lceil \frac{n}{2} \rceil \) is isometrically embedded in \( Q_n \), but is not extremal.

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

\[
\text{Sh}(F) = F_0 \cup F_1 \cup \cdots \cup F_k.
\]

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

The previous example can be generalized:

**Example 7.2** Let \( 0 < a < b < n \) be integers and \( F = F_a \cup F_{a+1} \cup \cdots \cup F_b \). We claim that \( F \) 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

\[
\text{Sh}(F) = F_0 \cup F_1 \cup \cdots \cup F_b,
\]

therefore \( F \) is not extremal. It is straightforward to see that \( F \) is isometrically embedded in \( Q_n \). ■

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

The Vapnik-Chervonenkis (VC) dimension is a widely known and used notion in mathematics.

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

As an example of its application, consider [10]. This says that the problem of computing the VC-dimension is in $\text{SAT}^{\log_2 n}$, 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 $\text{SAT}^{\log_2 n}_{\text{CNF}}$ (as $\text{SAT}^{\log_2 n}$, only with inputs in conjunctive normal form). This section is about the problem of computing $\text{VC} - \text{dim}(F)$ for a set system $F \subseteq 2^{[n]}$.

Proposition 8.1 (See [10].) For any set system $F \subseteq 2^{[n]}$, $\text{VC} - \text{dim}(F) \leq \log |F|$. (Here $\log$ stands for the logarithm with base 2.)

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

By this proposition, the simple algorithm for computing the VC-dimension of a set system $F$ which enumerates all possible sets to be shattered, shall terminate in $O(m \log |F|)$ time, where $m$ is the size of $F$ (see [11]). We give another algorithm with the same time bound. First let us recall that

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

From the proof of Proposition 3.2 we know that if we take a set $S$ from $\text{Sh}(F)$ then for the term order where the variables from $x_S$ are the smallest we have $S \in \text{Sm}(F)$ (i.e. $x_S \in \text{Sm}(I(F))$). 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 $\text{Sh}(F)$, and consequently $\text{VC} - \text{dim}(F)$, it is enough to compute $\text{Sm}(I(F))$ 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 $O(nm^{\log m})$.

9 Conclusion and future work

We have seen that algebraic methods can be very useful when studying combinatorial objects. We achieved several results. In Section 3 we gave a new characterization for extremal set systems using Gröbner bases. Then in Section 4 we proposed an efficient, $O(n^2 m)$, 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. In Section 7 we discussed the graph theoretical consequences of extremality, and characterized the event when downshifting makes \( \text{Sh}(F) \) shrink. Finally, in Section 8, we proposed a new algorithm for solving the problem of computing the Vapnik-Chervonenkis dimension of a set system \( F \subseteq 2^n \).

There are many ways to follow up this study. These include finding (slightly) faster (possibly linear time) algorithms for testing extremality or improving the \( O(mn \log m) \) time bound of the algorithm for computing the VC-dimension of a set system.

References


