The Number of Gröbner Bases in Finite Fields

Brandilyn Stigler, Anyu Zhang

Abstract In the field of algebraic systems biology, the number of minimal polynomial models constructed using discretized data from an underlying system is related to the number of distinct reduced Gröbner bases for the ideal of the data points. While the theory of Gröbner bases is extensive, what is missing is a closed form for their number for a given ideal. This work contributes connections between the geometry of data points and the number of Gröbner bases associated to small data sets. Furthermore we improve an existing upper bound for the number of Gröbner bases specialized for data over a finite field.

1 Introduction

Polynomial systems are ubiquitous across the sciences. While linear approximations are often desired for computational and analytic feasibility, certain problems may not permit such reductions. In 1965 Bruno Buchberger introduced Gröbner bases, which are multivariate nonlinear generalizations of echelon forms [3, 5]. Since this landmark thesis, the adoption of Gröbner bases has expanded into diverse fields, such as geometry [24], image processing [18], oil production [23], quantum field theory [20], and systems biology [17].

While working with a Gröbner basis (GB) of a system of polynomial equations is just as natural as working with a triangularization of a linear system, their complexity can make them cumbersome with which to work: for a general system, the complexity of Buchberger's Algorithm is doubly exponential in the number of variables [4]. The complexity improves in certain settings, such as systems with finitely many real-valued solutions ([6] is a classic example, whereas [12] is a more contemporary example), or solutions over finite fields [15]. Indeed much research has been devoted to improving Buchberger's Algorithm and analyzing the complexity and memory

Brandilyn Stigler, Southern Methodist University e-mail: bstigler@smu.edu · Anyu Zhang, Southern Methodist University, e-mail: anyuz@smu.edu

usage in more specialized settings (for example, [11, 19]), and even going beyond traditional ways of working with Gröbner bases [16]; however most results are for characteristic-0 fields, such $\mathbb{R}$ or $\mathbb{Q}$.

The goal of our work is to consider the *number* of Gröbner bases for a system of polynomial equations over a finite field (which has positive characteristic and consequently all systems have finitely many solutions). The motivation comes from the work of [17], in which the authors presented an algorithm to reverse engineer a model for a biological network from discretized experimental data and made a connection between the number of distinct *reduced* GBs and the number of (possibly) distinct *minimal* polynomial models. The number of reduced GBs associated to a data set gives a quantitative measure for how "underdetermined" the problem of reverse engineering a model for the underlying biological system is.

The Gröbner fan geometrically encapsulates all reduced Gröbner bases [21]. In [13] the authors provided an algorithm to compute all reduced GBs. When their number is too large for enumeration, the method in [9] allows one to sample from the fan. Finally in [22], the authors provide an upper bound for systems with finitely many solutions; however this bound is much too large for data over a finite field. To our knowledge, there is no closed form for the number of reduced Gröbner bases, in particular for systems over finite fields with finitely many solutions.

In this paper we make the following contributions:

1. a formula and some upper bounds of the number of reduced Gröbner bases for data sets over finite fields
2. geometric characterization of data associated with different numbers of reduced Gröbner bases.

In Section 2, we provide the relevant background, definitions, and results. In Section 3, we discuss the connection between the number of distinct reduced Gröbner bases for ideals of two points and the geometry of the points; furthermore, we provide a formula to two-point data sets. We provide upper bounds for data sets of three points in Section 4 and geometric observations for larger sets in Section 5. Then in Section 6, we consider the general setting of any fixed number of points over any finite field and provide an upper bound. We close with a discussion of possible future directions. We have verified all of the computations referenced in this work, provided illustrative examples throughout the text, and listed data tables in the Appendix.

2 Background

2.1 Algebraic Geometry Preliminaries

Let K be a field and let $R = K[x_1, \ldots, x_n]$ be a polynomial ring over K. Most definitions and known results in this section can be found in [8].

A *monomial order* $\prec$ is a total order on the set of all monomials in R that is closed with respect to multiplication and is a well-ordering. The *leading term* of a

polynomial $g \in R$ is thus the largest monomial for the chosen monomial ordering, denoted as $LT_\prec(g)$. Also we call $LT_\prec(I) = \langle LT_\prec(g) : g \in I \rangle$ the *leading term ideal* for an ideal I.

Definition 1 Let $\prec$ be a monomial order on R and let I be an ideal in R. Then $G \subset I$ is a *Gröbner basis* for I with respect to $\prec$ if for all $f \in I$ there exists $g \in G$ such that the leading term $LT_\prec(g)$ divides $LT_\prec(f)$.

It is well known that Gröbner bases exist for every $\prec$ and make multivariate polynomial division well defined in that remainders are unique; for example, see [8]. While there are infinitely many orders, there are only finitely many reduced GBs for a given ideal, that is monic polynomials whose leading terms do not divide other terms. This results in an equivalence relation where the leading terms of the representative of each equivalence class can be distinguished (underlined) [21]. In fact there is a one-to-one correspondence between *marked* reduced Gröbner bases and leading term ideals [7].

In this work all Gröbner bases are reduced.

Definition 2 The monomials which do not lie in $LT_\prec(I)$ are *standard* with respect to $\prec$; the set of standard monomials for an ideal I is denoted by $SM_\prec(I)$.

A set of standard monomials $SM_\prec(I)$ for a given monomial order forms a basis for R/I as a vector space over K. Given their construction, it follows that the sets of standard monomials associated to an ideal I are in bijection with the leading term ideals of I.

It is straightforward to check that standard monomials satisfy the following divisibility property: if $x^\alpha \in SM_\prec(I)$ and x^β divides x^α, then $x^\beta \in SM_\prec(I)$. This divisibility property on monomials is equivalent to the following geometric condition on lattice points.

Definition 3 A set $\lambda \subset \mathbb{N}^n$ is a *staircase* if for all $u \in \lambda$, $v \in \mathbb{N}^n$ and $v_i \le u_i$ for $1 \le i \le n$ imply $v \in \lambda$.

Let $\binom{\mathbb{N}^n}{m}$ denote the collection of all sets of m points in $\mathbb{N}^n$. Then for $\lambda = \{\lambda_1, \ldots, \lambda_m\} \in \binom{\mathbb{N}^n}{m}$, let $\sum \lambda$ denote the vector sum $\sum_{i=1}^{m} \lambda_i \in \mathbb{N}^n$. Let Λ denote the set of all staircases in $\binom{\mathbb{N}^n}{m}$. The *staircase polytope* of Λ is the convex hull of all points $\sum \lambda$ where $\lambda \in \Lambda$ (see [2, 22] for more details). For an ideal I, we call $\mathcal{P}$ the *staircase polytope of I* if $\mathcal{P}$ is the staircase polytope of the exponent vectors of the standard monomial sets associated to I for any monomial order.

For $S \subseteq K^n$, we call the set $I(S) := \{h \in R \mid h(s) = 0 \, \forall s \in S\}$ of polynomials that vanish on S an *ideal of points*. An ideal is *zero dimensional* if $\dim_K R/I < \infty$; when K is algebraically closed and $|S| = m < \infty$, then $m = \dim_K R/I(S)$. The number of reduced Gröbner bases for an ideal is in bijection with the number of vertices of the staircase polytope, which was proved for ideals of points in [22] and for all other zero-dimensional ideals in [2].

