Sparse Gröbner Bases: the Unmixed Case

Jean-Charles Faugere, Pierre-Jean Spaenlehauer, Jules Svartz

Sparse Gröbner Bases: the Unmixed Case

Jean-Charles Faugère[1], Pierre-Jean Spaenlehauer[2], and Jules Svartz[1]

[1]INRIA Paris-Rocquencourt, PolSys Project, Univ Paris 06, LIP6, CNRS
[2]INRIA, CNRS, Université de Lorraine, Caramel Project

Abstract

Toric (or sparse) elimination theory is a framework developped during the last decades to exploit monomial structures in systems of Laurent polynomials. Roughly speaking, this amounts to computing in a *semigroup algebra*, *i.e.* an algebra generated by a subset of Laurent monomials. In order to solve symbolically sparse systems, we introduce *sparse Gröbner bases*, an analog of classical Gröbner bases for semigroup algebras, and we propose sparse variants of the F_5 and FGLM algorithms to compute them. Our prototype "proof-of-concept" implementation shows large speed-ups (more than 100 for some examples) compared to optimized (classical) Gröbner bases software. Moreover, in the case where the generating subset of monomials corresponds to the points with integer coordinates in a normal lattice polytope $\mathscr{P} \subset \mathbb{R}^n$ and under regularity assumptions, we prove complexity bounds which depend on the combinatorial properties of $\mathscr{P}$. These bounds yield new estimates on the complexity of solving 0-dim systems where all polynomials share the same Newton polytope (*unmixed case*). For instance, we generalize the bound $\min(n_1, n_2) + 1$ on the maximal degree in a Gröbner basis of a 0-dim. bilinear system with blocks of variables of sizes (n_1, n_2) to the multilinear case: $\sum n_i - \max(n_i) + 1$. We also propose a variant of Fröberg's conjecture which allows us to estimate the complexity of solving overdetermined sparse systems. Finally, our complexity results apply in the dense (usual) case and, as a surprising by-product, we prove that restrictive assumptions in usual complexity estimates of classical inhomogeneous Gröbner bases algorithms can be removed.

1 Introduction

Context and problem statement. Many polynomial systems or systems of Laurent polynomials arising in applications do not have a dense monomial structure (*e.g* multi-homogeneous systems, fewnomials, systems invariant under the action of a linear group,...). The development of toric geometry during the 70s/80s has led to toric (or sparse) elimination theory [31], a framework designed to study and exploit algorithmically these monomial structures.

Central objects in toric geometry are *semigroup algebras* (also called toric rings). If $S \subset \mathbb{Z}^n$ is an affine semigroup (see Def. 2.1), then the semigroup algebra $k[S]$ is the set of finite sums $\sum_{s\in S} a_s X^s$, where X is a formal symbol, k is a field, $a_s \in k$ and $s \in S$. Semigroup algebras are isomorphic to subalgebras of $k[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]$ generated by a finite subset of monomials.

Our motivation is to propose fast algorithms to solve symbolically systems whose support lie in one of the following classes of semigroups: semigroups constructed from the points with integer coordinates in a normal lattice polytope $\mathscr{P} \subset \mathbb{R}^n$ (in that case, the algorithms we propose are well-suited for *unmixed* systems: the Newton polytopes of the input polynomials are all equal to $\mathscr{P}$) or semigroups generated by a scattered set of monomials (fewnomial systems).

Main results. Given a 0-dim. system of Laurent polynomials $f_1 = \cdots = f_m = 0$ and a finite subset $M \subset \mathbb{Z}^n$ such that each polynomial belongs to the subalgebra generated by $\{X_1^{\alpha_1} \cdots X_n^{\alpha_n} \mid \alpha \in M\}$, we associate to M two affine semigroups: $S_M \subset \mathbb{Z}^n$ generated by M and $S_M^{(h)} \subset \mathbb{Z}^{n+1}$ generated by $\{(\alpha, 1) \in \mathbb{Z}^{n+1} \mid \alpha \in M\}$. Under the assumption that S_M contains zero but no nonzero pairs $(s_1, s_2) \in S_M^2$ s.t. $s_1 + s_2 = \mathbf{0}$, our solving strategy proceeds by combining a sparse variant in the homogeneous algebra $k[S_M^{(h)}]$ of the `MatrixF5` algorithm and a sparse variant in $k[S_M]$ of the `FGLM` algorithm. We define a notion of *sparse Gröbner basis* (Def. 3.1) that is computed by the `sparse-MatrixF5` algorithm if we know a bound on its maximal degree (this maximal degree is called the *witness degree* of the system). An important feature of sparse GBs is that their definition depends only on the ambient semigroup algebra and not on an embedding in a polynomial algebra. In this sense, they differ conceptually from SAGBI bases, even though the `sparse-FGLM` algorithm has similarities with the `SAGBI-FGLM` algorithm proposed in [17]. In the special case $S_M = \mathbb{N}^n$, sparse Gröbner bases in $k[S_M]$ are classical Gröbner bases, and `sparse-FGLM` is the usual `FGLM` algorithm.

