Probabilistic analysis on Macaulay matrices over finite fields and complexity of constructing Gröbner bases

Igor Semaev and Andrea Tenti

igor.semaev@uib.no, andrea.tenti@uib.no
Department of Informatics, University of Bergen, Bergen, Norway

Abstract

Gröbner basis methods are used to solve systems of polynomial equations over finite fields, but their complexity is poorly understood. In this work an upper bound on the time complexity of constructing a Gröbner basis according to a total degree monomial ordering and finding a solution of a system is proved. A key parameter in this estimate is the degree of regularity of the leading forms of the polynomials. Therefore, we provide an upper bound to the degree of regularity for a sufficiently overdetermined system of forms of the same degree over any finite field. The bound holds for almost all polynomial system and depends only on the number of variables, the number of polynomials, and the degree. Our results imply that almost all sufficiently overdetermined systems of polynomial equations of the same degree are solvable in polynomial time.

Keywords — Polynomial equation systems, finite fields, Macaulay matrices, Groebner bases, multisets, complexity

1 Introduction

Let $x_1, \dots, x_n$ be variables over a field. Systems of polynomial equations

$$P_1(x_1, \dots, x_n) = 0, \dots, P_m(x_1, \dots, x_n) = 0 \tag{1}$$

are a main object to study in algebraic geometry and commutative algebra. Several methods to find an explicit solution to (1) were developed. In particular, Macaulay [Mac16] introduced multivariate resultants and used them to solve systems of polynomial equations by eliminating variables. The resultant of a system is the determinant of a matrix, obtained from the coefficients of the polynomials. Later, this construction was generalised so that a system has one such matrix for every degree (see e.g. [BFS15; CG17]).

These matrices are generally called Macaulay matrices and can be viewed as generalisations of the Sylvester Matrix, which is defined for two univariate polynomials [Syl53]. Buchberger [Buc65] defined the notion of a Gröbner basis for a polynomial ideal and showed how to construct such a basis. In some cases a solution to (1) may be instantly read from a reduced Gröbner basis. Lazard [Laz83] showed that a Gröbner basis according to a total degree monomial ordering may also be constructed by triangulating a suitable Macaulay matrix.

For a finite ground field $\mathbb{F}_q$, two problems are of special interest: how many $\mathbb{F}_q$-rational solutions does the system allow, and how do we compute them? The number of solutions may be estimated using the Lang-Weil bound [LW54]. The second problem is reducible to a satisfiability problem and is generally NP-hard.

Applications in cryptography renewed interest in solving polynomial equations over finite fields. Finding a solution is equivalent to breaking a cryptosystem. A particularly successful example, due to Faugère and Joux [FJ03], broke the HFE (Hidden Field Equations) cryptosystem with a Gröbner basis algorithm.

In some applications the problem may be reduced to overdetermined polynomial systems, where the number of equations m is larger than the number of variables n. For instance, one has to solve an overdetermined quadratic equation system over $\mathbb{F}_2$ to find an AES key given some plain-text and relevant cipher-text, [CP02]. In practice, among equation systems of the same degree, those which are overdetermined may be solved faster than those where $m \leq n$ using algorithms from Gröbner basis of XL families [BFS03; CKPS00]. Hence, it is interesting to study the time-complexity of those algorithms for overdetermined polynomial equation systems. However, the theoretical complexity of the Buchberger algorithm [CLO13] for constructing Gröbner bases over finite fields, and its well-known variations as F4 [Fau99] and F5 [Fau02], in general is unknown.

In order to avoid solutions in the extensions of the ground field we need to add $x_i^q - x_i, i = 1, \dots, n$ to the system (1). So we may assume that every monomial $x_1^{e_1} \dots x_n^{e_n}$ in (1) with a non-zero coefficient satisfies $e_i < q$ for every i. In this work a new algorithm to construct a Gröbner basis according to a total degree monomial ordering for (1) is presented. Its complexity is rigorously estimated through the degree of regularity for the leading forms of the polynomials.

Let $f_1, \dots, f_m$ be the leading forms of the polynomials $P_1, \dots, P_m$ and let I be an ideal in $R^h = \mathbb{F}_q[x_1, \dots, x_n]/(x_1^q, \dots, x_n^q)$ generated by $f_1, \dots, f_m$. By I_d we denote a vector space over $\mathbb{F}_q$ containing all forms in I of degree d. The degree of regularity of $I \neq 0$ is the smallest integer $d \geq 0$ for which $\dim_{\mathbb{F}_q} I_d = l_q(n, d)$, the number of monomials in R^h of total degree d, and it is denoted d_{reg}. It is easy to see that d_{reg} exists for every non-zero ideal in R^h generated by forms and $d_{\text{reg}} \leq (q-1)n$. The expression "the degree of regularity" was first used in [BFS03] and it is also called the index of R^h/I in [HMS17].

Assume that the polynomials (1) are linearly independent and of degree at most d_{reg}. If not, then the statement below is easy to adjust. In Theorem 2.1 we show that the time-complexity of constructing a total degree Gröbner basis for (1) is polynomial in $L_q(n, d_{\mathrm{reg}})$, where $L_q(n,d)$ is the number of monomials in R^h of total degree at most d. At least one solution to (1) over $\mathbb{F}_q$, if it exists, may then be computed faster than the Gröbner basis according to Theorem 2.2.

The notion of a semiregular system of polynomials (forms) was introduced by Bardet, Faugère, and Salvy in [BFS03]. The degree of regularity for a particular semiregular polynomial system may be computed by expanding a Hilbert series defined by n,m, and the degree of P_i. It was also conjectured that a random system of polynomials over $\mathbb{F}_2$ is semiregular with probability tending to 1 as n increases. The conjecture, in that form, was disproved in [HMS17]. Still it is believed that most systems behave like semiregular ones.

The present work gives an upper bound to the degree of regularity for an overdetermined system of forms $f_1, \ldots, f_m$ of the same degree D with coefficients in $\mathbb{F}_q$ taken uniformly at random. The bound holds with probability tending to 1; in other words, the bound holds for almost all such systems of forms for sufficiently large n. We do not impose any other restrictions on the polynomials such as semiregularity.

Theorem 1.1. *Let $q \geq 2$ and let D be fixed, and let $m \geq l_q(n, D{+}d)/l_q(n,d)$, where $D > d > 0$. Then*

$$\mathbb{P}(d_{reg} \leq D + d) \geq 1 - q^{l_q(n,D+d)-ml_q(n,d)} - A(n,D,d),$$

where $A(n,D,d) > 0$ and $A(n,D,d) = O(n^d q^{-Cn^D})$ for a positive constant C as $n \to \infty$.

The theorem implies that if $m \geq l_q(n, D+d)/l_q(n,d) + c$ for a positive constant c, then $\mathbb{P}(d_{\mathrm{reg}} \leq D{+}d) \geq 1 - q^{-cl_q(n,d)} - A(n,D,d) \to 1$ as $n \to \infty$.

Let $q = 2$. It is well known and easy to prove that almost all systems of $m \geq \frac{n(n-1)}{2} + c$ quadratic polynomials have $d_{\mathrm{reg}} = 2$.

Theorem 1.1 for $D = 2, d = 1$ implies that almost all systems of $m \geq \frac{(n-1)(n-2)}{6} + c$ quadratic polynomials have $d_{\mathrm{reg}} \leq 3$. Similarly, for $D = 3, d = 2$, almost all systems of $m \geq \frac{(n-2)(n-3)(n-4)}{60} + c$ cubic polynomials have $d_{\mathrm{reg}} \leq 5$, etc. According to Theorems 2.1 and 2.2, a total degree Gröbner basis and a solution to a relevant equation system may be then computed in polynomial time. In fact, our complexity bounds depend on the leading forms of the polynomials and do not depend on their lower degree terms.

Over $\mathbb{F}_2$ the bound on d_{reg} is as predicted in [BFS03] for a semiregular system with the same parameters (number of variables n, number of equations m, and of degree D). Modulo a conjecture from commutative algebra, a lower bound on the degree of regularity for homogeneous polynomial systems in $\mathbb{F}_q[x_1, \dots, x_n]$ is proved in [Die04]. Our result satisfies this bound as well.

The core of the proof of Theorem 1.1 is in Section 4, where we show that a Macaulay matrix of size $m\, l_q(n, d) \times l_q(n, d+D)$ constructed for the forms $f_1, \dots, f_m$ has linearly independent columns with probability tending to 1. The rows of the matrix are coefficients of the leading forms of $gf_i \in R^h$, where g runs over all monomials of degree d. For instance, let $n = 4, m = 1, D = 2$ and $d = 1$, and

$$f_1 = c_{12}x_1x_2 + c_{13}x_1x_3 + c_{14}x_1x_4 + c_{23}x_2x_3 + c_{24}x_2x_4 + c_{34}x_3x_4$$

over $\mathbb{F}_2$. Then the Macaulay matrix is

$$M = \begin{pmatrix} c_{23} & c_{24} & c_{34} & 0 \\ c_{13} & c_{14} & 0 & c_{34} \\ c_{12} & 0 & c_{14} & c_{24} \\ 0 & c_{12} & c_{13} & c_{23} \end{pmatrix}$$

and $\det M = c_{12}c_{34} + c_{13}c_{24} + c_{14}c_{23}$. So if the coefficients of f_1 are chosen uniformly and independently, then the columns of M are linearly independent with probability $28/64$, that is $\mathbb{P}(d_{\text{reg}} = 3) = 28/64$.

Section 3 contains an auxiliary combinatorial result used in the proof of the main Theorem 1.1. Each monomial $x_1^{a_1} \dots x_n^{a_n}$ of total degree d defines a d-multiset $(a_1, \dots, a_n)$, where $0 \le a_i < q$ and $\sum_{i=1}^n a_i = d$. Let v be a natural number and A a family of monomials of degree d such that $|A| = v$. By B we denote a family of monomials of degree $d + D$ divisible by at least one monomial from A. Theorem 3.1 implies that $|B|$ achieves its minimum when A is a family of the first (largest) v monomials of total degree d taken in a lexicographic order.

After this paper was submitted for possible publication, we realised that a statement equivalent to Theorem 3.1 was already proved in 1969 by Clements and Lindström [CL69]. The equivalent result is Corollary 1 in their paper. However, our proof is fundamentally different from theirs and we believe that it provides further insight into the problem.