The following results provide an upper bound for the number of reduced Gröbner bases for an ideal over any field.

Lemma 1 ([1]) *The number of vertices of a lattice polytope $P \subset \mathbb{R}^n$ is $\#vert(P) = O\left(vol(P)^{(n-1)/(n+1)}\right)$.*

Theorem 1 ([2, 22]) *Let I be an ideal such that $\dim_K R/I = m$. Let $\Lambda(I)$ be the set of standard monomial sets for I over all monomial orders. Then the number of distinct reduced Gröbner bases of I is in bijection with the number of vertices of the staircase polytope of I; that is, $\#GBs = O\left(m^{2n\frac{n-1}{n+1}}\right)$.*

Example 1 Let $S = \{(1,1),(2,3),(3,5),(4,6)\} \subset \mathbb{R}^2$. So $\dim_{\mathbb{R}} \mathbb{R}[x,y]/I(S) = 4$. Also $\Lambda(I(S)) = \{(1,x,x^2,x^3), (1,x,x^2,y), (1,x,y,y^2), (1,y,y^2,y^3)\}$. So the number of reduced Gröbner bases for $I(S)$ is four. Note that there are five stair-cases in $\binom{\mathbb{N}^2}{4}$, namely $\Lambda = \{\{(0,0),(1,0),(2,0),(3,0)\}, \{(0,0),(1,0),(2,0),(0,1)\}, \{(0,0),(1,0),(0,1),(1,1)\}, \{(0,0),(1,0),(0,1),(0,2)\}, \{(0,0),(0,1),(0,2),(0,3)\}\}$. The staircase polytope of Λ is the convex hull of the vector sums $\{(6,0),(3,1),(2,2),(1,3),(0,6)\}$, which has vertices $(6,0)$, $(3,1)$, $(1,3)$, and $(0,6)$, corresponding to the four standard monomial sets of $I(S)$.

Now we summarize the bijective correspondences for the number of reduced Gröbner bases for an ideal of points.

Theorem 2 *Let I be an ideal. There is a one-to-one correspondence among the following:*

1. *distinct marked reduced Gröbner bases of I*
2. *leading term ideals of I*
3. *sets of standard monomials for I*
4. *vertices of the staircase polytope of I.*

Proof Equivalence 1 $\Longleftrightarrow$ 2 is a result in [7]; 2 $\Longleftrightarrow$ 3 is by construction of standard monomials; and 1 $\Longleftrightarrow$ 4 was proved in [22] for ideals of points and in [2] for other zero-dimensional ideals.   □

2.2 Ideals over Finite Fields

In this section and following, we will work over a finite base field. Let F be a finite field of characteristic $p > 0$. We will typically consider the finite field $\mathbb{Z}_p = \{0, 1, \ldots, p-1\}$, that is the field of remainders of integers upon division by p with modulo-p addition and multiplication. Let $R = F[x_1, \ldots, x_n]$ be a polynomial ring over F. Finally let m denote the number of points in a subset of F^n.

A *polynomial dynamical system* (PDS) over F is a function $f = (f_1, \ldots, f_n) : F^n \to F^n$ where each component f_i is a polynomial in R. Below is an algorithm, first introduced in [17], to compute a PDS from a given set of data written using the ideal of the input points. This algorithm motivates the leading question in this work.

The general strategy is given input-output data $V = \{(s_1, t_1), \ldots, (s_m, t_m)\} \subset F^n \times F^n$, find all PDSs that fit V and select a minimal PDS with respect to polynomial

division. This is done as follows. For each x_j, compute one interpolating function $f_j \in R$ such that $f_j(s_i) = t_{ij}$; note that $s_i \in F^n$ while $t_{ij} \in F$. Then compute the ideal $I := I(\{s_1, \ldots, s_m\})$ of the inputs in V. The *model space* for V is the set

$$f + I := \{(f_1 + h_1, \ldots, f_n + h_n) : h_i \in I\}$$

of all PDSs which fit the data in V and where $f = (f_1, \ldots, f_n)$ is as computed above. A PDS can be selected from $f + I$ by choosing a monomial order $\prec$, computing a Gröbner basis G for I, and then computing the remainder (*normal form*) $\overline{f}^G$ of each f_i by dividing f_i by the polynomials in G. We call

$$(\overline{f_1}^G, \overline{f_2}^G, \ldots, \overline{f_n}^G)$$

the *minimal* PDS with respect to $\prec$, where G is a Gröbner basis for I with respect to $\prec$.

Changing the monomial order may change the resulting minimal PDS. While it is possible for two reduced Gröbner bases to give rise to the same normal form (see [17]), it is still the case that in general a set of data points may have *many* GBs associated to it. In this way, the number of distinct reduced GBs of I gives an upper bound for the number of different minimal PDSs. Therefore, we aim to find the number of distinct reduced Gröbner bases for a given data set.

Example 2 Consider two inputs $S = \{(0,0), (1,1)\} \subset (\mathbb{Z}_2)^2$. The corresponding ideal I of the points in S has 2 distinct reduced Gröbner bases, namely

$$G_1 = \{\underline{x_1} - x_2, \underline{x_2^2} - x_2\}, G_2 = \{\underline{x_2} - x_1, \underline{x_1^2} - x_1\}$$

Here, '_' marks the leading terms of the polynomials in a Gröbner basis. There are two resulting minimal models: any minimal PDS with respect to G_1 will be in terms of x_2 only as all x_1's are divided out, while any minimal PDS with respect to G_2 will be in terms of x_1 only as all x_2's are divided out. Instead if the inputs are $\{(0,0), (0,1)\}$, then I has a unique GB, $\{\underline{x_2^2} - x_2, \underline{x_1}\}$, resulting in a unique minimal PDS.

It is the polynomial $g = x_1 - x_2$ that has different leading terms for different monomial orders. In fact, for monomial orders with $x_1 > x_2$, the leading term of g is x_1, while for orders with $x_2 > x_1$ the opposite will be true. We say that g has *ambiguous* leading terms. We will mark only ambiguous leading terms.

As the elements of the quotient ring R/I are equivalence classes of functions defined over the inputs $S = \{s_1, \ldots s_m\}$ in V and since a set of standard monomials is a basis for R/I, it follows that each reduced polynomial $\overline{f}^G$ is written in terms of standard monomials. When working over a finite field, extensions of classic results in algebraic geometry state that when the number m of input points is finite, then m coincides with the dimension of the vector space $R/I(S)$ over F [14], which is stated below for convenience.

Theorem 3 ([14]) *Let $S \subseteq \mathbb{F}^n$ and $I(S)$ be the ideal of the points in S. Then $|S| = \dim_F R/I(S)$.*

Next we state a result about data sets and their complements.

Theorem 4 ([10]) *Let I be the ideal of input points S, and let I^c be ideal of the complement $F^n \setminus S$ of S. Then we have $SM_\prec(I) = SM_\prec(I^c)$ and $LT_\prec(I) = LT_\prec(I^c)$ for a given monomial order $\prec$. Hence, we have $\#GB(S) = \#GB(F^n \setminus S)$.*

We say that a polynomial $f \in R$ is *factor closed* if every monomial $m \in supp(f)$ is divisible by all monomials in $supp(f)$ smaller than m with respect to an order $\prec$. The following result gives an algebraic description of ideals with unique reduced Gröbner bases for any monomial order.