At the end of the solving process, we obtain a rational parametrisation of the form

$$Q(T) = 0 \quad \text{and} \quad \forall \alpha \in M \setminus \{\mathbf{0}\}, \quad X_1^{\alpha_1} \cdots X_n^{\alpha_n} - Q_\alpha(T) = 0$$

where $Q \in k[T]$ is a univariate polynomial, and for all $\alpha \in M$, $Q_\alpha \in k(T)$ is a rational function. Consequently, the solutions of the input sparse system can be expressed in terms of the roots of the univariate polynomial Q by inverting a monomial map.

The next main result addresses the question of the complexity of this solving process when M is given as the set $\mathscr{P} \cap \mathbb{Z}^n$, where $\mathscr{P} \subset \mathbb{R}^n$ is a lattice polytope of dimension n. It turns out that the complexities of `sparse-MatrixF5` and `sparse-FGLM` algorithms depend mainly on intrinsic combinatorial properties of $\mathscr{P}$:

- the normalized volume $\mathrm{vol}(\mathscr{P}) \in \mathbb{N}$;
- the Castelnuovo-Mumford regularity $\mathrm{reg}(k[S^{(h)}_{\mathscr{P}\cap\mathbb{Z}^n}]) = n + 1 - \ell$ where ℓ is the smallest integer such that the intersection of $\mathbb{Z}^n$ with the interior of $\ell \cdot \mathscr{P}$ is nonempty;
- the Ehrhart polynomial $\mathrm{HP}_{\mathscr{P}}(\ell)$ which equals the cardinality of $(\ell \cdot \mathscr{P}) \cap \mathbb{Z}^n$ for $\ell \in \mathbb{N}$.

We use as indicator of the complexity the *witness degree* which bounds the maximal "sparse degree" (corresponding to an $\mathbb{N}$-grading on $k[S^{(h)}_{\mathscr{P}\cap\mathbb{Z}^n}]$) in a reduced sparse Gröbner basis. More precisely, we obtain the following complexity estimates:

Theorem 1.1. *Let $\mathscr{P} \subset \mathbb{R}^n$ be a normal lattice polytope of dimension n with one vertex at $\mathbf{0} \in \mathbb{Z}^n$, $(d_1, \ldots, d_n)$ be a sequence of positive integers and $(f_1, \ldots, f_n)$ be a regular sequence of Laurent polynomials in $k[X_1^{\pm 1}, \ldots, X_n^{\pm 1}]^n$, such that the support of f_i is included in $\{X_1^{s_1} \cdots X_n^{s_n} \mid s \in (d_i \cdot \mathscr{P}) \cap \mathbb{Z}^n\}$. Then a sparse GB of the ideal $\langle f_1, \ldots, f_n\rangle \subset k[S_{\mathscr{P}\cap\mathbb{Z}^n}]$ can be computed within*

$$O\left(n \operatorname{HP}_{\mathscr{P}}(\mathrm{d}_{\mathrm{wit}})^{\omega}\right)$$

*arithmetic operations in k, where $\omega < 2.373$ is a feasible exponent for the matrix multiplication and $\mathrm{d}_{\mathrm{wit}} \leq \operatorname{reg}(k[\mathscr{P}]) + 1 + \sum_{j=1}^{n}(d_j - 1)$. Moreover, if $\mathbf{0}$ is a simple vertex of $\mathscr{P}$ (*i.e. *a vertex which is the intersection of n facets), then the* `sparse-FGLM` *algorithm executes at most*

$$O\left(\operatorname{HP}_{\mathscr{P}}(1)\left(\operatorname{vol}(\mathscr{P})\prod_{j=1}^{n} d_j\right)^3\right)$$

arithmetic operations in k.

Direct consequences of these formulas allow us to derive new complexity bounds for solving regular multi-homogeneous systems. We show that the witness degree of a regular system of n multi-homogeneous polynomials of multi-degree $(d_1, \ldots, d_p)$ w.r.t. blocks of variables of sizes $(n_1, \ldots, n_p)$ (with $\sum n_i = n$) is bounded by $n+2-\max_{i\in\{1,\ldots,p\}}(\lceil (n_i+1)/d_i \rceil)$ (which generalizes the bound $\min(n_1, n_2) + 1$ in the bilinear case [18]). We also propose a variant of Fröberg's conjecture for sparse systems and a notion of semi-regularity, which yield complexity estimates for solving sparse overdetermined systems. A surprising by-product of our approach is that it also yields new results for classical Gröbner bases algorithms of inhomogeneous systems: the assumption that the part of highest degree has to be regular is actually not needed for known complexity bounds.