Theorem 2.1 was proved by Semaev. The main idea of the proof of Theorem 1.1 belongs to Semaev too, who first proved it for $\mathbb{F}_2$ and $D = 2, d = 1$. The generalisation for every $\mathbb{F}_q$ and $D > d$ is due to Tenti, who also proved Theorem 2.2. Tenti conjectured the statement of Theorem 3.1 for $k = k_1 = \dots = k_n$ and proved it for $k = 1, d = 2$. With a different method, presented in Section 3, the theorem in its generality was proved by Semaev.

An extended abstract of this paper was presented at WCC2019 [ST19].

2 Complexity of constructing Gröbner bases

Let

$$I = (P_1, \ldots, P_m, x_1^q - x_1, \ldots, x_n^q - x_n) \tag{2}$$

be an ideal in $\mathbb{F}_q[x_1, \ldots, x_n]$ and $R = \mathbb{F}_q[x_1, \ldots, x_n]/(x_1^q - x_1, \ldots, x_n^q - x_n)$. Let $N = L_q(n, d_{\mathrm{reg}})$, the number of monomials in R^h of degree $\le d_{\mathrm{reg}}$ as above.

In this section we show how to construct a Gröbner basis for I for a total degree monomial ordering and rigorously estimate the complexity of the construction in arithmetic operations in $\mathbb{F}_q$, where both the parameters n, q may grow. The new algorithm incorporates three stages.

1. Compute d_{reg} for the leading forms $f_1, \ldots, f_m$ of $P_1, \ldots, P_m$. For every monomial h of degree d_{reg}, compute the forms $t_1, \ldots, t_m$ such that

   $$h = t_1 f_1 + \ldots + t_m f_m$$

   in R^h, where $\deg t_i = d_{\mathrm{reg}} - \deg f_i$ or $t_i = 0$. Compute $g = t_1 P_1 + \ldots + t_m P_m$ in R. The degree of g is d_{reg} and its leading form is the monomial h. One adds g to the initial basis of I and gets a new basis $\{U_1, \ldots, U_r, x_1^q - x_1, \ldots, x_n^q - x_n\}$ for I in $\mathbb{F}_q[x_1, \ldots, x_n]$.

2. Compute a basis B of the ideal I with the following properties. First, $\deg g \le d_{\mathrm{reg}}$ for every $g \in B$. Second, B incorporates $l_q(n, d_{\mathrm{reg}})$ polynomials $g = h + f$ such that $\deg f < d_{\mathrm{reg}}$ and their leading forms h are all possible monomials of degree d_{reg}.

3. A Buchberger algorithm application to B gives a Gröbner basis for I.

The theoretical complexity of the algorithm is a function in d_{reg} computed for the leading forms of the polynomials as this is explained below. Testing the algorithm on large instances and comparison with the plain Buchberger algorithm or F4 as implemented in Magma [BCP97] is out of scope of this paper.

In order to simplify the argument and the complexity bound in Theorem 2.1 the polynomials $P_1, \ldots, P_m$ are assumed linearly independent and of degree $\le d_{\mathrm{reg}}$. Otherwise, the statement is easy to adjust. So $m \le N$ and one may assume that there are at most $l_q(n, k)$ forms of degree k among $f_1, \ldots, f_m$. To compute d_{reg}, one gradually triangulates with elimination Macaulay matrices for the forms f_i multiplied by monomials of degree $d - \deg f_i$ in R^h for $d \le d_{\mathrm{reg}}$. The number of rows is at most

$$\sum_{k=1}^{d} l_q(n, k) l_q(n, d-k) \le dN^2$$

and the number of columns is $l_q(n,d) \le N$. It takes $O(dN^4)$ operations in $\mathbb{F}_q$ to triangulate the matrix. The cost for all $d \le d_{\text{reg}}$ is $O(d_{\text{reg}}^2 N^4)$ operations. Within the same cost one constructs the polynomials $U_1, \dots, U_r$ of degree $\le d_{\text{reg}}$ according to the algorithm. Exactly $l_q(n, d_{\text{reg}})$ polynomials U_i are of degree d_{reg} and their leading forms are all possible monomials of degree d_{reg}. The polynomials $\{U_1, \dots, U_r, x_1^q - x_1, \dots, x_n^q - x_n\}$ give a basis for I in $\mathbb{F}_q[x_1, \dots, x_n]$.

If $q \le d_{\text{reg}}$, then that basis is B. Let $q > d_{\text{reg}}$. One replaces each $x_i^q - x_i$ in the basis with its residue after division by the polynomials $U_j = h + f$, where h is a monomial of degree d_{reg} and $\deg f < d_{\text{reg}}$. That produces the basis B. When computing the residue of $x_i^q - x_i$, it might be that the intermediate polynomials after each division by such U_j incorporate only monomials $x_i^b x_1^{a_1} \dots x_n^{a_n}$, where $b < q$ and $\sum_{j=1}^n a_j < d_{\text{reg}}$. So the number of monomials at each division step is at most qN. The division cost is $O(qN^2)$ for each $x_i^q - x_i$ and $O(nqN^2)$ overall.

However, B is not generally a Gröbner basis for I. For instance, the polynomial system

$$P_1 = x_1x_2 + 1, \quad P_2 = x_1x_3, \quad P_3 = x_2x_3, \quad x_1^2 - x_1, \quad x_2^2 - x_2, \quad x_3^2 - x_3$$

from $\mathbb{F}_2[x_1, x_2, x_3]$ has $d_{\text{reg}} = 2$. However, that is not a Gröbner basis, as the ideal contains the polynomial $x_3 = x_3P_1 + x_2P_2$ and its leading term is not divisible by the leading terms of the basis. So the argument in [BFS03, Section 2.2] on the complexity of constructing a Gröbner basis is not valid. In order to compute a Gröbner basis one generally has to work with polynomials of degree $> d_{\text{reg}}$ as well. The following theorem, in particular, proves that with the basis B one can construct a Gröbner basis for I at maximum degree $\le 2d_{\text{reg}}$ by an application of the Buchberger algorithm. We estimate the complexity of the construction.

Theorem 2.1. *Time-complexity of constructing a Gröbner basis for I is polynomial in N and q.*

Proof. One can make the polynomials in B linearly independent. It is then enough to prove that an application of the Buchberger algorithm to the polynomials B takes $O(N^6)$ operations in $\mathbb{F}_q$. For each $Q_1, Q_2 \in B$, the algorithm computes a residue T of the S-polynomial $S(Q_1, Q_2)$ after division by the polynomials B. Each monomial of degree d_{reg} occurs as a leading monomial of some polynomial in B, so the degree of the residue is less than d_{reg}. If $T \ne 0$, then B is augmented with T and the step repeats. If the residue is 0 for each pair, then B is a Gröbner basis. At each step of the algorithm the polynomials in B are linearly independent.

One has to examine $\le N^2$ pairs before finding a non-zero residue or terminating. The number of possible linearly independent residues is $\le N$, so the number of divisions is $\le N^3$. The number of intermediate monomials is $\leq N^2$ when an S-polynomial is divided by the polynomials from B. So computing its residue takes $O(N^3)$ operations. Overall complexity is that stated. $\square$

A more careful analysis shows that one can work with polynomials of degree $\leq 2d_{\mathrm{reg}} - 2$ and the time-complexity is

$$O(N^2\, L_q^2(n, d_{\mathrm{reg}} - 1)\, L_q(n, 2d_{\mathrm{reg}} - 2))$$

operations.

2.1 From a Gröbner basis to a solution of the system

Let $Z(I) \subseteq \mathbb{F}_q^n$ be the set of zeroes for the ideal I defined by (2). Here we show how to compute $(a_1, \ldots, a_n) \in Z(I)$ and estimate the time complexity. Let G be a Gröbner basis for I according to a fixed total degree ordering computed as above. Then $\deg(g) \leq d_{\mathrm{reg}}$ for every $g \in G$.

Theorem 2.2. *One can compute $(a_1, \ldots, a_n) \in Z(I)$ or prove $Z(I) = \emptyset$ in $O(nN^3)$ operations.*

Proof. If G' is a reduced Gröbner basis for I according to the total degree ordering, then $G' = \{1\}$ if and only if $Z(I) = \emptyset$. So the algorithm we employ is the following. First, we compute the reduced Gröbner basis G' of I. If $G' = \{1\}$, then the system has no solutions. Otherwise, we take $a_n \in \mathbb{F}_q$ and compute the reduced Gröbner basis G' of $I + (x_n - a_n)$. If $G' = \{1\}$, then we take another a_n and compute the reduced Gröbner basis, etc. Otherwise, if $G' \neq \{1\}$, we replace I with $I + (x_n - a_n)$ and repeat the previous step. This repeats until a solution $(a_1, \ldots, a_n)$ is found.

Obviously, the algorithm produces a zero of I if it exists or proves $Z(I) = \emptyset$. One has to compute up to qn reduced Gröbner bases of ideals $I + (x_n - a_n)$. We will now prove that it is possible to compute the reduced Gröbner basis in $O(N^3)$ operations at any step. Let LT denote the leading term of a polynomial or the set of the leading terms of a set of polynomials.

According to [KR00], to reduce G we first remove from G all g such that $LT(g) \in (LT(G \setminus \{g\}))$ and make the rest of the polynomials monic. As $|G| \leq N$, it takes $\leq N^2$ monomial divisions. We call the new set G', which is still a Gröbner basis for I. Next, for every $g \in G'$ one computes its residue g' after division by $G' \setminus \{g\}$ and sets $G' = G' \setminus \{g\} \cup \{g'\}$. As every polynomial in G' incorporates $\leq N$ monomials, computing the residue takes $O(N^2)$ operations. Since $LT(g) = LT(g')$, once an element is modified, it does not change further. The overall cost is $O(N^3)$ operations.

Let $G = \{g_1, \ldots, g_t\}$ be a reduced Gröbner Basis for I according to a fixed total degree ordering constructed as above and let $I_{\leq d}$ denote the space of polynomials in I of degree $\leq d$.

Lemma 2.3. The set of polynomials $x^\alpha g_i$ such that $|\alpha| + \deg(g_i) \leq d$ generates $I_{\leq d}$ as a vector space over $\mathbb{F}_q$.

The lemma follows directly from the properties of polynomial division and Gröbner basis. For $d = d_{\text{reg}}$ one can extract a basis for $I_{\leq d}$ from the generators in $O(N^2 \log N)$ operations by sorting their leading monomials.