Theorem 5 ([10]) *A reduced Gröbner basis G with factor-closed generators is reduced for every monomial order; that is, G is the unique reduced Gröbner basis for its corresponding ideal.*

We end this section with a discussion on the number of distinct reduced Gröbner bases for extreme cases. The set $\mathbb{Z}_p^n$ contains p^n points. For $n = 1$, all ideals have a unique reduced GB since all polynomials are single-variate and as such are factor closed. We consider cases for $n > 1$. For empty sets or singletons in $\mathbb{Z}_p^n$, it is straightforward to show that the ideal of points has a unique reduced GB for any monomial order; that is, for a point $s = (s_1, \ldots, s_n)$, the ideal of s is $I = \langle x_1 - s_1, \ldots, x_n - s_n \rangle$ whose generators form a Gröbner basis and hence is unique (via Theorem 5). According to Theorem 4, the same applies to $p^n - 1$ points. In the rest of this work, we consider the number of reduced Gröbner bases for an increasing number of points.

Note that over a finite field, the relation $x^p - x$ always holds.

3 Data Sets with $m = 2$ Points

In this section we consider bounds for the number of Gröbner bases for ideals of 2 points and relate the geometry of the points to these numbers.

Define $NGB(p, n, m)$ to be the number of reduced Gröbner bases for ideals of m points in $\mathbb{Z}_p^n$. The following theorem provides a formula for sets with $m = 2$ points in any number of coordinates and over any finite field $\mathbb{Z}_p$.

Theorem 6 *Let $P = (p_1, \ldots, p_n), Q = (q_1, \ldots, q_n) \in \mathbb{Z}_p^n$ where $P \neq Q$, and let $I \subset \mathbb{Z}_p[x_1, \ldots, x_n]$ be the ideal of the points P, Q. The number of distinct reduced Gröbner bases for I is given by*

$$NGB(p, n, 2) = \sum_{\substack{p_i \neq q_i \\ i=1,\ldots,n}} 1.$$

Proof Let $S = \{P, Q\} \subset \mathbb{Z}_p^n$ with $P = (p_1, \ldots, p_n), Q = (q_1, \ldots, q_n)$. Let $I \subset \mathbb{Z}_p[x_1, \ldots, x_n]$ be the ideal of the points in S. By Theorem 3, the number of elements of any set of standard monomials for I is $|S| = 2$. Since sets of standard monomials

must be closed under division, the only option for such a set is $\{1, x_i\}$ for some $i = 1, \ldots, n$. So the possible associated minimally generated leading term ideals are of the form $\langle x_1, \ldots, x_{i-1}, x_i^2, x_{i+1}, \ldots, x_n \rangle$. We consider the number of leading terms ideals in regards to the number of coordinate changes between the points.

If P and Q have one different coordinate, say $p_1 \neq q_1$, then the only possible minimal generating set for the leading term ideal of I is $\{x_1^2, x_2, \ldots, x_n\}$. If P, Q have two different coordinates, say $p_i \neq q_i$ for $i = 1, 2$, then the possible minimal generating sets for the leading term ideal of I are $\{x_1^2, x_2, \ldots, x_n\}$ when $x_1 \prec x_2$ and $\{x_1, x_2^2, x_3, \ldots, x_n\}$ when $x_2 \prec x_1$. Increasing the number of coordinate changes will add another leading term ideal. In general, if $p_i \neq q_i$ for $i = 1, \ldots, k$ where $k \leq n$, then the possible minimal generating sets for the leading term ideal of I are as follows:

1. $\{x_1^2, x_2, \ldots, x_n\}$ when x_1 is the smallest variable in the monomial order among $x_1, \ldots, x_k$
2. $\{x_1, x_2^2, x_3, \ldots, x_n\}$ when x_2 is smallest among $x_1, \ldots, x_k$
$\vdots$
k. $\{x_1, \ldots, x_{k-1}, x_k^2, x_{k+1}, \ldots, x_n\}$ when x_k is smallest among $x_1, \ldots, x_k$. $\qquad \square$

Corollary 1 *The maximum number of distinct reduced Gröbner bases for an ideal of two points in $\mathbb{Z}_p^n$ is $NGB(p, n, 2) \leq n$.*

With different choices of smallest coordinate, there are up to n different sets of standard monomials, each corresponding to a distinct reduced Gröbner basis. So, there are up to n reduced Gröbner bases, with the maximum achieved by two points with no coordinates in common.

In applications, modeling is often driven by data. So geometric descriptions of data sets can reveal essential features in the underlying network. We illustrate the above results by considering different configurations of points. We begin with Boolean data.
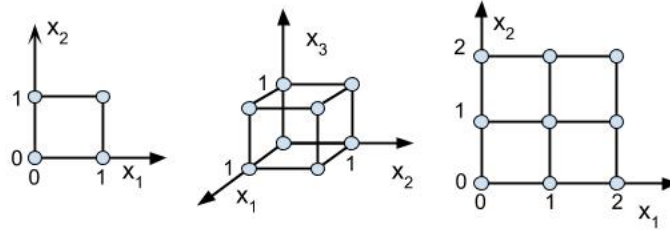

Fig. 1 The lattice of points in $\mathbb{Z}_2^2$ (left), in $\mathbb{Z}_2^3$ (center), and in $\mathbb{Z}_3^2$ (right).

Example 3 Consider two points in $\mathbb{Z}_2^2$. The left graph in Figure 1 is the plot of all points in $\mathbb{Z}_2^2$. By decomposing the 2-square on which they lie, we find that pairs

of points that lie along horizontal lines have unique reduced Gröbner bases for any monomial order; see Figure 2. For example, $\{(0,0),(0,1)\}$ has ideal of points $\langle x_1, x_2^2 - x_2 \rangle$. By Theorem 5 we see that the generators of I form a unique reduced GB. Similarly $\{(1,0),(1,1)\}$ has ideal of points $\langle x_1 - 1, x_2^2 - x_2 \rangle$, which also has a unique reduced GB. Note that while they have different GBs, they have the same leading term ideal, namely, $\langle x_1, x_2^2 \rangle$. In the same way, pairs of points that lie along vertical lines have unique reduced GBs: sets $\{(0,0),(1,0)\}$ and $\{(0,1),(1,1)\}$ have the unique leading term ideal $\langle x_1^2, x_2 \rangle$. In each case, these sets have points with one coordinate change.