We have implemented in C a prototype of the `sparse-MatrixF5` algorithm, that runs several times faster than the original F_5 algorithm in the FGb software. For instance, we report speed-up ratios greater than 100 for instances of overdetermined bihomogeneous systems. The implementation also works well for fewnomial systems (although this case is not covered by our complexity analysis).

Related works. Computational aspects of toric geometry and Gröbner bases are investigated in [32]. In particular, [32, Subroutine 11.18] gives an algorithm to compute syzygies of monomials in toric rings, which is an important routine for critical-pairs based algorithms.

Other approaches have been designed to take advantage of the sparse structure in Gröbner bases computations. For instance, the Slim Gröbner bases in [4] describes strategies to avoid increasing the number of monomials during computations. This approach improves practical computations, but does not lead to new asymptotic complexity bounds for classes of sparse systems.

The sparse structure and the connection with toric geometry have also been incorporated to the theory of resultants, and a vast literature has been written on this topic, see *e.g.* [7, 8, 13, 14]. In particular, mixed monomials structures are well-understood in this context. Although we do not know how to extend the algorithms proposed in this paper to mixed structures, Gröbner-type algorithms enjoy the property of extending without any modification to the overdetermined case.

Perspectives. Our approach is for the moment limited to *unmixed systems*: all input polynomials have to lie in the same semigroup algebra. A possible extension of this work would be the generalization to mixed systems (where the algorithms would depend on the Newton polytope of each of the polynomials of the system). Some results seem to indicate that such a generalization may be possible: for instance, under genericity assumptions, mixed monomial bases of quotient algebras are explicitly described in [27]. Also, a bound on the witness degree and the complexity analysis is for the moment restricted to the polytopal case. Merging the approach in this paper with a Buchberger's type approach such as [32, Algo. 11.17] could lead to a termination criterion of the `sparse-MatrixF5` algorithm in the non-regular cases and for positive dimensional systems. Finally, finding complexity bounds which explain the efficiency of the sparse Gröbner bases approach for fewnomial systems (see Table 3) remains an open problem.

Organisation of the paper. We recall in Section 2 the background material on semigroup algebras and convex geometry that will be used throughout this paper. Section 3 introduces sparse Gröbner bases and describes a general solving process for sparse systems. The main algorithms are described in Section 4 and their complexities are analyzed in Section 5. Finally, we describe in Section 6 some results that are direct consequences of this new framework and experimental results in Section 7.

Acknowledgements. We are grateful to Kaie Kubjas, Guillaume Moroz and Bernd Sturmfels for helpful discussions and for pointing out important references. This work was partly done while the second author was supported and hosted by the Max Planck Institute for Mathematics (Bonn, Germany).

2 Preliminaries and notations

In this paper, the basic algebraic objects corresponding to monomials in classical polynomial rings are *affine semigroups*. We always consider them embedded in $\mathbb{Z}^n$. We refer the reader to [9, 20, 25] for a more detailed presentation of this background material. First, we describe the main notations that will be used throughout the paper:

Definition 2.1. *An* affine semigroup *S is a finitely-generated additive subsemigroup of $\mathbb{Z}^n$ for some $n \in \mathbb{N}$ containing $\mathbf{0} \in \mathbb{Z}^n$ and no nonzero invertible element (*i.e. *for all $s, s' \in S \setminus \{\mathbf{0}\}, s + s' \neq \mathbf{0}$). Any affine semigroup has a unique minimal set of generators, called the* Hilbert basis *of S and denoted by* $\mathrm{Hilb}(S)$. *Let* $\mathrm{gp}(S)$ *denote the smallest subgroup of $\mathbb{Z}^n$ containing S. Then S is called* normal *if $S = \{q \in \mathrm{gp}(S) \mid \exists c \in \mathbb{N}, c \cdot q \in S\}$. For a field k, we let $k[S]$ denote the associated* semigroup algebra *of finite formal sums $\sum_{s \in S} a_s X^s$ where $a_s \in k$. An element $X^s \in k[S]$ is called a* monomial.

We use the letter M to denote a finite subset of $\mathbb{Z}^n$ such that $\mathbf{0} \in M$ and the semigroup S_M generated by M contains no nonzero invertible element. Also, we let $S_M^{(h)}$ denote the affine semigroup generated by $\{(\alpha, 1) \mid \alpha \in M\} \subset \mathbb{Z}^{n+1}$. The semigroup algebra $k[S_M^{(h)}]$ is homogeneous *(*i.e. $\mathbb{N}$*-graded and generated by degree* 1 *elements): the degree of a monomial $X^{(s_1,\ldots,s_n,d)}$ is $d \in \mathbb{N}$. The vector space of homogeneous elements of degree $d \in \mathbb{N}$ in $k[S_M^{(h)}]$ is denoted by $k[S_M^{(h)}]_d$.*