Lemma 2.4. Let g be a linear polynomial. The vector space $(I + (g))_{\leq d_{\text{reg}}}$ is generated by $x^\alpha g_i$ and $x^\beta g$, with $|\alpha| + \deg(g_i) \leq d_{\text{reg}}$ and $|\beta| < d_{\text{reg}}$.

Proof. First we show that every $f \in (I+(g))$ may be represented as $f = p+gr$ for some $p \in I$ and r with $\deg(r) < d_{\text{reg}}$. Obviously, $f = f_1 + f_2 g$ with $f_1 \in I$, $f_2 \in \mathbb{F}_q[x_1, \ldots, x_n]$. Let r be a residue of f_2 after division by G. Then $f_2 = h + r$, where $h \in I$ and $\deg(r) < d_{\text{reg}}$. Hence $f = p + rg$, with $p = f_1 + gh \in I$.

Therefore, $f = p + gr$ is in $(I + (g))_{\leq d_{\text{reg}}}$ if and only if $\deg(p) \leq d_{\text{reg}}$. Hence

$$(I + (g))_{\leq d_{\text{reg}}} \subseteq I_{\leq d_{\text{reg}}} + (g)_{\leq d_{\text{reg}}}.$$

The first vector space is generated by $x^\alpha g_i$ with $|\alpha| + \deg(g_i) \leq d_{\text{reg}}$ thanks to Lemma 2.3. On the other hand, $(g)_{\leq d_{\text{reg}}}$ is trivially generated by $x^\beta g$ with $|\beta| + \deg(g) \leq d_{\text{reg}}$. The proof is complete. □

Corollary 2.5. Let $B = \{b_1, \ldots, b_k\}$ be a basis for the vector space $(I + (g))_{\leq d_{\text{reg}}}$. Then $G^+ = \{b_1, \ldots, b_k, g_1, \ldots, g_t\}$ is a Gröbner basis for $(I + (g))$.

Proof. Obviously, G^+ is a basis for $I + (g)$. Let $f \in I + (g)$. If $\deg(f) \leq d_{\text{reg}}$, then $LT(f) = LT(b_i)$ for some $b_i \in B$. If $\deg(f) > d_{\text{reg}}$, then $LT(f)$ is divisible with some $LT(g_i)$ by the definition of d_{reg}.

Therefore, every leading term of $f \in I + (g)$ is divisible by the leading term of one of the elements in G^+. Hence the latter is a Gröbner basis for $I + (g)$. □

We can now complete the proof of Theorem 2.2. In order to compute B, one triangulates a matrix with $\leq N$ columns and $\leq 2N$ rows (the first N are given by S and the second ones are given by $x^\beta g$). The size of G^+ is $\leq 2N$. So each computation of a reduced Gröbner basis that we perform has a cost of $O(N^3)$ operations. In order to find one zero in $Z(I)$, we need to perform at most qn iterations. Hence the total cost is $O(nN^3)$ as claimed. □

Remark 2.6. The algorithm just presented returns only one of the zeroes in $Z(I)$. The entire set can be found by using the Shape Lemma [KR00] after a linear change of coordinates in an extension of $\mathbb{F}_q$. This approach has the drawback that if the system has many solutions, then the extension is large. The full process is described in [KR00, Section 3.7].

3 Minimal covering family of multisets

Let $\{1, 2, \ldots, n\}$ be a set of n elements, equipped with the standard ordering $\leq$ and let $k_1, \ldots, k_n, d$ be non-zero integers. The tuple $X = (x_1, x_2, \ldots, x_n)$ is a d-multiset if $0 \leq x_i \leq k_i$ and $\sum_{i=1}^{n} x_i = d$. We say n is the length of X. The family of all d-multisets is denoted $\mathcal{X} = \mathcal{X}^d$.

Let $Y = (y_1, y_2, \ldots, y_n)$ be a D-multiset for some $D \geq d$. We say X is a subset of Y, denoted $X \subseteq Y$, if $x_i \leq y_i, 1 \leq i \leq n$. If $x_i + y_i \leq k_i$ for every i, one defines $X + Y = (x_1 + y_1, \ldots, x_n + y_n)$ and if $X \subseteq Y$, then $Y \setminus X = (y_1 - x_1, \ldots, y_n - x_n)$.

The reverse ordering on $\{1, 2, \ldots, n\}$ induces a lexicographic ordering $>$ on the family of all d-multisets $\mathcal{X}$. Let $\mathcal{X}_v = \{X_1, \ldots, X_v\}$ denote the family of the first (largest) v multisets according to that ordering, that is $X_1 > \ldots > X_v$. We call $\mathcal{X}_v$ a minimal family of size v. Let $\mathcal{Y} = \mathcal{Y}^D$ denote the lexicographically ordered family of all D-multisets. Then $\mathcal{Y}_u$ denote the family of the first (largest) u elements in $\mathcal{Y}$ according to the ordering. For instance, ordered 2- and 3-multisets ($d = 2$ and $D = 3$) of length 3, where $k_1 = k_2 = k_3 = 2$, are presented in Table 1.

Table 1: Ordered 2- and 3-multisets of length 3, with $k_1 = k_2 = k_3 = 2$

X_6	002
X_5	011
X_4	020
X_3	101
X_2	110
X_1	200

Y_7	012
Y_6	021
Y_5	102
Y_4	111
Y_3	120
Y_2	201
Y_1	210

By $Y_{\ell(v)}$ we denote the smallest D-multiset such that $Y_{\ell(v)} \supseteq X_v$ (we say covered by X_v). For instance, $\ell(2) = 4$ in Table 1. So $\mathcal{Y}_{\ell(v)} = \{Y_1, \ldots, Y_{\ell(v)}\}$ is the ordered family of $Y \geq Y_{\ell(v)}$ in $\mathcal{Y}$. Let $\mathcal{A} = \{X_{i_1}, \ldots, X_{i_v}\}$ be a family of d-multisets. By $||\mathcal{A}||$ we denote the number of D-multisets which contain at least one element from $\mathcal{A}$ (we say covered by $\mathcal{A}$). The goal of this section is to prove

Theorem 3.1. *If $k_1 \leq k_2 \leq \ldots \leq k_n$ and $|\mathcal{A}| = v$, then $||\mathcal{A}|| \geq ||\mathcal{X}_v|| = \ell(v)$.*

If the condition $k_1 \leq k_2 \leq \ldots \leq k_n$ is not satisfied, then the theorem is not generally true. For example, let $k_1 = 3$, $k_2 = 1$, $d = 1$, $D = 3$, $v = 1$. Then $(1, 0), (0, 1)$ are all 1-multisets and $(3, 0), (2, 1)$ are all 3-multisets ordered lexicographically. The family $\mathcal{A} = \{(0, 1)\}$ covers only $(2, 1)$, while the family $\mathcal{X}_1 = \{(1, 0)\}$ covers $(3, 0), (2, 1)$. So $||\mathcal{A}|| = 1$ and $||\mathcal{X}_v|| = 2$. The statement does not hold.

We will prove several lemmas first. We can assume that d is sufficiently large, otherwise the proofs below may be easily adjusted.

Lemma 3.2. The family of D-multisets covered by $\mathcal{X}_v$ is $\mathcal{Y}_{\ell(v)}$. In particular, $||\mathcal{X}_v|| = \ell(v)$.

Proof. Let $X \in \mathcal{X}_v$ or, in other words, $X \geq X_v$. First, we will prove that for every D-multiset $Y \supseteq X$ we have $Y \in \mathcal{Y}_{\ell(v)}$. If $X = X_v$, that holds by the definition of $\ell(v)$. If $X < X_v$, then there exists $i \in \{1, \dots, n-1\}$ such that

$$X_v = (x_1, \dots, x_{i-1}, x_i, \dots, x_n), \quad X = (x_1, \dots, x_{i-1}, x'_i, \dots, x'_n),$$

where $x'_i > x_i$ and

$$Y_{\ell(v)} = (y_1, \dots, y_{i-1}, y_i, \dots, y_n), \quad Y = (y'_1, \dots, y'_{i-1}, y'_i, \dots, y'_n).$$

Let $i > 1$. If $y'_1 > y_1$, then $Y > Y_{\ell(v)}$ and there is nothing to prove. Assume $y'_1 \leq y_1$. If $y'_1 < y_1$, then $x_1 \leq y'_1 \leq y_1 - 1$. There exists a D-multiset $Y' = (y_1 - 1, y_2 \dots, y_j + 1, \dots, y_n)$ for some $j > 1$ or $Y_{\ell(v)} = (y_1, k_2 \dots, k_n)$. The latter is impossible as $Y_{\ell(v)}$ and Y are both D-multisets. Therefore, $Y' < Y_{\ell(v)}$ and $X_v \subseteq Y'$ which contradicts the definition of $Y_{\ell(v)}$. We conclude that $y'_1 = y_1$. By the same argument one proves $y'_j = y_j$ for all $1 \leq j \leq i-1$.

So we can assume $i = 1$ or $i > 1$ and $y'_j = y_j$ for $1 \leq j \leq i-1$. If $y'_i > y_i$, then $Y > Y_{\ell(v)}$ and the statement holds. Otherwise, if $y'_i \leq y_i$, then $x_i < x'_i \leq y'_i \leq y_i$. As $i < n$, there exists a D-multiset $Y' = (y_1, \dots, y_{i-1}, y_i - 1, \dots, y_j + 1, \dots, y_n)$ for some j such that $Y' < Y_{\ell(v)}$ and $X_v \subseteq Y'$, a contradiction to the definition of $\ell(v)$.

Secondly, it is easy to see that for every D-multiset $Y \geq Y_{\ell(v)}$ there exists a d-multiset $X \geq X_v$ such that $X \subseteq Y$. Therefore, the family of D-multisets covered by $\mathcal{X}_v$ is exactly $\mathcal{Y}_{\ell(v)}$. That proves the lemma. $\square$

Lemma 3.3. It suffices to prove Theorem 3.1 for $D = d+1$.