On the other hand, pairs of points that lie on diagonals have 2 distinct reduced Gröbner bases as such points have two coordinate changes. For example, the set of points $\{(0,0),(1,1)\}$ has GBs $\{\underline{x_1} - x_2, x_2^2 - x_2\}$ and $\{x_1^2 - x_1, \underline{x_2} - x_1\}$ with leading term ideals $\langle x_1, x_2^2 \rangle$ and $\langle x_1^2, x_2 \rangle$ respectively. Similarly the set $\{(0,1),(1,0)\}$ has $\{\underline{x_1} - x_2 - 1, x_2^2 - x_2\}$ and $\{x_1^2 - x_1, \underline{x_2} - x_1 - 1\}$ as Gröbner bases with leading term ideals $\langle x_1, x_2^2 \rangle$ and $\langle x_1^2, x_2 \rangle$ respectively.
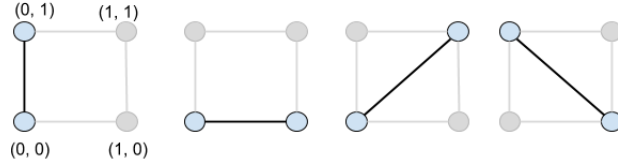

Fig. 2 Four configurations of pairs of points in $\mathbb{Z}_2^2$. From left to right: $\{(1,0),(0,1)\}$ and $\{(0,0),(1,0)\}$ each have 1 GB, while $\{(0,0),(1,1)\}$ and $\{(1,0),(0,1)\}$ have 2 distinct GBs.

Example 4 Now consider two points in $\mathbb{Z}_2^3$. The center graph in Figure 1 is the plot of all points in $\mathbb{Z}_2^3$. In Figure 3, pairs of points that lie on edges of the 3-cube have 1 reduced Gröbner basis, as the points have one coordinate change: for example the set $\{(1,0,1),(1,1,1)\}$ (first from the left in Figure 3) has the unique reduced GB $\{x_1 - 1, x_2^2 - x_2, x_3 - 1\}$ and $\{(0,0,0),(0,0,1)\}$ (second) has the unique GB $\{x_1, x_2, x_3^2 - x_3\}$. Points that lie on faces of the 3-cube have 2 GBs as they have 2 coordinate changes: the third set $\{(1,1,1),(0,1,0)\}$ in Figure 3 has GBs $\{\underline{x_1} - x_3, x_2 - 1, x_3^2 - x_3\}$ and $\{x_1^2 - x_1, x_2 - 1, \underline{x_3} - x_1\}$. Finally points that lie on lines through the interior have 3 GBs as they have 3 coordinate changes: the fourth set $\{(1,0,1),(0,1,0)\}$ has GBs $\{\underline{x_1} - x_3, \underline{x_2} - x_3 - 1, x_3^2 - x_3\}$, $\{\underline{x_1} - x_2 - 1, x_2^2 - x_2, \underline{x_3} - x_2 - 1\}$, and $\{x_1^2 - x_1, \underline{x_2} - x_1 - 1, \underline{x_3} + x_1\}$.
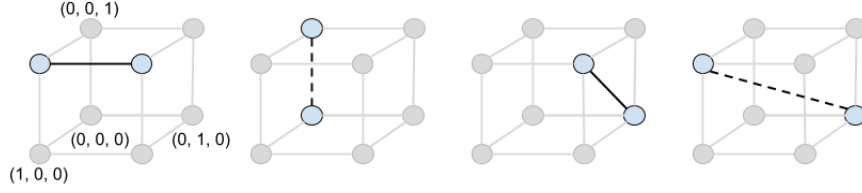
Fig. 3 Four configurations of pairs of points in $\mathbb{Z}_2^3$. From left to right: $\{(1, 0, 1), (1, 1, 1)\}$ and $\{(0, 0, 0), (0, 0, 1)\}$ have 1 GB; $\{(1, 1, 1), (0, 1, 0)\}$ has 2 GBs; and $\{(1, 0, 1), (0, 1, 0)\}$ has 3 GBs.

Next we consider data over the field $\mathbb{Z}_3$.

Example 5 Let $p = 3$ and $n = 2$. The right graph in Figure 1 is the plot of all points in $\mathbb{Z}_3^2$. Similar to the Boolean case in Figure 2, pairs of points that lie on horizontal or vertical lines have one associated reduced Gröbner basis for any monomial order, while pairs of points that lie on any skew line have two distinct GBs. For example, the set $\{(0, 0), (0, 2)\}$ in Figure 4 has ideal of points $\langle x_1, x_2^2 + x_2 \rangle$, which has a unique reduced Gröbner basis via Theorem 5. On the other hand, the set of points $\{(1, 2), (2, 1)\}$ has two GBs, namely $\{\underline{x_1} + x_2, x_2^2 + 1\}$ and $\{x_1^2 - 1, \underline{x_2} + x_1\}$ with leading term ideals $\langle x_1, x_2^2 \rangle$ and $\langle x_1^2, x_2 \rangle$ respectively.
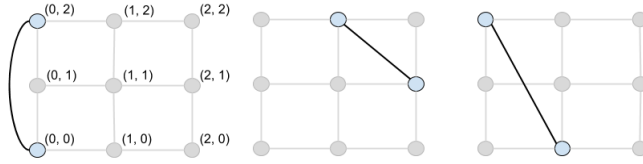

Fig. 4 Three configurations of points in $\mathbb{Z}_3^2$. From left to right: $\{(0, 0), (0, 2)\}$ has 1 GB, while $\{(1, 2), (2, 1)\}$ and $\{(0, 2), (1, 0)\}$ each have 2 distinct GBs.

In the case of $m = 2$ points, we see that data that lie on horizontal or vertical edges have ideals of points with unique Gröbner bases, that is unique models, while data whose coordinates change simultaneously have multiple models associated with them. Though the number n of coordinates impacts the number of resulting models, the field cardinality p does not.

4 Data Sets with $m = 3$ Points

Theorem 7 *The number of distinct reduced Gröbner bases for ideals of 3 points in* $\mathbb{Z}_p^n$ *is*

$$NGB(p, n, 3) \leq \begin{cases} \frac{n(n-1)}{2} & \text{for } p = 2 \\ \frac{n(n+1)}{2} & \text{for } p \geq 3. \end{cases}$$

Proof We begin by considering the Boolean base field. By Theorem 3, the form of a set of standard monomials for an ideal of 3 points is $\{1, x_i, x_j\}$ for $x_i \neq x_j$. Considering the choice of x_i and x_j, there are up to $\frac{n(n-1)}{2}$ different standard monomial sets, each corresponding to a distinct reduced Gröbner basis by Theorem 2.

For a base field with $p > 2$, the two possible forms of standard monomial sets are $\{1, x_i, x_j\}$ for $x_i \neq x_j$, and $\{1, x_i, x_i^2\}$. As we showed above, there are up to $\frac{n(n-1)}{2}$ distinct reduced Gröbner bases corresponding to $\{1, x_i, x_j\}$. Further, the maximum number for the standard monomial form $\{1, x_i, x_i^2\}$ is n. As the two standard monomial forms can both be associated to the same data set, the upper bound for a non-Boolean field is $\frac{n(n-1)}{2} + n = \frac{n(n+1)}{2}$. $\qquad\square$

Example 6 Let $p = 2$ and $n = 2$. Then $NGB(2, 2, 3) \leq 1$; that is, all ideals of 3 points in $\mathbb{Z}_2^2$ have a unique reduced Gröbner basis, which is corroborated by Theorem 4 and the fact that ideals of a single point have only one distinct Gröbner basis for any monomial order.

Unlike the bound for two points, there are sets of three points for which the upper bound is not sharp. For example when $n = 4$, the upper bound is $NGB(2, 4, 3) \leq 6$; however the maximum number is 5, which we tested exhaustively (data not shown).