Depending on the articles on this topic, the condition "S contains no invertible element" is not always included in the definition of an affine semigroup. However, this is a necessary condition for the algorithms we propose in this paper. Also, the term "Hilbert basis" is sometimes reserved for affine semigroups of the form $\mathscr{C} \cap \mathbb{Z}^n$ where $\mathscr{C}$ is a rational cone (see *e.g.* [25, Prop. 7.15] and the discussion after this statement). We always assume implicitely that $\mathrm{gp}(S) \subset \mathbb{Z}^n$ is a full rank lattice (this does not lose any generality since this case can be reached by embedding S in a lower dimensional $\mathbb{Z}^{n'}$). Note that $k[\mathbb{N}^n]$ is the classical polynomial ring $k[X_1, \ldots, X_n]$. Semigroup algebras are integral domains [25, Thm. 7.4] of Krull dimension n and play an important role in toric geometry: they are precisely the coordinate rings of *affine toric varieties*.

The normality of the semigroup S is an important property which implies that $k[S]$ is Cohen-Macaulay by a theorem by Hochster [22]. An important feature of normal affine semigroups is that they can be represented by the intersection of $\mathbb{Z}^n$ with a pointed rational polyhedral cone (also called *strongly convex rational polyhedral cone* [26, Sec 1.1]).

Definition 2.2. *A* cone *$\mathscr{C} \subset \mathbb{R}^n$ is a convex subset of $\mathbb{R}^n$ stable by multiplication by $\mathbb{R}_+$, the set of non-negative real numbers. The dimension $\dim(\mathscr{C})$ of a cone $\mathscr{C}$ is the dimension of the linear subspace spanned by $\mathscr{C}$. A cone is called* pointed *if it does not contain any line. A pointed cone of dimension* 1 *is called a ray. A ray is called* rational *if it contains a point in $\mathbb{Z}^n$. A* rational polyhedral cone *is the convex hull of a finite number of rational rays. Pointed rational polyhedral cones will be abbreviated* PRPC*.*

We shall use PRPCs in Section 3 to define admissible monomial orderings in semigroup algebras. We now recall the definition of *simplicial affine semigroups*, for which we will be able to derive tight complexity bounds for the `sparse-FGLM` algorithm (Section 5).

Definition 2.3. *An affine semigroup $S \subset \mathbb{Z}^n$ is called simplicial if the convex hull of $\mathbb{R}_+ S$ is a simplicial PRPC,* i.e. *the convex hull of n linearly independant rays.*

Another important family of objects are *projective toric varieties*. Their homogeneous coordinate rings are associated to a lattice polytope, which we shall assume to be normal in order to ensure that the coordinate ring is Cohen-Macaulay. As in the classical case, homogeneity is a central concept to analyze the complexity of Gröbner bases algorithms. All lattice polytopes will be assumed to be full dimensional.

Definition 2.4. *A* lattice polytope *$\mathscr{P} \subset \mathbb{R}^n$ is the convex hull of a finite number of points in $\mathbb{Z}^n$. Its* normalized volume*,* i.e. *$n!$ times its Euclidean volume, is denoted by* $\mathrm{vol}(\mathscr{P}) \in \mathbb{N}$*.*

To a lattice polytope $\mathscr{P} \subset \mathbb{R}^n$ is associated an affine semigroup $S^{(h)}_{\mathscr{P}\cap\mathbb{Z}^n} \subset \mathbb{Z}^{n+1}$ generated by $\{(\alpha, 1) \mid \alpha \in \mathscr{P} \cap \mathbb{Z}^n\}$. The polytope $\mathscr{P}$ is called normal *if $S^{(h)}_{\mathscr{P}\cap\mathbb{Z}^n}$ is a normal semigroup. The associated semigroup algebra is called a* polytopal algebra *and abbreviated $k[\mathscr{P}]$.*

If $\mathscr{P} \subset \mathbb{R}^n$ is a lattice polytope containing $\mathbf{0}$ as a vertex, then $k[\mathscr{P}] = k[S^{(h)}_{\mathscr{P}\cap\mathbb{Z}^n}]$ (Def. 2.1). Moreover, if $\mathscr{P}$ is normal, then so is $S_{\mathscr{P}\cap\mathbb{Z}^n}$ [9, Prop. 2.17]. Also, note that if $\mathscr{P}'$ is a translation of $\mathscr{P}$, then the homogeneous algebras $k[\mathscr{P}]$ and $k[\mathscr{P}']$ are isomorphic. Consequently, we shall assume w.l.o.g. in the sequel that one of the vertices of $\mathscr{P}$ is the origin, so that $M = \mathscr{P} \cap \mathbb{Z}^n$ verifies the assumptions of Def. 2.1. We also introduce a few more notations for lattice polytopes:

Notation 2.5. *The number of lattice points in a polytope $\mathscr{P} \subset \mathbb{R}^n$ (*i.e. *the cardinality of $\mathscr{P} \cap \mathbb{Z}^n$) is denoted by $\#\mathscr{P}$. The* Minkowsky sum *of two lattice polytopes $\mathscr{P}_1, \mathscr{P}_2 \subset \mathbb{R}^n$ is the lattice polytope $\{p_1 + p_2 \mid p_1 \in \mathscr{P}_1, p_2 \in \mathscr{P}_2\}$. For all $\ell \in \mathbb{N}$ we write $\ell \cdot \mathscr{P}$ for the Minkowski sum $\mathscr{P} + \cdots + \mathscr{P}$ with ℓ summands. For $n \in \mathbb{N}$, we let $\Delta_n \subset \mathbb{R}^n$ denote the* standard simplex*, namely the convex hull of $\mathbf{0}$ and of the points $\mathbf{e}_i \in \mathbb{R}^n$ whose entries are zero except for the ith coefficient which is equal to 1. For $\mathscr{P}_1 \subset \mathbb{R}^i$,$\mathscr{P}_2 \subset \mathbb{R}^j$ we write $\mathscr{P}_1 \times \mathscr{P}_2 \subset \mathbb{R}^{i+j}$ for the lattice polytope whose points are $\{(p_1, p_2) \mid p_1 \in \mathscr{P}_1, p_2 \in \mathscr{P}_2\}$.*

Next, we recall several useful classical properties of polytopal algebras. We refer to [25, Ch. 12] for a detailed presentation of the connections between Ehrhart theory and computational commutative algebra.

Proposition 2.6. *Let $\mathscr{P} \subset \mathbb{R}^n$ be a lattice polytope. For $d \in \mathbb{N}$, we let $\mathrm{HP}_{\mathscr{P}} \in \mathbb{Q}[d]$ denote the* Ehrhart polynomial *of $\mathscr{P}$,* i.e. *$\mathrm{HP}_{\mathscr{P}}(d) = \#(d \cdot \mathscr{P})$. Also, let $\mathrm{HS}_{\mathscr{P}}(t) \in \mathbb{Z}[[t]]$ denote the generating series*

$$\mathrm{HS}_{\mathscr{P}}(t) = \sum_{d\in\mathbb{N}} \mathrm{HP}_{\mathscr{P}}(d)t^d.$$

Then the Hilbert series of the polytopal algebra $k[\mathscr{P}]$, namely

$$\mathrm{HS}_{k[\mathscr{P}]}(t) = \sum_{d\in\mathbb{N}} \dim_k(k[\mathscr{P}]_d)t^d$$

is equal to $\mathrm{HS}_{\mathscr{P}}$ and there exists a polynomial $Q \in \mathbb{Z}[t]$ with non-negative coefficients such that

$$\mathrm{HS}_{\mathscr{P}}(t) = \frac{Q(t)}{(1-t)^{n+1}}, \quad \deg(Q) \leq n.$$

Proof. The fact that the map $\mathrm{HP}_{\mathscr{P}} : d \mapsto \#(d \cdot \mathscr{P})$ is polynomial is a classical result by Ehrhart [11]. The second statement $\mathrm{HS}_{\mathscr{P}} = \mathrm{HS}_{k[\mathscr{P}]}$ follows from the definition of $k[\mathscr{P}]$. The last statement is Stanley's non-negativity theorem [29, Thm. 2.1]. □

We let $\mathrm{reg}(k[\mathscr{P}])$ denote the *Castelnuovo-Mumford regularity* of $k[\mathscr{P}]$. The Castenuovo-Mumford regularity of a graded module is an important measure of its "complexity": it is related to the degrees where its local cohomology modules vanish. We refer to [5, Ch. 15] for a more detailed presentation. The following classical proposition relates the regularity with a combinatorial property of the polytope $\mathscr{P}$ and with the degree of the numerator of $\mathrm{HS}_{\mathscr{P}}$:

Proposition 2.7. *Let $\mathscr{P}$ be a normal lattice polytope. The regularity* $\mathrm{reg}(k[\mathscr{P}])$ *equals* $n-\ell+1$*, where* ℓ *is the smallest integer such that* $\ell\cdot\mathscr{P}$ *contains an integer point in its interior. Moreover, with the same notations as in Proposition 2.6,* $\deg(Q)=\mathrm{reg}(k[\mathscr{P}])$.