Proof. Let the theorem be true for $D = d+1$ and every d. We prove it is true for $D = d+2$. Let $\ell_{01}, \ell_{12}, \ell_{02}$ be the above function for $d, d+1$, and $d+1, d+2$, and $d, d+2$ respectively. Assume a family $\mathcal{A}$ of d-multisets covers a family $\mathcal{B}$ of $(d+1)$-multisets, and $\mathcal{B}$ covers a family $\mathcal{C}$ of $(d+2)$-multisets. Then $\mathcal{C}$ consists of all $(d+2)$-multisets covered by $\mathcal{A}$. In particular, $\ell_{12}(\ell_{01}(s)) = \ell_{02}(s)$. Let $|\mathcal{A}| = s, |\mathcal{B}| = r, |\mathcal{C}| = t$. Then

$$t \geq \ell_{12}(r), \quad r \geq \ell_{01}(s)$$

as Theorem 3.1 holds for $D = d+1$ by the assumption. Therefore, $t \geq \ell_{12}(r) \geq \ell_{12}(\ell_{01}(s)) = \ell_{02}(s)$ and the lemma is true for $D = d+2$. One uses the same argument to prove it for $D > d+2$. $\square$

Let s be a natural number and

$$f_s(v) = |\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}|$$

for $0 \leq v \leq |\mathcal{X}| - s$. The family $\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}$ incorporates all D-multisets covered by $\{X_{v+1}, \ldots, X_{v+s}\}$ and not covered by $\{X_1, \ldots, X_v\}$.

Lemma 3.4. $f_s(|\mathcal{X}| - s) \leq f_s(v) \leq f_s(0)$.

Proof. We will only prove the right hand side inequality

$$|\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}| \leq |\mathcal{Y}_{\ell(s)}|. \tag{3}$$

The left hand side inequality is proved by a similar argument. The proof is by induction. The statement is correct for $s = 0$, any v, and $v = 0$, any s.

We will reduce (3) to a "smaller" problem $|\mathcal{Y}_{\ell(v_1+s_1)} \setminus \mathcal{Y}_{\ell(v_1)}| \leq |\mathcal{Y}_{\ell(s_1)}|$, where $s_1 = s$ and $v_1 < v$ or $s_1 < s$. If $v < s$, then it is enough to prove $|\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(s)}| \leq |\mathcal{Y}_{\ell(v)}|$ as

$$|\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}| = |\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(s)}| + |\mathcal{Y}_{\ell(s)} \setminus \mathcal{Y}_{\ell(v)}| \leq |\mathcal{Y}_{\ell(v)}| + |\mathcal{Y}_{\ell(s)} \setminus \mathcal{Y}_{\ell(v)}| = |\mathcal{Y}_{\ell(s)}|.$$

So the problem is reduced to a "smaller" problem.

Assume $v \geq s$. Let u be the largest index such that $X_u = (1, 0, a_3, \ldots, a_n)$ for some $a_3, \ldots, a_n$. So z is the largest index such that $X_z = (0, 1, a_3, \ldots, a_n)$ and therefore $X_u > X_z$. If such u does not exist, then the proof is easily reduced to one of the cases below.

1. First, $u \leq v$. Then the first entry in each of $\{X_{v+1}, \ldots, X_{v+s}\}$ is 0. If $u < v$, then by induction (right hand side inequality of the lemma for a smaller n) $|\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}| \leq |\mathcal{Y}_{\ell(u+s)} \setminus \mathcal{Y}_{\ell(u)}|$ and the problem is reduced to a "smaller" problem $|\mathcal{Y}_{\ell(u+s)} \setminus \mathcal{Y}_{\ell(u)}| \leq |\mathcal{Y}_{\ell(s)}|$.

   So one can assume $v = u$. Let $X_{v+s} = (0, x_2, x_3, \ldots, x_n)$. One defines a mapping

   $$\varphi : (0, y_2, y_3, \ldots, y_n) \to (1, y_2 - 1, y_3, \ldots, y_n). \tag{4}$$

   If $x_2 \geq 1$, then φ is well defined on $\{X_{v+1}, \ldots, X_{v+s}\}$ and maps it to $\{X_{w+1}, \ldots, X_{w+s}\}$ for some $w < v$. It is not difficult to see that φ is a bijection between $\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}$ and $\mathcal{Y}_{\ell(w+s)} \setminus \mathcal{Y}_{\ell(w)}$. So $|\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}| = |\mathcal{Y}_{\ell(w+s)} \setminus \mathcal{Y}_{\ell(w)}|$. We obtain a reduction to a "smaller" problem $|\mathcal{Y}_{\ell(w+s)} \setminus \mathcal{Y}_{\ell(w)}| \leq |\mathcal{Y}_{\ell(s)}|$.

   Let $x_2 = 0$. So $X_{v+1} \leq X_z < X_{v+s}$. As $\varphi(X_z) = X_u = X_v$,

   $$\begin{aligned} |\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}| &= |\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(z)}| + |\mathcal{Y}_{\ell(z)} \setminus \mathcal{Y}_{\ell(v)}| \\ &\leq |\mathcal{Y}_{\ell(2v+s-z)} \setminus \mathcal{Y}_{\ell(v)}| + |\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(w)}| \\ &= |\mathcal{Y}_{\ell(w+s)} \setminus \mathcal{Y}_{\ell(w)}|, \end{aligned}$$

where $|\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(z)}| \leq |\mathcal{Y}_{\ell(2v+s-z)} \setminus \mathcal{Y}_{\ell(v)}|$ comes by induction (right hand side inequality of the lemma for a smaller n) and $|\mathcal{Y}_{\ell(z)} \setminus \mathcal{Y}_{\ell(u)}| = |\mathcal{Y}_{\ell(u)} \setminus \mathcal{Y}_{\ell(w)}|$ for some $w < v$ as φ is a bijection between these two sets. We obtain a reduction to a "smaller" problem $|\mathcal{Y}_{\ell(w+s)} \setminus \mathcal{Y}_{\ell(w)}| \leq |\mathcal{Y}_{\ell(s)}|$.

2. Secondly, $v < u$. If $v + s \leq u$, then the first entry in each of $\{X_{v+1}, \ldots, X_{v+s}\}$ is > 0. The statement follows by induction (right hand side inequality of the lemma for a smaller k_1).

   We may assume $v < u < v + s$. By induction (left hand side inequality of the lemma for a smaller k_1), $|\mathcal{Y}_{\ell(u)} \setminus \mathcal{Y}_{\ell(v)}| \leq |\mathcal{Y}_{\ell(s)} \setminus \mathcal{Y}_{\ell(v+s-u)}|$ as $s \leq v < u$. It is enough now to show that $|\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(u)}| \leq |\mathcal{Y}_{\ell(v+s-u)}|$, where $0 < v + s - u < s$, and that is a "smaller" problem. It implies (3) as

$$\begin{aligned} |\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(v)}| &= |\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(u)}| + |\mathcal{Y}_{\ell(u)} \setminus \mathcal{Y}_{\ell(v)}| \\ &\leq |\mathcal{Y}_{\ell(v+s)} \setminus \mathcal{Y}_{\ell(u)}| + |\mathcal{Y}_{\ell(s)} \setminus \mathcal{Y}_{\ell(v+s-u)}| \\ &= |\mathcal{Y}_{\ell(s)}|. \end{aligned}$$

That finishes the proof of the lemma. □

Proof. We will now prove Theorem 3.1 by induction. Let $\{1, 2, \ldots, n\} = \{i_1, \ldots, i_r\} \cup \{j_1, \ldots, j_{n-r}\}$, where $1 \leq r < n$.

One splits $\mathcal{A} = \bigcup_Z \mathcal{A}_Z$ into subfamilies $\mathcal{A}_Z$, where Z are t-multisets $(z_{i_1}, \ldots, z_{i_r})$, $0 \leq t \leq d$. Each $(x_1, x_2, \ldots, x_n) \in \mathcal{A}_Z$ satisfies $(x_{i_1}, \ldots, x_{i_r}) = Z$ and $(x_{j_1}, \ldots, x_{j_{n-r}})$ is a $(d-t)$-multiset.

We construct a new family $\mathcal{C}$ of multisets of the same size as $\mathcal{A}$. Let $\mathcal{C}_Z$ be a family of d-multisets $(x_1, x_2, \ldots, x_n)$, where $(x_{i_1}, \ldots, x_{i_r}) = Z$ and $(x_{j_1}, \ldots, x_{j_{n-r}})$ are the first (largest) $|\mathcal{A}_Z|$ $(d-t)$-multisets according to a lexicographic order. Then $\mathcal{C} = \bigcup_Z \mathcal{C}_Z$. Obviously, $|\mathcal{C}| = |\mathcal{A}|$. We say $\mathcal{C}$ satisfies the condition $[i_1, \ldots, i_r]$.

Lemma 3.5. $||\mathcal{C}|| \leq ||\mathcal{A}||$

Proof. Let $\mathcal{B}$ be a family of D-multisets $(y_1, y_2, \ldots, y_n)$ covered by $\mathcal{A}$. One splits $\mathcal{B} = \bigcup_U \mathcal{B}_U$ into subfamilies $\mathcal{B}_U$, where U runs over T-multisets $(u_{i_1}, \ldots, u_{i_r})$. Each D-multiset $(y_1, y_2, \ldots, y_n) \in B_U$ satisfies $(y_{i_1}, \ldots, y_{i_r}) = U$ and $(y_{j_1}, \ldots, y_{j_{n-r}})$ is a $(D-T)$-multiset. One further splits $\mathcal{B}_U = \bigcup_{Z \subseteq U} \mathcal{B}_{U,Z}$ into subfamilies $\mathcal{B}_{U,Z}$ covered by $\mathcal{A}_Z$, where Z is a t-multiset and $0 \leq t \leq d$.

Let $\ell_{U,Z}(s)$ be the number of $(D-T)$-multisets of length $n-r$ covered by a minimal family of $(d-t)$-multisets of length $n-r$ and of size s. By induction, Theorem 3.1 is true for multisets of length $n-r < n$. So $|\mathcal{B}_{U,Z}| \geq \ell_{U,Z}(|\mathcal{A}_Z|)$ and therefore $|\bigcup_Z \mathcal{B}_{Z,U}| \geq \max_Z \ell_{Z,U}(|\mathcal{A}_Z|)$. Hence

$$||\mathcal{A}|| = |\bigcup_{Z,U} \mathcal{B}_{Z,U}| = \sum_U |\bigcup_{Z \subseteq U} \mathcal{B}_{Z,U}| \geq \sum_U \max_Z \ell_{Z,U}(|\mathcal{A}_Z|) = ||\mathcal{C}||. \qquad \square$$

If the family $\mathcal{A}$ does not satisfy a condition $[i_1, \ldots, i_r]$, then one transforms $\mathcal{A}$ into a family of d-multisets with the same size for which this condition is satisfied. Note $||\mathcal{A}||$ does not grow by Lemma 3.5. After each transformation, the members of $\mathcal{A}$ become larger (according to the lexicographic order), so this process stops at some point. We may assume $\mathcal{A}$ satisfies all the conditions $[i_1, \ldots, i_r]$ for $1 \leq r < n$.