Next we connect configurations of three points to the number of associated Gröbner bases. We start with Boolean data.

Example 7 Let $p = 2$ and $n = 3$. In this case, $NGB(2, 3, 3) \leq 3$. Consider the configurations of points in $\mathbb{Z}_2^3$ in Figure 5. The data set corresponding to the green triangle on the top "lid" of the leftmost 3-cube is $S_1 = \{(0, 0, 1), (0, 1, 1), (1, 0, 1)\}$ and its ideal of points has a unique Gröbner basis, namely $\{x_2^2 + x_2, x_3 + 1, x_1 x_2, x_1^2 + x_1\}$. The data set corresponding to the pink triangle in the center 3-cube is $S_2 = \{(0, 0, 1), (0, 1, 1), (1, 1, 0)\}$ and has two distinct associated GBs, with ambiguous leading terms distinguished:

$$\{x_3^2 + x_3, x_2 x_3 + x_2 + x_3 + 1, x_2^2 + x_2, \underline{x_1} + x_3 + 1\}, \{x_1 + \underline{x_3} + 1, x_2^2 + x_2, x_1 x_2 + x_1, x_1^2 + x_1\}.$$

Finally the data set corresponding to the red triangle in the rightmost 3-cube is $S_3 = \{(1, 0, 0), (0, 1, 0), (1, 1, 1)\}$ and has three GBs:

$$\{x_3^2 + x_3, x_2 x_3 + x_3, x_2^2 + x_2, \underline{x_1} + x_2 + x_3 + 1\},$$

$$\{x_3^2 + x_3, x_1 + \underline{x_2} + x_3 + 1, x_1 x_3 + x_3, x_1^2 + x_1\},$$

$$\{x_1 + x_2 + \underline{x_3} + 1, x_2^2 + x_2, x_1 x_2 + x_1 + x_2 + 1, x_1^2 + x_1\}.$$
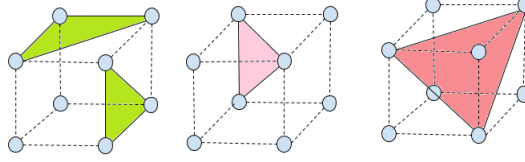
Fig. 5 Configurations of sets of 3 points in $\mathbb{Z}_2^3$ corresponding to different numbers of GBs. Points that are in configurations similar to the green triangles (left) have a unique reduced Gröbner basis for any monomial order; the pink triangle (center) has two distinct GBs; and the red triangle (right) has three distinct GBs.

The example illustrates that points that lie on faces of the 3-cube have 1 Gröbner basis; points forming a triangle which lies in the interior with 2 collinear vertices have 2 distinct GBs, and points in other configurations have 3 GBs.

Now we consider data in $\mathbb{Z}_3$.

Example 8 Let $p = 3$ and $n = 2$. By Theorem 7, we have that $NGB(3, 2, 3) \leq 3$. Consider the point configurations in Figure 6. The data set corresponding to the green triangle (left) is $S_1 = \{(0,0), (0,1), (1,1)\}$ and has a unique reduced Gröbner basis: $\{x_2^2 - x_2, x_1 x_2 - x_1, x_1^2 - x_1\}$. The data set corresponding to the pink triangle (center) is $S_2 = \{(0,1), (1,2), (2,0)\}$ and has two distinct associated reduced GBs:

$$\{x_2^3 - x_2, \underline{x_1} - x_2 + 1\}, \{-x_1 + \underline{x_2} - 1, x_1^3 - x_1\}.$$

The data set corresponding to the pink triangle (right) is $S_3 = \{(0,1), (1,2), (2,0)\}$ and also has two GBs:

$$\{x_2^3 - x_2, x_1 x_2^2 - x_1 x_2 + x_2^2 - x_2, x_1^2 - x_1 x_2 + x_1 - x_2\}, \{x_2^3 - x_2, -x_1^2 + x_1 x_2 - x_1 + x_2, x_1^3 - x_1\}.$$
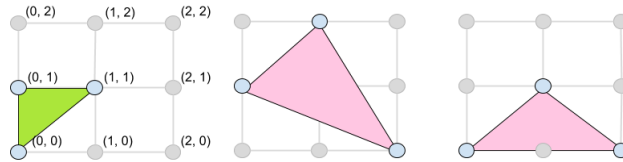

Fig. 6 Configurations of sets of 3 points in $\mathbb{Z}_3^2$ corresponding to unique and non-unique Gröbner bases. Points that are in configurations similar to the green triangle (left) have a unique reduced Gröbner basis for any monomial order; the pink triangles (center and right) have two distinct GBs.

Using Figure 6, we see that 3 points that are collinear or have two adjacent collinear points have unique Gröbner bases, while other configurations result in 2 distinct ones. There are no data sets of 3 points in $\mathbb{Z}_3^2$ that have 3 associated Gröbner bases which we verified exhaustively (data not shown). Therefore the upper bound in Theorem 7 is not sharp for $p = 3, n = 2$.

5 Geometric Observations for Larger Sets

In this section, we offer empirical observations for the number r of distinct reduced Gröbner bases for data sets of m points, where $2 \leq m \leq 6$. Furthermore, we state a conjecture for decreasing r by adding points in so-called linked positions, using the geometric insights from $m = 2, 3$ points.

To generalize the observations from small data sets to larger data sets, we start with configurations of two points, and then consider changes in r as points are added.

Definition 4 Given a set S of points, we say that a point q is in a *linked* position with respect to the points in S if q is adjacent to a point in S and has minimal sum of distances to the points in S.

Figure 7 shows the changes in the number of Gröbner bases when points are added at either linked or non-linked positions.

Example 9 Consider the set $S = \{(0, 1), (1, 2)\}$, which has $r = 2$ Gröbner bases associated to it. We aim to add a point so that the augmented set has $r = 1$. There are four points adjacent to the points in S, namely $(0, 0)$, $(0, 2)$, $(1, 1)$ and $(2, 2)$; see the green points in the top panel of Figure 7. The sum of the distances between $(0, 0)$ and the points in S is $\sqrt{5} + 1$; similarly for $(2, 2)$. On the other hand, $(0, 2)$ and $(1, 1)$ both have a distance sum of 2. So $(0, 2)$ and $(1, 1)$ are in linked positions with respect to S. Note that inclusion of either $(0, 2)$ or $(1, 1)$ to S reduces r to 1, while inclusion of either of $(0, 0)$ or $(2, 2)$ keeps $r = 2$.

Example 10 Consider the the set $S = \{(0, 1), (1, 1)\}$, which has a unique Gröbner basis. There are five points adjacent to S, namely $(0, 0)$, $(0, 2)$, $(1, 0)$, $(1, 2)$, and $(2, 1)$; see the green points in the bottom panel of Figure 7. The first four points have a distance sum of $\sqrt{2} + 1$, while the last point $(2, 1)$ has a distance sum of 3. So these four points are in linked positions with respect to S and inclusion of any one of them keeps $r = 1$. On the other hand, $(2, 1)$ is not in linked position; nevertheless adding it to S results in a unique Gröbner basis due to it being collinear to the points in S.