Proof. The first claim follows from [6, Sec. 5.4]. To prove the second claim, we use the partial fraction expansion of $\mathrm{HS}_{\mathscr{P}}$ which is of the form $\sum_{\ell=n+1-\deg(Q)}^{n+1}\frac{a_\ell}{(1-t)^\ell}$ with $a_{n+1-\deg(Q)}\neq 0$. Then we obtain the equality $\mathrm{HP}_{\mathscr{P}}(d)=\sum_{\ell=n+1-\deg(Q)}^{n+1}\frac{a_\ell}{(\ell-1)!}\prod_{j=1}^{\ell-1}(d+j)$, and hence $d=n-\deg(Q)+1$ is the smallest positive integer such that $\mathrm{HP}_{\mathscr{P}}(-d)\neq 0$. The Ehrhart-MacDonald reciprocity [24] concludes the proof. □

3 Sparse Gröbner bases

In this section, we show that classical Gröbner bases algorithms extend to the context of semigroup algebras. First, we need to extend the notion of admissible monomial ordering and of Gröbner bases. We recall that the monomials of a semigroup algebra $k[S]$ are the elements X^s for $s\in S$.

Definition 3.1. *Let S be an affine semigroup. A total ordering on the monomials of $k[S]$ is called* admissible *if*

- *it is compatible with the internal law of S: for any $s_1,s_2,s_3\in S$, $X^{s_1}\prec X^{s_2}\Rightarrow X^{s_1+s_3}\prec X^{s_2+s_3}$;*
- *for any $s\in S\setminus\{\mathbf{0}\}$, $X^{\mathbf{0}}\prec X^s$.*

For a fixed admissible ordering $\prec$ and for any element $f\in k[S]$, we let $\mathrm{LM}(f)$ *denote its leading monomial. Similarly, for any ideal $I\subset k[S]$,* $\mathrm{LM}(I)$ *denotes the ideal generated by* $\{\mathrm{LM}(f)\mid f\in I\}$*. A finite subset $G\subset I$ is called a* sparse Gröbner basis *(abbreviated sGB) of I with respect to $\prec$ if the set* $\{\mathrm{LM}(g)\mid g\in G\}$ *generates* $\mathrm{LM}(I)$ *in* $k[S]$.

Note that admissible orderings exist for any semigroup algebra: the convex hull of a semigroup $S\subset\mathbb{Z}^n$ is a PRPC $\mathscr{C}\subset\mathbb{R}^n$ (this is a consequence of the fact that there is no nonconstant invertible monomial in $k[S]$). Now one can pick n independant linear forms $(\ell_1,\ldots,\ell_n)$ with integer coefficients in the dual cone $\mathscr{C}^*=\{\text{linear forms }\ell:\mathbb{R}^n\to\mathbb{R}\mid\forall\mathbf{x}\in\mathscr{C},\ell(\mathbf{x})\geq 0\}$, and set $X^{s_1}\prec X^{s_2}$ if and only if the vector $(\ell_1(s_1),\ldots,\ell_n(s_1))$ is smaller than $(\ell_1(s_2),\ldots,\ell_n(s_2))$ for a classical admissible ordering on $\mathbb{N}^n$.

Note that the assumption that $k[S]$ contains no nonconstant invertible monomial is a necessary and sufficient condition for the existence of an admissible ordering.

We describe now an algorithmic framework to solve sparse systems of Laurent polynomials. Let $M\subset\mathbb{Z}^n$ be a finite subset verifying the assumptions of Definition 2.1, and $f_1,\ldots,f_m\in k[X_1^{\pm1},\ldots,X_n^{\pm1}]$ be Laurent polynomials such that the supports of the f_i are included in $\{X_1^{\alpha_1}\cdots X_n^{\alpha_n}\mid\alpha\in S_M\}$. Note that translating M amounts to multiplying the Laurent polynomials by Laurent monomials: this does not change the set of solutions of the system in the torus $\left(\overline{k}\setminus\{0\}\right)^n$.

Assuming that the system $f_1 = \cdots = f_m = 0$ has finitely-many solutions in $(\overline{k} \setminus \{0\})^n$, we proceed as follows:

1. homogenize $(f_1, \ldots, f_m)$ via Def.-Prop. 3.3 (note that the homogenization depends on the choice of the (not necessarily minimal) generating set M;
2. compute a sparse Gröbner basis w.r.t. a graded ordering of the homogeneous ideal $I = \langle f_1^{(h)}, \ldots, f_m^{(h)} \rangle \subset k[S_M^{(h)}]$ by using a variant of F_4/F_5 algorithm (Algo. 1).
3. dehomogenize the output to obtain a sGB of the ideal $\langle f_1, \ldots, f_m \rangle \subset k[S_M]$ (Prop. 3.5);
4. use a sparse variant of FGLM to obtain a 0-dim. triangular system (hence containing a univariate polynomial) whose solutions are the image of the toric solutions of $f_1 = \cdots = f_m = 0$ by monomial maps (Algo. 2);
5. compute the non-zero roots of the univariate polynomial and invert the monomial map to get the solutions.

We focus on the four first steps of this process. The fifth step involves computing the roots of a univariate polynomial, for which dedicated techniques exist and depend on the field k. It also involves inverting a monomial map, which can be achieved by solving a consistent linear system of $\#\operatorname{Hilb}(S_M)$ equations in n unknowns.