The family $\mathcal{A}$ may be split into $\mathcal{A} = \bigcup_{z=0}^{k_1} \mathcal{A}_z$, where $\mathcal{A}_z$ incorporates multisets with the first entry z. As $\mathcal{A}$ satisfies the condition [1], each $\mathcal{A}_z$ is a minimal family of $(d-z)$-multisets of length $n-1$.

Let $s_0 = |\mathcal{A}_0|, s_1 = |\mathcal{A}_{k_1}|$, and u denote the number of all d-multisets the first entry of which is k_1. If $s_0 = 0$ or $s_1 = u$, then the theorem is true by induction for a smaller k_1. Assume $s_0 > 0$ and $s_1 < u$.

Let $\mathcal{A}_0 = \{X_{v-s_0+1}, \ldots, X_v\}$, where $X_v = (0, x_2, x_3, \ldots, x_n)$ for some $x_2, x_3, \ldots, x_n$. If $x_2 < k_1$, then $\mathcal{A}_0$ contains all d-multisets $(0, k_1, *, \ldots, *)$ as $\mathcal{A}_0$ is a minimal family. By condition $[3, \ldots, n]$, the family $\mathcal{A}$ contains all d-multisets $(k_1, 0, *, \ldots, *)$ and therefore all d-multisets $(k_1, *, *, \ldots, *)$. The latter is impossible as $s_1 < u$. So we can assume $x_2 \geq k_1$. Consider a mapping

$$\varphi : (0, y_2, y_3, \ldots, y_n) \to (k_1, y_2 - k_1, y_3, \ldots, y_n).$$

The mapping is well defined on $\mathcal{A}_0$. By condition $[3, \ldots, n]$, it maps $\mathcal{A}_0$ to

$$\{X_{w-s_0+1}, \ldots, X_w\} \subseteq A_{k_1}$$

for some $w \leq u$. It also maps $\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(v-s_0)}$ to $\mathcal{Y}_{\ell(w)} \setminus \mathcal{Y}_{\ell(w-s_0)}$. It is not difficult to see that φ is a bijection between those two sets. So $|\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(v-s_0)}| = |\mathcal{Y}_{\ell(w)} \setminus \mathcal{Y}_{\ell(w-s_0)}|$. This is also true for any subinterval of $\{X_{v-s_0+1}, \ldots, X_v\}$. We now consider two cases.

1. First, $u \geq s_0 + s_1$. Then

$$|\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(v-s_0)}| = |\mathcal{Y}_{\ell(w)} \setminus \mathcal{Y}_{\ell(w-s_0)}| \geq |\mathcal{Y}_{\ell(s_1+s_0)} \setminus \mathcal{Y}_{\ell(s_1)}|.$$

   The inequality comes from the left hand side inequality of Lemma 3.4 applied for $(d-k_1)$-multisets of length $n-1$ and defined by the numbers $k_2 - k_1, k_3, \ldots, k_n$.

   The multisets in $\mathcal{A}_0$ cover D-multisets in $\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(v-s_0)}$, the first entry of which is 0, and some other D-multisets, the first entry of which is $>$ 0. The latter are covered by $\mathcal{A} \setminus A_0$ as well. By Lemma 3.3 it suffices to consider $D = d+1$. If $(0, y_2, y_3, \ldots, y_n) \in A_0$, then it covers D-multiset $(1, y_2, y_3, \ldots, y_n)$. The latter is covered by $(1, y_2 - 1, y_3, \ldots, y_n)$, which belongs to A_1 by condition $[3, \ldots, n]$. We define a new family

$$\mathcal{C} = (\mathcal{A} \setminus \mathcal{A}_0) \cup \{X_{s_1+1}, \ldots, X_{s_1+s_0}\}.$$

   Then $|\mathcal{C}| = |\mathcal{A}|$ and $||\mathcal{C}|| \leq ||\mathcal{A}||$ by the inequality above. As $|C_0| = 0$, the theorem follows as above.

2. Secondly, $u < s_0 + s_1$. As φ is a bijection between $\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(v-u+s_1)}$ and $\mathcal{Y}_{\ell(w)} \setminus \mathcal{Y}_{\ell(w-u+s_1)}$,

$$|\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(v-u+s_1)}| = |\mathcal{Y}_{\ell(w)} \setminus \mathcal{Y}_{\ell(w-u+s_1)}| \geq |\mathcal{Y}_{\ell(u)} \setminus \mathcal{Y}_{\ell(s_1)}|.$$

The inequality comes from the left hand side inequality of Lemma 3.4 applied for $(d-k_1)$-multisets of length $n-1$ and defined by the integers $k_2, k_3, \ldots, k_n$. The multisets in $\{X_{v-u+s_1+1}, \ldots, X_v\}$ cover D-multisets in $\mathcal{Y}_{\ell(v)} \setminus \mathcal{Y}_{\ell(v-u+s_1)}$ and some other D-multisets. The latter are also covered by $\mathcal{A} \setminus \{X_{v-u+s_1+1}, \ldots, X_v\}$, which is easy to show under the condition $[3, \ldots, n]$ and $D = d + 1$ as above. We define a new family

$$\mathcal{C} = (\mathcal{A} \setminus \{X_{v-u+s_1+1}, \ldots, X_v\}) \cup \{X_{s_1+1} \ldots, X_u\}.$$

Then $|\mathcal{C}| = |\mathcal{A}|$ and $||\mathcal{C}|| \leq ||\mathcal{A}||$ by the inequality above. As $|C_{k_1}| = u$, the theorem follows in this case.

The proof is now complete. $\square$

4 Analysis of the probability

We consider a system of forms $f_1, \ldots, f_m$ of degree D. Let d be a natural number. The degree $d + D$ Macaulay matrix of the system is the matrix M, whose rows are labelled by pairs (r, f_i), where r is a monomial of degree d, and columns are labelled by the monomials t of degree $d + D$. The entry of the matrix M in row (r, f_i) and column t is equal to the coefficient of the monomial t in rf_i computed in R^h, see the Introduction. The size of the matrix M is $m\, l_q(n, d) \times l_q(n, d + D)$. If the columns of M are linearly independent, then $d_{\text{reg}} \leq d + D$.

Let $f_1, \ldots, f_m$ be taken independently and uniformly at random and let p denote the probability that the columns of M are linearly dependent. We prove that if $d < D$ and $m \geq l_q(n, d + D)/l_q(n, d)$, then

$$p \leq q^{l_q(n,d+D)-ml_q(n,d)} + O(n^d q^{-Cn^D})$$

for a positive constant C as n tends to infinity. This implies Theorem 1.1.

The matrix M can be divided into m blocks $M_1, \ldots, M_m$, each with $l_q(n, d)$ rows. The matrix M_j is the Macaulay matrix for the single polynomial f_j. Its rows are indexed by the multisets $\mathcal{X}^d$ and the columns by the multisets $\mathcal{X}^{d+D}$. For instance, let $q = 3, n = 3, D = 2$ and

$$f = c_{200}x_1^2 + c_{110}x_1x_2 + c_{101}x_1x_3 + c_{020}x_2^2 + c_{011}x_2x_3 + c_{002}x_3^2.$$

The degree 3 Macaulay matrix for f is in Table 2.

Table 2: The degree 3 Macaulay matrix for f

	012	021	102	111	120	201	210
001	c_{011}	c_{020}	c_{101}	c_{110}	0	c_{200}	0
010	c_{002}	c_{011}	0	c_{101}	c_{110}	0	c_{200}
100	0	0	c_{002}	c_{011}	c_{020}	c_{101}	c_{110}

As the f_j are chosen independently, the matrices M_j are independent. Let u be a vector over $\mathbb{F}_q$ and of length $l_q(n, d+D)$. Its entries are indexed by the multisets $\mathcal{X}^{d+D}$. Then

$$p_u = \mathbb{P}(Mu = 0) = p_{1u}^m,$$

where $p_{ju} = \mathbb{P}(M_j u = 0)$. Therefore, $p \leq \sum_{u \neq 0} p_u = \sum_{u \neq 0} p_{1u}^m$.

Let c denote a vector of coefficients of f_1. It has length $l_q(n, D)$, and its entries c_L are indexed by the multisets $L \in \mathcal{X}^D$. Let m_{JI} denote the entry of M_1 in the row $J \in \mathcal{X}^d$ and the column $I \in \mathcal{X}^{d+D}$. By the definition of M_1, we have $m_{JI} = c_{I \setminus J}$ if $J \subseteq I$ and $m_{JI} = 0$ otherwise, see Table 2 for an example. So $M_1 u = 0$ is equivalent to the following equalities which hold for every row of M_1 indexed by $J \in \mathcal{X}^d$. Observe that

$$\sum_{I \in \mathcal{X}^{d+D}} m_{JI}\, u_I = \sum_{J \subseteq I} c_{I \setminus J}\, u_I = \sum_{J+L \in \mathcal{X}^{d+D}} c_L\, u_{J+L} = 0, \tag{5}$$

where the second sum is over $I \in \mathcal{X}^{d+D}$ such that $J \subseteq I$, and the third sum is over $L \in \mathcal{X}^D$ such that $L + J \in \mathcal{X}^{d+D}$.