Adding a red point in Figure 7, which is not in a linked position with respect to the starting data set, will not reduce the number of Gröbner bases as its inclusion does not aid in removing ambiguous leading terms. In fact, the pink triangles in the last column in Figure 7 give instances in which r increases.

For $p = 3$ and $n = 2$, we computed the number of Gröbner bases for data sets up to 6 points; see Figure 8. The points at the vertices of the green polygons have $r = 1$. The uniqueness can be maintained by adding points in linked positions; however the points at the vertices of the pink polygons have non-unique Gröbner bases.

Based on the geometric observations from Figures 7 and 8, we provide heuristic rules to aid in decreasing the number of candidate models as enumerated by the number of Gröbner bases:

1. *For two points*, fewer changing coordinates in the data points will lead to fewer Gröbner bases. In the simplest case, if only one coordinate changes, a unique model will be generated.
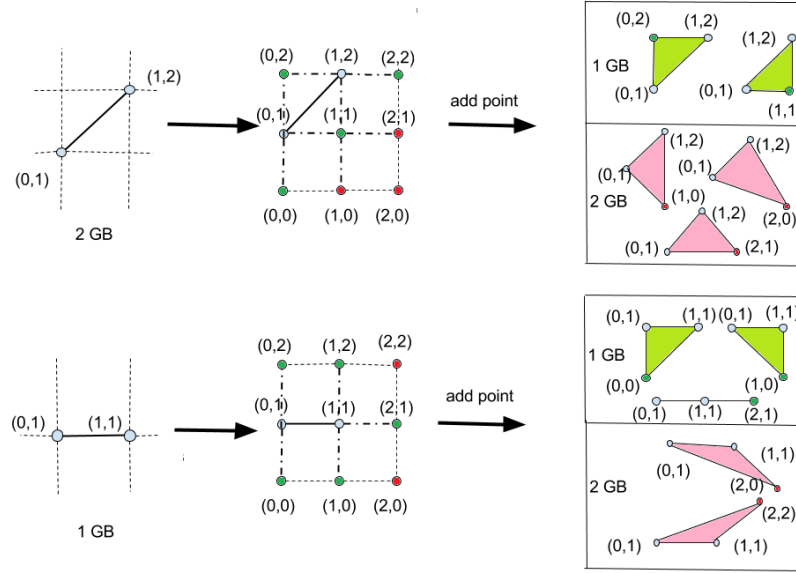
Fig. 7 The green points are adjacent to the blue points. Green triangles are associated with a unique GB, while pink triangles are associated with non-unique GBs.

2. *For three points*, more points lying on horizontal or vertical edges will reduce the number of Gröbner bases. A unique Gröbner basis arises when the data lie on a horizontal line, a vertical line or form a right triangle.
3. *In the process of adding points*, to decrease or maintain the number of minimal models, add points in linked positions with respect to an existing data set: this guarantees more points lying on horizontal or vertical edges.

By adding points in linked positions, data sets with multiple Gröbner bases can be transformed to data sets with unique one, as the following example suggests.

Example 11 Consider data sets in $\mathbb{Z}_2^4$. Let S_{max} be a data set whose ideal of points has the maximum number of Gröbner bases. Define $S_{unique} = S_{max} \cup S_{add}$ where S_{add} is a collection of points such that the augmented data set S_{unique} has an ideal of points with a unique GB. The table summarizes for different sized sets how many points must be added to guarantee a unique Gröbner basis from a data set associated with the maximum number of Gröbner bases.

max(#*GBs*)	4	5	6	13	12	13	9	13	12	13	6	5	4
$|S_{max}|$	2	3	4	5	6	7	8	9	10	11	12	13	14
$|S_{unique}|$	5	5	8	11	11	11	11	12	15	15	15	15	15
$|S_{add}|$	3	2	4	6	5	4	3	3	5	4	3	2	1

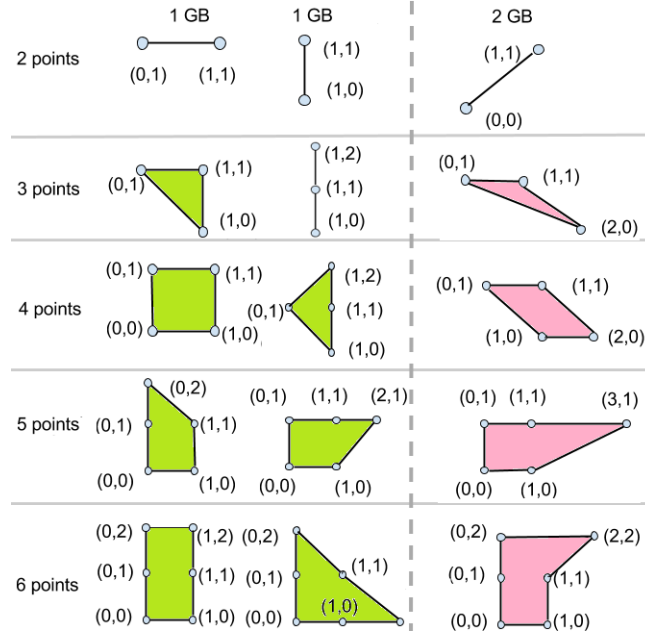We end this discussion with a conjecture about points in linked positions.

Fig. 8 Point configurations based on the number of Gröbner bases for $2 \leq m \leq 6$. The left two columns contain points that form green polygons and correspond to a unique Gröbner basis. The right column contains the pink polygons corresponding to non-unique GBs.

Conjecture 1 Let S be a set of points, $q \notin S$, and $T = S \cup \{q\}$. If q is in a linked position and the convex hull of the points in T does not contain "holes" (*i.e*, lattice points not in T), then $\#GB(T) \leq \#GB(S)$.

6 Upper Bound for the Number of Gröbner Bases

We now focus on the general setting of subsets of any size m in $\mathbb{Z}_p^n$, for any p and any n.

In Theorem 1, the stated upper bound for the number of Gröbner bases for an ideal I of m points in K^n is $m^{2n\frac{n-1}{n+1}}$, where K is any field; furthermore the number of Gröbner bases coincides with the number of vertices of the staircase polytope of I. When the base field is finite, however, this bound becomes unnecessarily large for even small m. Unlike in characteristic-0 fields, all coordinates in positive-characteristic fields are bounded above by p; for example see Figure 9. We will use the fact that staircases in a finite field are contained in a hypercube of volume p^n to modify the bound. The only part of the construction of the staircase polytope that is affected by the field characteristic is the maximum value of any vertex. As a vertex

is a vector sum $\sum \lambda$ of points in a staircase λ, the modification comes from placing staircase points aimed to maximize the sum.
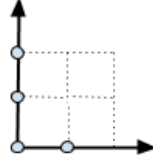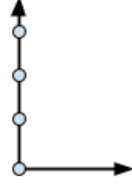

Fig. 9 The staircase $\lambda \subset \mathbb{R}^2$ (left) has $\sum \lambda = (0, 6)$ while the staircase $\lambda \subset \mathbb{Z}_3^2$ (right) has $\sum \lambda = (1, 3)$.
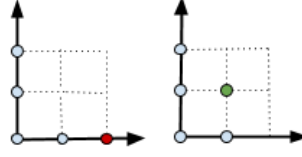
Fig. 10 The staircase $\lambda \subset \mathbb{Z}_3^2$ with red point (left) has $\sum \lambda = (3, 3)$ while the staircase $\lambda \subset \mathbb{Z}_3^2$ with green point (right) has $\sum \lambda = (2, 4)$.