In the sequel of this section, we investigate the behavior of sparse Gröbner bases under homogenization and dehomogeneization (Steps 1 and 3). We refer the reader to [9, Ch. 2] for geometrical aspects of projective toric varieties and their affine charts. If M verifies the assumptions of Def. 2.1, then there is a canonical dehomogenization map:

Definition 3.2. *With the notations of Def. 2.1, there is a* dehomogeneization morphism χ_M *defined by*

$$\begin{array}{rccc} \chi_M : & k[S_M^{(h)}] & \to & k[S_M] \\ & X^{(s,d)} & \mapsto & X^s \end{array}$$

Definition-Proposition 3.3. *With the notations of Def. 2.1, for any $f \in k[S_M]$, we call degree of f, the number* $\deg(f) = \min\{d \in \mathbb{N} \mid \chi_M^{-1}(f) \cap k[S_M^{(h)}]_d \neq \emptyset\}$. *Moreover the set $\chi_M^{-1}(f) \cap k[S_M^{(h)}]_{\deg(f)}$ contains a unique element, called the* homogenization *of f.*

Proof. The only statement to prove is that $\chi_M^{-1}(f) \cap k[S_M^{(h)}]_{\deg(f)}$ contains a unique element. Let $f_1^{(h)}, f_2^{(h)} \in \chi_M^{-1}(f) \cap k[S_M^{(h)}]_{\deg(f)}$. Then $\chi_M(f_1^{(h)} - f_2^{(h)}) = 0$, which implies $f_1^{(h)} = f_2^{(h)}$. □

The next step is to prove that dehomogenizing a homogeneous Gröbner basis (with respect to a graded ordering) gives a Gröbner basis of the dehomogenized ideal.

Definition 3.4. *An admissible monomial ordering $\prec$ on $k[S_M^{(h)}]$ is called* graded *if there exists an associated ordering $\prec'$ on $k[S_M]$ such that*

$$X^{(s_1,d_1)} \prec X^{(s_2,d_2)} \Leftrightarrow \left\{ \begin{array}{c} d_1 < d_2 \text{ or} \\ d_1 = d_2 \text{ and } X^{s_1} \prec' X^{s_2} \end{array} \right.$$

Proposition 3.5. *Let G be an homogeneous sGB of an homogeneous ideal $I \subset k[S_M^{(h)}]$ with respect to a graded ordering. Then $\chi_M(G)$ is a sGB of $\chi_M(I)$ with respect to the associated ordering on $k[S_M]$.*

Proof. First, notice that χ_M commutes with leading monomials on homogeneous components of $k[S_M^{(h)}]$: for any $f \in k[S_M^{(h)}]_d$, $\chi_M(\mathrm{LM}(f)) = \mathrm{LM}(\chi_M(f))$. Let $f \in \chi_M(I)$ and $f^{(h)} \in I$ be a homogeneous polynomial such that f is equal to $\chi_M(f^{(h)})$. Consequently, there exists $g \in G$ such that $\mathrm{LM}(g)$ divides $\mathrm{LM}(f^{(h)})$. Applying χ_M, we obtain that $\mathrm{LM}(\chi_M(g))$ divides $\mathrm{LM}(\chi_M(f^{(h)})) = \mathrm{LM}(f)$. Therefore $\chi_M(G)$ is a sGB of $\chi_M(I)$ for the associated ordering. □

4 Algorithms

4.1 Sparse-MatrixF5 algorithm

As pointed out in [23], classical Gröbner bases algorithms are related to linear algebra via the Macaulay matrices. Since $k[S_M^{(h)}]$ is generated by elements of degree 1, the following proposition shows that similar matrices can be constructed in the case of semigroup algebras:

Proposition 4.1. *Any monomial of degree d in $k[S_M^{(h)}]$ is equal to a product of a monomial of degree $d-1$ by a monomial of degree 1.*

With the notations of Def. 2.1, $k[S_M^{(h)}]$ has the following property: for any $f_1 \in k[S_M^{(h)}]_d$, and for all $\ell \geq d$, there exists $f_2 \in k[S_M^{(h)}]_\ell$ s.t. $\chi_M(f_1) = \chi_M(f_2)$. This leads to the following definition of a D-Gröbner basis:

Definition 4.2. *Let $I \subset k[S_M^{(h)}]$ be a homogeneous ideal and $\prec$ be an admissible monomial ordering on $k[S_M]$. Then a finite subset $G \subset I$ is called a D-sGB of I if for any homogeneous polynomial $f \in I$ with $\deg(f) \leq D$, there exists $g \in G$ such that $\mathrm{LM}(g)$ divides $\mathrm{LM}(f)$.*