Let $Y^{(u)}$ be a matrix of size $l_q(n, d) \times l_q(n, D)$ whose rows and columns are labelled by the multisets from $\mathcal{X}^d$ and $\mathcal{X}^D$ respectively. The entries of $Y^{(u)}$ are defined by

$$Y^{(u)}_{J,L} = \begin{cases} u_{J+L} & \text{if } J + L \in \mathcal{X}^{d+D}, \\ 0 & \text{otherwise.} \end{cases}$$

For $n = 3, q = 3$, $d = 1$, and $D = 2$ the matrix $Y^{(u)}$ is in Table 3. By (5),

Table 3: Matrix $Y^{(u)}$

	002	011	020	101	110	200
001	0	u_{012}	u_{021}	u_{102}	u_{111}	u_{201}
010	u_{012}	u_{021}	0	u_{111}	u_{120}	u_{210}
100	u_{102}	u_{111}	u_{120}	u_{201}	u_{210}	0

the equality $M_1 u = 0$ is equivalent to $Y^{(u)} c = 0$. So $p_{1u} = q^{-\text{rank}(Y^{(u)})}$ and therefore

$$p \le \sum_{u \ne 0} q^{-m \, \text{rank}(Y^{(u)})} = \sum_{v=0}^{l_q(n,d)-1} N_v q^{-m(l_q(n,d)-v)}, \tag{6}$$

where N_v denotes the number of vectors u such that $\text{rank}(Y^{(u)}) = l_q(n,d) - v$. The value N_v is bounded above by the size of

$$S_v = \left\{ u \mid \text{rank}(Y^{(u)}) \le l_q(n,d) - v \right\}.$$

In particular, $u \in S_v$ if and only if there exists a row vector subspace $V \subseteq \mathbb{F}_q^{l_q(n,d)}$ of dimension v in the kernel of $Y^{(u)}$. Let $B = (b_1, \dots, b_v)$ be a basis of this subspace. We index the coordinates of b_i with $J \in \mathcal{X}^d$ according to the lexicographic order from left to right. Then $b_i Y^{(u)} = 0$ is equivalent to the following equality which holds for every $L \in \mathcal{X}^D$:

$$\sum_{J+L \in \mathcal{X}^{d+D}} b_{i,J} \, u_{J+L} = 0, \tag{7}$$

where the sum is over $J \in \mathcal{X}^d$ such that $J + L \in \mathcal{X}^{d+D}$. The basis B may be represented as a matrix of size $v \times l_q(n,d)$ in row echelon form, where every leading coefficient is 1:

$$B = \begin{pmatrix} 0 & \dots & 0 & 1 & * & \dots & * & 0 & * & \dots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & * & \dots \\ \dots \end{pmatrix}.$$

For each vector b_i, $i = 1, \dots, v$, in the basis B we now define a matrix A_i. The matrix A_i has $l_q(n, d+D)$ rows and $l_q(n,D)$ columns, indexed by $I \in \mathcal{X}^{d+D}$ and by $L \in \mathcal{X}^D$ respectively. The indices are ordered according to the lexicographic order from left to right and from top to bottom. The entry I, L of A_i is defined by

$$A_{i,I,L} = \begin{cases} b_{i, I \setminus L} & \text{if } L \subseteq I, \\ 0 & \text{otherwise.} \end{cases}$$

For $n = 3, q = 3$ and $d = 1, D = 2$ the matrix A_i constructed for $b_i = (b_{100}, b_{010}, b_{001})$ is in Table 4. Let A_V denote the horizontal concatenation of the matrices $A_1, \dots, A_v$, that is $A_V = A_1 | A_2 | \dots | A_v$. The equalities (7) are equivalent to $u A_V = 0$ and therefore

$$|S_v| \le \sum_{\dim(V) = v} q^{l_q(n,d+D) - \text{rank}(A_V)},$$

where the sum is over subspaces V of dimension v in $\mathbb{F}_q^{l_q(n,d)}$. Let the multiset $J_i \in \mathcal{X}^d$ index the first nonzero entry of the vector $b_i \in B$. As B is in row echelon form, the multisets $J_1, \dots, J_v$ are pairwise different. We denote $\mathcal{I} = \bigcup_{i=1}^{v} \left\{ I \in \mathcal{X}^{d+D} | I \supseteq J_i \right\}$.

Table 4: Matrix A_i

	002	011	020	101	110	200
012	b_{010}	b_{001}	0	0	0	0
021	0	b_{010}	b_{001}	0	0	0
102	b_{100}	0	0	b_{001}	0	0
111	0	b_{100}	0	b_{010}	b_{001}	0
120	0	0	b_{100}	0	b_{010}	0
201	0	0	0	b_{100}	0	b_{001}
210	0	0	0	0	b_{100}	b_{010}

Lemma 4.1. $\operatorname{rank}(A_V) \geq |\mathcal{I}|$.

Proof. For $I \in \mathcal{I}$ we fix some multiset $J_k \subseteq I$ and take a column in the block A_k indexed by $L = I \setminus J_k$. We show that those $|\mathcal{I}|$ columns in A_V are linearly independent. It is enough to prove that the row with index I has 1 in the column L of the block A_k and that $A_{k,I',L} = 0$ if $I' < I$. First, $A_{k,I,L} = b_{k,J_k} = 1$ since $J_k = I \setminus L$. Let $I' < I$. We consider two cases.

1. Let $I' \not\supseteq L$, so $A_{k,I',L} = 0$ by the definition of A_k.
2. Let $I' \supseteq L$, so $I' = J + L$ for some d-multiset J. As $I = J_k + L$ and $I' < I$, we deduce that $J < J_k$ by the properties of the lexicographic order. Hence $A_{k,I',L} = b_{k,J} = 0$.

The lemma is proved. $\square$

Similar to Section 3, let $\ell(v)$ denote the minimal number of $(d + D)$-multisets covered by v of d-multisets. By Lemma 4.1 and Theorem 3.1, $\operatorname{rank}(A_V) \geq |\mathcal{I}| \geq \ell(v)$. So

$$N_v \leq \sum_{\dim(V)=v} q^{l_q(n,d+D)-\operatorname{rank}(A_V)} \leq s_v q^{l_q(n,d+D)-\ell(v)},$$

where s_v is the number of subspaces of dimension v in $\mathbb{F}_q^{l_q(n,d)}$. It is easy to see that $s_v \leq q^{(l_q(n,d)-v+1)v}$. By (6),

$$\begin{aligned} p \leq \sum_{v=0}^{l_q(n,d)-1} q^{(l_q(n,d)-v+1)v+l_q(n,d+D)-\ell(v)-(l_q(n,d)-v)m} &= q^{l_q(n,d+D)-ml_q(n,d)} \\ + \sum_{v=1}^{l_q(n,d)-1} q^{(l_q(n,d)-v+1)v+l_q(n,d+D)-\ell(v)-(l_q(n,d)-v)m}. & \qquad (8) \end{aligned}$$

In Section 5 we prove that the second term is $O(n^d q^{-Cn^D})$ for fixed $d < D, q$, a positive constant C and n tending to infinity. That will finish the proof of Theorem 1.1. $\square$

Remark 4.2. If $m < l_q(n, d+D)/l_q(n, d)$, then the regularity degree for m polynomials of degree D must be larger than $d+D$, for the Macaulay matrix of degree $d+D$ cannot have linearly independent columns.

5 The second term

In this section we bound the second term in (8). In order to simplify notation we combine inequalities with O-notation in what follows. By convention, the expression $f(n) \le g(n) + O(h(n))$ means that there exists $t(n)$ such that $|t(n)| \le c|h(n)|$ and $f(n) \le g(n) + t(n)$ for a positive constant c and all sufficiently large n. Similarly, $f(n) \ge g(n) + O(h(n))$ means that there exists $t(n)$ such that $|t(n)| \le c|h(n)|$ and $f(n) \ge g(n) + t(n)$ for a positive constant c and all sufficiently large n.

Let $d < D$ and $q \ge 2$ be fixed and let n tend to infinity. Let $\mathcal{X} = \{(a_1, \ldots, a_n) | 0 \le a_i < q, \sum_{i=1}^{n} a_i = d\}$ be a family of d-multisets as defined in Section 3 for $k_i = q-1$. By $\ell(v)$ we denote the number of $(d+D)$-multisets covered by the family of the first v of lexicographically ordered d-multisets $\mathcal{X}$. Let $S(t) = \sum_{i=0}^{\infty} \alpha_i t^i$ and $[t^d]S(t)$ denote the coefficient at t^d. Obviously,

$$l_q(n, d) = [t^d] \frac{(1-t^q)^n}{(1-t)^n}. \tag{9}$$

Let $\begin{bmatrix} n \\ j \end{bmatrix}$ denote the number of solutions to $j = x_1 + \ldots + x_n$ in integer $x_i \ge 0$. Then $\frac{1}{(1-t)^n} = \sum_{j=0}^{\infty} \begin{bmatrix} n \\ j \end{bmatrix} t^j$.

Lemma 5.1. $l_q(n, d) = \begin{bmatrix} n \\ d \end{bmatrix} + O(n^{d-q+1})$ as $n \to \infty$.

Proof. By (9),

$$l_q(n, d) = [t^d] \left(\sum_{i=0}^{n} (-1)^i \binom{n}{i} t^{qi} \cdot \sum_{j=0}^{\infty} \begin{bmatrix} n \\ j \end{bmatrix} t^j \right) = \sum_{i=0}^{\lfloor d/q \rfloor} (-1)^i \binom{n}{i} \begin{bmatrix} n \\ d-iq \end{bmatrix}$$
$$= \begin{bmatrix} n \\ d \end{bmatrix} + \sum_{i=1}^{\lfloor d/q \rfloor} (-1)^i \binom{n}{i} \begin{bmatrix} n \\ d-iq \end{bmatrix} = \begin{bmatrix} n \\ d \end{bmatrix} + O(n^{d-q+1}). \qquad \square$$

Let $s \ge 1$ and $l_{s,q}(n, d)$ denote the number of monomials of degree d in $\mathbb{F}_q[x_1, \ldots, x_n]/(x_1^s, x_2^q, \ldots, x_n^q)$. Obviously,

$$l_{s,q}(n, d) = \sum_{i=0}^{s-1} l_q(n-1, d-i). \tag{10}$$

Let $S = \{1, \ldots, l_q(n, d)\}$ and let X_v denote the v-th largest multiset in the family of d-multisets $\mathcal{X}$ according to the lexicographic order. We will partition S into disjoint intervals.

Let $0 \le \delta \le d$. By division with remainder, $d - \delta = \sigma(q-1) + t$ for some $\sigma \ge 0$ and $0 \le t < q-1$. We consider a family of all d-mutisets

$$(q-1, \dots, q-1, u, a_{\sigma+2}, \dots, a_n),$$

where $u \ge t$, for some $a_{\sigma+2}, \dots, a_n$. Let v_δ denote the largest index v such that X_v belongs to that family. If that does not exist, then we put $v_\delta = v_{\delta-1}$, where $v_{-1} = 0$. Obviously, $v_\delta = l_{q-t,q}(n-\sigma, \delta)$. In particular, $v_0 = 1, v_d = l_q(n, d)$, and $v_{-1} < v_0 \le v_1 \le \ldots \le v_d$.