Consider any staircase λ of 5 elements. In the following discussion, we will consider the placement of points so that the vector sum is maximized. We proceed in a "greedy" manner by maximizing a fixed coordinate. Suppose four (blue) points have already been placed so as to maximize the value of the second coordinate of $\sum \lambda$; see Figure 10. Placing the green point $(1, 1)$ contributes 1 to the running sum, that is, $\sum_{j=1}^{m} \lambda_{j2} = 4$ while placing the red point $(2, 0)$ keeps the sum of the coordinate unchanged. In fact, to maximize the sum of second coordinate, choose any point whose second coordinate is largest among the available positions, that is so that the configuration continues to be a staircase.

Theorem 8 *The number of distinct reduced Gröbner bases for an ideal of m points in $\mathbb{Z}_p^n$ is*

$$NGB(p, n, m) = \begin{cases} O\left(\left(p^2 \lfloor m/p \rfloor + (m \ (\mathrm{mod} \ p))^2\right)^{n^{\frac{n-1}{n+1}}}\right) & : 0 < m \le \lfloor p^n/2 \rfloor \\ O\left(\left(p^2 \lfloor (p^n - m)/p \rfloor + (-m \ (\mathrm{mod} \ p))^2\right)^{n^{\frac{n-1}{n+1}}}\right) & : \lfloor p^n/2 \rfloor \le m < p^n \\ 1 & : m = 0, p^n. \end{cases}$$

Proof Let I be an ideal of m points in $\mathbb{Z}_p^n$. Recall that the number of Gröbner bases of I is bijective with the number of vertices of the staircase polytope $\mathcal{P}$ of I by Theorem 2. The cases $m = 0, p^n$ are trivial. So we proceed with $0 < m \le \lfloor p^n/2 \rfloor$.

As $\mathcal{P}$ is the convex hull of the points $\sum \lambda$ where λ is a staircase corresponding to the exponent vectors of the standard monomial sets of I, we will show that the staircase polytope of I is contained in a larger convex body whose volume can be computed easily.

Let $\lambda = \{\lambda_1, \ldots, \lambda_m\}$. Then $\sum \lambda = \sum_{i=1}^{m} \lambda_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{m} \lambda_{ji}\right) e_i$ where λ_{ji} denotes the i-th coordinate of the j-th point and e_i is the standard basis vector. Note that the maximum sum of the i-th coordinate is

$$M := \max \sum_{j=1}^{m} \lambda_{ji} = \underbrace{(1 + \ldots + p - 1)\lfloor m/p \rfloor}_{p \lfloor m/p \rfloor \text{ points}} + \underbrace{(1 + \ldots + m \pmod{p} - 1)}_{\text{remaining } m \pmod{p} \text{ points}}$$

$$= \frac{p(p-1)}{2} \lfloor m/p \rfloor + \frac{(m \pmod{p})(m \pmod{p} - 1)}{2}.$$

So the staircase polytope $\mathcal{P} \subset \mathbb{R}^n$ is contained in the hypercube $[0, M]^n$, which has volume M^n. Therefore $vol(\mathcal{P}) \leq M^n$. By Lemma 1 and Theorem 1, we have that

$$NGB(p, n, m) = O\left(vol(\mathcal{P})^{(n-1)/(n+1)}\right)$$

$$= O\left((M^n)^{\frac{n-1}{n+1}}\right)$$

$$= O\left(\left(p^2 \lfloor m/p \rfloor + (m \pmod{p})^2\right)^{n\frac{n-1}{n+1}}\right). \tag{1}$$

For the final case when $m \geq \lfloor p^n/2 \rfloor$, the number of Gröbner bases can be computed by plugging $p^n - m$ into the second argument of the above bound, according to Theorem 4. $\qquad\square$

It is straightforward to show that our bound grows much slower than the bound $O\left(m^{2n\frac{n-1}{n+1}}\right)$ reported in [22], which we have also verified computationally. In the Appendix is a table of selected numerical results of the new upper bound in comparison to the values of the original upper bound in [22]. Figure 11 provides a comparison for selected cases among $p = 2, 3$ and $n = 2, 3, 4$.

Not only are the values from Theorem 8 closer to the actual number of GBs, including an application of Theorem 4 in our bound retains the symmetric nature of the maximum number of Gröbner bases for ideals of points in $\mathbb{Z}_p^n$. For example, for $p = 2$, $n = 4$, and $m = 5$ in Figure 11, the original bound is over 2000, while the modified bound is in the same order of magnitude as the actual maximum number of GBs.

The significance of this result is that Theorem 8 provides a more accurate representation of the maximum number of models associated to a data set, which may aid in experimental design.

7 Discussion

This work relates the geometric configuration of data points with the number of associated Gröbner bases. In particular we provide some insights into which configurations lead to uniqueness. We give an upper bound for the number of Gröbner bases for any set over a finite field. We also provide a heuristic for decreasing the number by adding points in so-called linked positions. An implication of this work is a more computationally accurate way to predict the number of distinct minimal
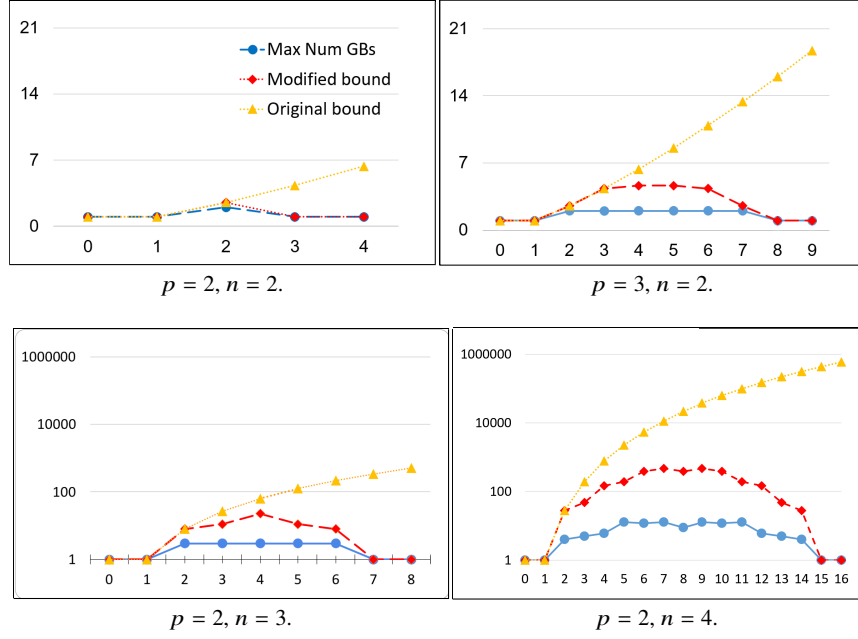
Fig. 11 Plots comparing the maximum number of Gröbner bases. The caption in each plot indicates the values of p and n for $\mathbb{Z}_p^n$. In each case, all subsets of size m are computed, where m ranges from 0 to p^n and listed on the horizontal axis. The vertical axis is the maximum number of GBs for a set of size m. The blue solid line with dots shows the actual maximum number of GBs. The yellow dotted line with triangles is the original upper bound given by Theorem 1, where the red dashed line with squares is the modified upper bound given by Theorem 8.

models which may aid modelers in estimating the computational cost before running physical experiments.

Increasing p, n or m inflates the difference between the estimated number of Gröbner bases and the actual number. The performance of the bound in Theorem 8 works well for large p and m. Though the bound is tighter than the original bound in [22], it still has large differences from the actual values for $n > 4$; see Table 5 in the Appendix. Hence, improving this bound further or finding a closed form for the number of Gröbner bases remains an important direction for future work.

References

1. G. Andrews. A lower bound for the volume of strictly convex bodies with many boundary lattice points. *Transactions of the American Mathematical Society*, 106(2):270–279, February 1963.
2. E. Babson, S. Onn, and R. Thomas. The Hilbert zonotope and a polynomial time algorithm for universal Gröbner bases. *Advances in Applied Mathematics*, 30(3):529–544, 2003.

3. B. Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*. PhD thesis, Universität Innsbruck, 1965.
4. B. Buchberger. A note on the complexity of constructing Groebner-Bases. In J. von Hulzen, editor, *Computer Algebra: Proceedings of EUROCAL 83*, volume 162 of *Lecture Notes in Computer Science*, pages 137–145. Springer Berlin, 1983.
5. B. Buchberger. Bruno Buchberger's PhD thesis 1965: An algorithm for finding the basis elements of the residue class ring of a zero dimensional polynomial ideal. *Journal of Symbolic Computation*, 41(3-4):475–511, March-April 2006.
6. B. Buchberger and M. Möller. The construction of multivariate polynomials with preassigned zeroes. In J. Calmet, editor, *Computer Algebra: EUROCAM '82*, volume 144 of *Lecture Notes in Computer Science*, pages 24–31. Springer Berlin, 1982.
7. D. Cox, J. Little, and D. O'Shea. *Using Algebraic Geometry*, volume 185. Springer Science & Business Media, 2006.
8. D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties, and Algorithms*. Springer, 2007.
9. E.S. Dimitrova. Estimating the relative volumes of the cones in a Gröbner fan. *Special issue of Mathematics in Computer Science: Advances in Combinatorial Algorithms II*, 3(4):457–466, 2010.
10. E.S. Dimitrova, Q. He, L. Robbiano, and B. Stigler. Small Gröbner fans of ideals of points. *Journal of Algebra and Its Applications*, 2019.
11. C. Eder and J.-C. Faugére. A survey on signature-based Gröbner basis computations. *Journal of Symbolic Computation*, 80(3):719–784, May 2014.
12. J. Farr and S. Gao. Computing Gröbner bases for vanishing ideals of finite sets of points. In M Fossorier, H Imai, S Lin, and et al., editors, *Applied Algebra, Algebraic Algorithms and Error-Correcting Codes*, pages 118–127. Ann N Y Acad Sci., Springer, Berlin, 2006.
13. F. Fukuda, A. Jensen, and R. Thomas. Computing Gröbner fans. *Mathematics of Computation*, 76(260):2189–2212, 2007.
14. S. Gao, A. Platzer, and E. Clarke. Quantifier elimination over finite fields using Gröbner bases. In Franz Winkler, editor, *Algebraic Informatics*, pages 140–157, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
15. W. Just and B. Stigler. Computing Gröbner bases of ideals of few points in high dimensions. *ACM Communications in Computer Algebra*, 40(3/4):67–78, September/December 2006.
16. V. Larsson, M. Oskarsson, K. Astrom, A. Wallis, T. Pajdla, and Z. Kukelova. Beyond Gröbner bases: Basis selection for minimal solvers. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3945–3954. IEEE, June 2018.
17. R. Laubenbacher and B. Stigler. A computational algebra approach to the reverse engineering of gene regulatory networks. *J. Theor. Biol.*, 229(4):523–537, 2004.
18. Z. Lin, L. Xu, and Q. Wu. Applications of Gröbner bases to signal and image processing: A survey. *Linear Algebra and its Applications*, 391:169–202, 2004.
19. R. Makarim and M. Stevens. M4GB: An efficient Gröbner-basis algorithm. In *Proceedings of the 2017 ACM on International Symposium on Symbolic and Algebraic Computation*, ISSAC 17, pages 293–300. ACM, July 2017.
20. M. Maniatis, A. von Manteuffel, and O. Nachtmann. Determining the global minimum of Higgs potentials via Groebner bases - applied to the NMSSM. *The European Physical Journal C*, 49(4):1067–1076, 2007.
21. T. Mora and L. Robbiano. The Gröbner fan of an ideal. *Journal of Symbolic Computation*, 6(2-3):183–208, 1988.
22. S. Onn and B. Sturmfels. Cutting corners. *Advances in Applied Mathematics*, 23(1):29–48, 1999.
23. M. Torrente. *Applications of Algebra in the Oil Industry*. PhD thesis, Scuola Normale Superiore di Pisa, 2009.
24. Y.-L. Tsai. Estimating the number of tetrahedra determined by volume, circumradius and four face areas using Groebner basis. *Journal of Symbolic Computation*, 77:162–174, 2016.

Appendix

We provide tables comparing of the maximum number of distinct reduced Gröbner bases to the predictions made by the original bound (third column) in Theorem 1 and the modified bound (last column) in Theorem 8. The second column shows the actual maximum number as computed for all sets in $\mathbb{Z}_p^n$ of size given in the first column. All values are rounded up to 2 decimal places.

Table 1 $p = 2, n = 2$

# of points	max # of GBs	original bound	modified bound
1	1	1	1
2	2	2.52	2.52
3	1	4.33	1
4	1	6.35	1

Table 2 $p = 2, n = 3$

# of points	max # of GBs	original bound	modified bound
1	1	1	1
2	3	8	8
3	3	27	11.18
4	3	64	22.63
5	3	125	11.18
6	3	216	8
7	1	343	1
8	1	512	1

Table 3 $p = 2, n = 4$. Half of the table is listed due to space constraints.

# of points	max # of GBs	original bound	modified bound
1	1	1	1
2	4	27.86	27.86
3	5	195.07	47.59
4	6	776.05	147.03
5	13	2264.94	195.07
6	12	5434.08	389.08
7	13	11388.61	471.48
8	9	21618.82	389.08

Table 4 $p = 3$, $n = 2$

# of points	max # of GBs	original bound	modified bound
1	1	1	1
2	2	2.52	2.52
3	2	4.33	4.33
4	2	6.35	4.64
5	2	8.55	4.64
6	2	10.90	4.33
7	2	13.39	2.52
8	1	16	1
9	1	18.72	1

Table 5 $p = 2$ and $m = 4$. Here we show how the number of Gröbner bases changes as the number of coordinates changes.

# of coordinates	max # of GBs	original bound	modified bound
2	1	6.35	1
3	3	64	22.63
4	6	776.05	147.03
5	8	10321.27	1024