Let I_δ denote all v such that $v_{\delta-1} < v \le v_\delta$. Clearly, $v \in I_\delta$ if and only if X_v belongs to the family of d-multisets

$$(q-1, \dots, q-1, t, a_{\sigma+2}, \dots, a_n)$$

for some $a_{\sigma+2}, \dots, a_n$. So $|I_\delta| = v_\delta - v_{\delta-1} = l_q(n-\sigma-1, \delta)$. Observe $S = \bigcup_{\delta=0}^{d} I_\delta$. Let $0 \le x \le n-\sigma-1$. We consider a family of all d-multisets

$$(q-1, \dots, q-1, t, 0, \dots, 0, a_{\sigma+x+2}, \dots, a_n),$$

where $a_{\sigma+x+2} \ne 0$. Let $v_{\delta,x}$ denote the largest v such that X_v belongs to that family. If the family is empty, then we put $v_{\delta,x} = v_{\delta,x-1}$, where $v_{\delta,-1} = v_{\delta-1}$. Then $v_{\delta-1} = v_{\delta,-1} \le v_{\delta,0} \le \ldots \le v_{\delta,n-\sigma-1} = v_\delta$. Obviously, $v_{\delta,x} = v_\delta - l_q(n-\sigma-x-2, \delta)$. Let $I_{\delta,x}$ denote the set of all v such that $v_{\delta,x-1} < v \le v_{\delta,x}$. Then $I_\delta = \bigcup_{x=0}^{n-\sigma-1} I_{\delta,x}$.

Proposition 5.2. If $\delta = 0$, then $\ell(v_{0,n-\sigma-1}) = l_{q-t,q}(n-\sigma, \delta+D)$, and $\ell(v_{0,x}) = 0$ for $x < n-\sigma-1$. If $\delta > 0$, then

$$\ell(v_{\delta,x}) = l_{q-t,q}(n-\sigma, \delta+D) - l_q(n-\sigma-x-2, \delta+D).$$

Proof. For $\delta = 0$ the statement is obviously correct. Let $\delta > 0$. We notice that the family of d-multisets X_v, where $1 \le v \le v_{\delta,x}$, consists of d-multisets

$$(q-1, \dots, q-1, t + a_{\sigma+1}, a_{\sigma+2}, \dots, a_n),$$

where at least one among $a_{\sigma+1}, \dots, a_{\sigma+x+2}$ is non-zero and $\sum a_i = \delta$. That family covers precisely all $(d+D)$-multisets of the form

$$(q-1, \dots, q-1, t + a_{\sigma+1}, a_{\sigma+2}, \dots, a_n),$$

where at least one among $a_{\sigma+1}, \dots, a_{\sigma+x+2}$ is non-zero and $\sum a_i = \delta + D$. The number of such $(d+D)$-multisets is

$$l_{q-t,q}(n-\sigma, \delta+D) - l_q(n-\sigma-x-2, \delta+D).$$

That implies the statement for $\delta > 0$. $\square$

Lemma 5.3. If $v \in I_{\delta,x}$, then $\ell(v+1) - \ell(v) \le l_q(n-\sigma-x-2, D)$.

Proof. Since $v \in I_{\delta,x}$,

$$X_v = (q-1, \dots, q-1, t, 0, \dots, 0, a_{\sigma+x+2}, \dots, a_n),$$

for some $a_{\sigma+x+2}, \dots, a_n$, where $a_{\sigma+x+2} \neq 0$. It follows that

$$X_{v+1} = (q-1, \dots, q-1, t, 0, \dots, 0, a_{\sigma+x+2}, \dots, a_{j-1}, a_j - 1, b_{j+1}, \dots, b_n),$$

for $j \geq \sigma + x + 2$ and some $b_{j+1}, \dots, b_n$. Every $(d+D)$-multiset covered by X_{v+1} and not covered by $\{X_1, \dots X_v\}$ is in the family of $(d+D)$-multisets

$$(q-1, \dots, q-1, t, 0, \dots, 0, a_{\sigma+x+2}, \dots, a_{j-1}, a_j - 1, c_{j+1}, \dots, c_n),$$

for some $c_{j+1}, \dots, c_n$. The size of that family is at most $l_q(n-\sigma-x-2, D)$. That implies the lemma. □

Lemma 5.4. Let $1 < s \leq q$.

1. $l_{s,q}(n,\delta) - l_q(n-x,\delta) \geq x l_q(n-x, \delta-1)$.
2. For $x \leq \sqrt{n}$ and sufficiently large n,

$$l_{s,q}(n,\delta) - l_q(n-x,\delta) \leq x(l_q(n-1,\delta-1) + (q-2) l_q(n-1,\delta-2)).$$

Proof. By (10),

$$\begin{aligned}
& l_{s,q}(n,\delta) - l_q(n-x,\delta) \\
= \quad & (l_{s,q}(n,\delta) - l_q(n-1,\delta)) + \sum_{i=1}^{x-1} (l_q(n-i,\delta) - l_q(n-i-1,\delta)) \\
= \quad & \sum_{j=1}^{s-1} l_q(n-1,\delta-j) + \sum_{i=1}^{x-1}\sum_{j=1}^{q-1} l_q(n-i-1,\delta-j) \geq x l_q(n-x,\delta-1)
\end{aligned}$$

by considering only summands for $j = 1$. On the other hand, for $x < \sqrt{n}$ and sufficiently large n, $l_q(n-x,\delta-i) > l_q(n-x,\delta-i-1)$. Therefore,

$$\begin{aligned}
l_{s,q}(n,\delta) - l_q(n-x,\delta) &= \sum_{j=1}^{s-1} l_q(n-1,\delta-j) + \sum_{i=1}^{x-1}\sum_{j=1}^{q-1} l_q(n-i-1,\delta-j) \\
&\leq \sum_{i=0}^{x-1}\sum_{j=1}^{q-1} l_q(n-i-1,\delta-j) \\
&\leq x l_q(n-1,\delta-1) + (q-2)\sum_{i=0}^{x-1} l_q(n-i-1,\delta-2) \\
&\leq x(l_q(n-1,\delta-1) + (q-2) l_q(n-1,\delta-2)). \quad \square
\end{aligned}$$

We consider the exponent in the second term of (8). As $m \geq \frac{l_q(n,d+D)}{l_q(n,d)}$,

$$(l_q(n,d) - v + 1)\, v + l_q(n,d+D) - \ell(v) - (l_q(n,d) - v)\, m \leq E_n(v),$$

where $E_n(v) = Pv - v^2 - \ell(v)$ and $P = \left(l_q(n,d) + 1 + \frac{l_q(n,d+D)}{l_q(n,d)}\right)$. Assume $v \in I_\delta$, that is $v_{\delta-1} < v \leq v_\delta$. First, if $\delta = 0$, then $v = 1$ and

$$E_n(1) = l_q(n,d) + \frac{l_q(n,d+D)}{l_q(n,d)} - \ell(1),$$

where, by Proposition 5.2, $\ell(1) = l_{t,q}(n-\sigma, D) = \frac{n^D}{D!} + O(n^{D-1})$ for large n. Therefore,

$$E_n(1) = -n^D \left(\frac{1}{D!} - \frac{d!}{(d+D)!}\right) + O(n^{D-1}). \tag{11}$$

Let $\delta > 0$ and $v \in I_{\delta,x}$. This implies that $v_{\delta,x-1} < v \leq v_{\delta,x}$.

Lemma 5.5. Let $0 < \alpha < \sqrt[D]{\frac{d!D!}{(d+D)!}}$. For $x > n(1-\alpha)$ and $v \in I_{\delta,x}$, we have $E_n(v+1) - E_n(v) > 0$ for all sufficiently large n. In particular, the maximum on the given intervals of the function E_n can be found at $v = v_\delta$.

Proof. Using Lemma 5.3, we can see that

$$\begin{aligned} E_n(v+1) - E_n(v) &= P - 2v - 1 - \ell(v+1) + \ell(v) \\ &\geq \frac{l_q(n,d+D)}{l_q(n,d)} - l_q(n,d) - l_q(n-\sigma-x-2, D). \end{aligned}$$

As $x > n(1-\alpha_0)$,

$$\begin{aligned} E_n(v+1) - E_n(v) &\geq \frac{\begin{bmatrix} n \\ d+D \end{bmatrix}}{\begin{bmatrix} n \\ d \end{bmatrix}} - \begin{bmatrix} \alpha n - \sigma - 2 \\ D \end{bmatrix} + O(n^{D-1}) \\ &\geq n^D \left(\frac{d!}{(d+D)!} - \frac{\alpha^D}{D!}\right) + O(n^{D-1}). \end{aligned}$$

So, for sufficiently large n, $E_n(v+1) - E_n(v) > 0$ for $v \in I_{\delta,x}$ and $x > n(1-\alpha)$. $\square$

Proposition 5.6. There exists positive C and n_0 such that $E_n(v) < -Cn^D$ for $n \geq n_0$ and $1 \leq v \leq l_q(n,d) - 1$.

Proof. Let $v \in I_{\delta,x}$, that is $v_{\delta,x-1} < v \leq v_{\delta,x}$. Then $E_n(v) < Pv_{\delta,x} - \ell(v_{\delta,x-1})$. Let $0 < \alpha < \sqrt[D]{\frac{d!D!}{(d+D)!}}$ be fixed. We will divide I_δ into three intervals: $0 \leq x \leq \sqrt{n}$, $\sqrt{n} < x \leq n(1-\alpha)$, $n(1-\alpha) < x \leq n - \sigma - 1$ and bound $E_n(v)$ from above on each of them.

Case 1. Let $0 \le x \le \sqrt{n}$. By Lemma 5.4,

$$\begin{aligned}E_n(v) &\le P\, v_{\delta,x} - \ell(v_{\delta,x-1}) \le P\,(x+2)(l_q(n-\sigma-1,\delta-1)+\\ &+(q-2)l_q(n-\sigma-1,\delta-2)) - (x+1)l_q(n-\sigma-x-1,\delta+D-1))\\ &\le (x+1)\left(n^{\delta+D-1}\left(\frac{2d!}{(d+D)!(\delta-1)!}-\frac{1}{(\delta+D-1)!}\right)+\right.\\ &\left.+\,O(n^{\delta+D-3/2})\right). \qquad (12)\end{aligned}$$

The summand with the highest power of n of the last expression is negative for every $x \ge 0$, since

$$2(\delta+D-1)! = (2\delta)(\delta+D-1)\dots(\delta+1)(\delta-1)! < (d+D)\dots(d+1)(\delta-1)!$$

Hence, for n sufficiently large the maximum of (12) is achieved for $x=0$.

Case 2. Let $\sqrt{n} < x \le n(1-\alpha_0)$. For simplicity, we replace $n-\sigma-x-2 = y$, so $\alpha_0 n - 2 - \sigma \le y < n - \sqrt{n} - 2 - \sigma$. By rearranging the terms,

$$\begin{aligned}E_n(v) &\le P\, v_{\delta,x} - \ell(v_{\delta,x-1})\\ &= P\, l_{q-t,q}(n-\sigma,\delta) - l_{q-t,q}(n-\sigma,\delta+D) - P\, l_q(y,\delta) + l_q(y+1,\delta+D).\end{aligned}$$

Hence

$$\begin{aligned}Pl_{q-t,q}(n-\sigma,\delta) - l_{q-t,q}(n-\sigma,\delta+D) &= n^{\delta+D}\left(\frac{d!}{(D+d)!\delta!}-\frac{1}{(D+\delta)!}\right)+\\ &+O(n^{\delta+D-1}). \qquad (13)\end{aligned}$$

Then

$$\begin{aligned}-Pl_q(y,\delta)+l_q(y+1,\delta+D) &= \begin{bmatrix} y\\ \delta\end{bmatrix}\left(-\frac{\begin{bmatrix} n\\ d+D\end{bmatrix}}{\begin{bmatrix} n\\ d\end{bmatrix}}+\frac{\begin{bmatrix} y\\ \delta+D\end{bmatrix}}{\begin{bmatrix} y\\ \delta\end{bmatrix}}\right)+O(n^{\delta+D-1})\\ &\le \begin{bmatrix} y\\ \delta\end{bmatrix}\left(-\frac{n^D d!}{(D+d)!}+\frac{(n-\sqrt{n})^D\delta!}{(D+\delta)!}\right)+O(n^{\delta+D-1})\\ &= \begin{bmatrix} y\\ \delta\end{bmatrix}\left(\frac{n^D\delta!}{(D+\delta)!}-\frac{n^D d!}{(D+d)!}-\frac{n^{D-1/2}D\delta!}{(\delta+D)!}\right)+O(n^{\delta+D-1}).\end{aligned}$$

We notice that for sufficiently large n (this choice depends only on δ, d, and D) the sum in the parenthesis is positive if $\delta < d$ and negative if $\delta = d$. If $\delta < d$, then

$$\begin{aligned}&-Pl_q(y,\delta)+l_q(y+1,\delta+D)\\ \le\;& n^{\delta+D}\left(\frac{1}{(\delta+D)!}-\frac{d!}{(D+d)!\delta!}\right)-\frac{n^{\delta+D-1/2}D}{(\delta+D)!}+O(n^{\delta+D-1}). \qquad (14)\end{aligned}$$

If $\delta = d$, then

$$-Pl_q(y,d) + l_q(y+1, d+D) \leq -\frac{n^{d+D-1/2}D\alpha^d}{(d+D)!} + O(n^{d+D-1}). \tag{15}$$

Overall for $\delta < d$, by putting together (13) and (14),

$$E_n(v) \leq -\frac{n^{\delta+D-1/2}D}{(\delta+D)!} + O(n^{\delta+D-1}) \tag{16}$$

for sufficiently large n. For $\delta = d$, by putting together (13) and (15),

$$E_n(v) \leq -\frac{n^{d+D-1/2}D\alpha^d}{(d+D)!} + O(n^{d+D-1}) \tag{17}$$

for sufficiently large n.

Case 3. Let $n(1-\alpha) < x \leq n - \sigma - 1$. By Lemma 5.5, $E_n(v) \leq E_n(v_\delta)$.

For $\delta = d$, since $v_d = l_q(n,d)$ is not in the domain of E_n, we use $E_n(v_d - 1)$ as an upper bound, where

$$\begin{aligned} E_n(v_d - 1) &= 2l_q(n,d) - 2 - \frac{l_q(n, d+D)}{l_q(n,d)} \\ &= -\frac{n^D d!}{(d+D)!} + O(n^{D-1}) \end{aligned} \tag{18}$$

since $l_q(n, d+D) = \ell(v_d - 1)$. For $\delta < d$, the maximum of E_n on the interval is achieved at v_δ:

$$\begin{aligned} E_n(v_\delta) &\leq P\, l_{q-t,q}(n-\sigma, \delta) - l_{q-t,q}(n-\sigma, \delta+D) \\ &= -n^{\delta+D}\left(\frac{1}{(\delta+D)!} - \frac{d!}{(D+d)!\delta!}\right) + O(n^{\delta+D-1}). \end{aligned} \tag{19}$$

Overall, by combining (11),(12),(16),(17),(18),(19), we get $E_n(v) < -Cn^D$ for a positive C, and sufficiently large n uniformly in $v \in \{1, \ldots, l_q(v,d) - 1\}$ (that means $n \geq n_0$ and n_0 is independent of v). $\square$

We conclude the proof of Theorem 1.1:

$$p \leq q^{l_q(n,d+D) - ml_q(n,d)} + \sum_{v=1}^{l_q(n,d)-1} q^{E_n(v)} \leq q^{l_q(n,d+D)-ml_q(n,d)} + O(n^d q^{-Cn^D}).$$

Acknowledgements

The research was funded by the Norwegian Research Council under the project *Modern Methods and Tools for Theoretical and Applied Cryptology.* The authors are grateful to the anonymous reviewers for their helpful comments.

References

[BCP97] Wieb Bosma, John Cannon, and Catherine Playoust, *The Magma algebra system. I. The user language*, Computational algebra and number theory (London, 1993), 1997, DOI: 10.1006/jsco.1996.0125, URL: http://dx.doi.org/10.1006/jsco.1996.0125.

[BFS03] Magali Bardet, Jean-Charles Faugère, and Bruno Salvy, "Complexity of Gröbner basis computation of Semi-Regular Overdetermined sequences over F_2 with solutions in F_2", *INRIA Research Report 5049*, 2003, pp. 71–74.

[BFS15] Magali Bardet, Jean-Charles Faugère, and Bruno Salvy, "On the complexity of the F5 Gröbner basis algorithm", *Journal of Symbolic Computation* 70 (2015), pp. 49–70.

[Buc65] Bruno Buchberger, "Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal, (An algorithm for finding a basis for the residue class ring of a zero-dimensional polynomial ideal)", *PhD thesis, University of Innsbruck* (1965).

[CG17] Alessio Caminata and Elisa Gorla, "Solving multivariate polynomial systems and an invariant from commutative algebra", *arXiv preprint arXiv:1706.06319* (2017).

[CKPS00] Nicolas Courtois, Alexander Klimov, Jacques Patarin, and Adi Shamir, "Efficient algorithms for solving overdefined systems of multivariate polynomial equations", *EUROCRYPT 2000. LNCS 1807*, Springer, Heidelberg, 2000, pp. 392–407.

[CL69] George F. Clements and Bernt Lindström, "A generalization of a combinatorial theorem of Macaulay", *Journal of Combinatorial Theory* 7.3 (1969), pp. 230–238.

[CLO13] David Cox, John Little, and Donal O'Shea, *Ideals, Varieties, and Algorithms: An introduction to Computational Algebraic Geometry and Commutative Algebra*, 4th ed., Springer, Heidelberg, 2013.

[CP02] Nicolas Courtois and Josef Pieprzyk, "Cryptanalysis of block ciphers with overdefined systems of equations", *ASIACRYPT 2002. LNCS 2501*, Springer, Heidelberg, 2002, pp. 267–287.

[Die04] Claus Diem, "The XL-algorithm and a conjecture from commutative algebra", *ASIACRYPT 2004. LNCS 3329*, Springer, Heidelberg, 2004, pp. 323–337.

[Fau02] Jean-Charles Faugère, "A new efficient algorithm for computing Gröbner bases without reduction to zero (F5)", *ISSAC 2002*, ACM Press, 2002, pp. 75–83.

[Fau99] Jean-Charles Faugère, "A new efficient algorithm for computing Gröbner bases (F4)", *Journal of pure and applied algebra* 139.1-3 (1999), pp. 61–88.

[FJ03] Jean-Charles Faugère and Antoine Joux, "Algebraic cryptanalysis of hidden field equation (HFE) cryptosystems using Gröbner bases", *CRYPTO 2003. LCNS 2729*, Springer, Heidelberg, 2003, pp. 44–60.

[HMS17] Timothy J. Hodges, Sergio D. Molina, and Jacob Schlather, "On the existence of homogeneous semi-regular sequences in - $F_2[X_1, ..., X_n]/(X_1^2, ..., X_n^2)$", *Journal of Algebra* 476 (2017), pp. 519 –547.

[KR00] Martin Kreuzer and Lorenzo Robbiano, *Computational Commutative Algebra 1*, Springer, Heidelberg, 2000.

[Laz83] Daniel Lazard, "Gröbner bases, Gaussian elimination and resolution of systems of algebraic equations", *EUROCAL 1983. LNCS 162*, Springer, Heidelberg, 1983, pp. 146–156.

[LW54] Serge Lang and André Weil, "Number of points of varieties in finite fields", *American Journal of Mathematics* (1954), pp. 819–827.

[Mac16] Francis S. Macaulay, *The algebraic theory of modular systems*, Cambridge University Press, 1916.

[ST19] Igor Semaev and Andrea Tenti, "Probabilistic analysis on Macaulay matrices over finite fields and complexity of constructing Gröbner bases", *WCC 2019*, 2019, URL: `https://www.lebesgue.fr/sites/default/files/proceedings_WCC/WCC_2019_paper_37.pdf`.

[Syl53] James J. Sylvester, "XVIII. On a theory of the syzygetic relations of two rational integral functions, comprising an application to the theory of Sturm's functions, and that of the greatest algebraical common measure", *Philosophical transactions of the Royal Society of London* 143 (1853), pp. 407–548.