Note that for any $D \in \mathbb{N}$ there always exists a homogeneous D-sGB of I. A D-sGB of I can be deduced from a row echelon basis of the k-vector space $I \cap k[S_M^{(h)}]_D$, and can be computed via the Macaulay matrix:

Definition 4.3. *Let $f_1, \ldots, f_m$ be homogeneous polynomials in $k[S_M^{(h)}]$. Then the* Macaulay matrix *in degree $d \in \mathbb{N}$ of $f_1, \ldots, f_m$ is a matrix with $\sum_{i=1}^m \max(\mathrm{HF}_{k[S_M^{(h)}]}(d - \deg(f_i)), 0)$ rows, $\mathrm{HF}_{k[S_M^{(h)}]}(d)$ columns and entries in k (where $\mathrm{HF}_{k[S_M^{(h)}]}$ is the Hilbert function of $k[S_M^{(h)}]$). Rows are indexed by the products $X^{(s,d-\deg(f_i))} \cdot f_i$ where $X^{(s,d-\deg(f_i))} \in k[S_M^{(h)}]$. Columns are indexed by monomials of degree d and are sorted in decreasing order w.r.t. an admissible monomial ordering. The entry at the intersection of the row $X^{(s,d-\deg(f_i))} \cdot f_i$ and the column $X^{(s',d)}$ is the coefficient of $X^{(s',d)}$ in $X^{(s,d-\deg(f_i))} \cdot f_i$.*

By a slight abuse of notation, we identify implicitely a row in the Macaulay matrix of degree d with the corresponding polynomial in $k[S_M^{(h)}]_d$. The relation between the Macaulay matrix and a D-sGB is given by:

Definition-Proposition 4.4. *Let $f_1, \ldots, f_m \in k[S_M^{(h)}]$ be homogeneous polynomials, $\prec$ a graded monomial ordering, and for $d \in \mathbb{N}$, let G_d be the set of polynomials corresponding to the rows of the reduced row echelon form of the Macaulay matrix in degree d of $f_1, \ldots f_m$. Then we have*

$$\begin{gathered}\text{for any } D \in \mathbb{N}\ , G_0 \cup \cdots \cup G_D \text{ is a } D\text{-sGB of } I,\\ \text{and } \chi_M(G_0) \subset \chi_M(G_1) \subset \chi_M(G_2) \subset \ldots\end{gathered}$$

The smallest integer ℓ such that $\chi_M(G_\ell)$ is a sGB of the ideal $\chi_M(\langle f_1, \ldots, f_m\rangle)$ is called the witness degree *and noted* $\mathrm{d}_{\mathrm{wit}}$.

Proof. The first statement ($G_0 \cup \cdots \cup G_D$ is a D-sGB of I) follows from the fact that G_d is a triangular basis of the vector space $k[S_M^{(h)}]_d$. The second statement is deduced from the inclusions $\chi_M(k[S_M^{(h)}]_0) \subset \chi_M(k[S_M^{(h)}]_1) \subset \ldots$. Let G be a sGB of $\langle f_1, \ldots, f_m\rangle$. Then $\mathrm{d}_{\mathrm{wit}}$ is bounded above by $\max\{\deg(g) \mid g \in G\}$ and is therefore finite. □

As in the original F_5 algorithm [15], many lines are reduced to 0 during row-echelon form computations of Macaulay matrices. The F_5 criterion [10, 15][2, Prop. 6] extends without any major difficulty in this context and identifies all reductions to zero when the input system is a regular sequence in $k[S_M^{(h)}]$:

Lemma 4.5 (F_5-criterion)**.** *With the notations of Algorithm 1, if m is the leading monomial of a row in $\widetilde{\mathcal{M}}_{d-d_i,i-1}$ then the polynomial mf_i belongs to the vector space*

$$\mathrm{Span}_k(\mathit{Rows}(\mathcal{M}_{d,i-1}) \cup \{uf_i \mid u \in k[S_M^{(h)}]_{d-d_i} \text{ and } u \prec m\}).$$

Algorithm 1: `sparse-MatrixF5`

Input : Homogeneous $f_1, \ldots, f_m \in k[S_M^{(h)}]$ of resp. degrees $(d_1, \ldots, d_m)$, a graded monomial ordering $\prec$ on $k[S_M^{(h)}]$, a maximal degree D
Output: a D-Gröbner basis of $\langle f_1, \ldots, f_m\rangle$ w.r.t. $\prec$
for $i = 1$ **to** m **do** $\mathcal{G}_i := \emptyset$;
for $d = 1$ **to** D **do**
  $\mathcal{M}_{d,0} := \emptyset$, $\widetilde{\mathcal{M}}_{d,0} := \emptyset$;
  for $i = 1$ **to** m **do**
    case $d_i > d$: $\mathcal{M}_{d,i} := \widetilde{\mathcal{M}}_{d,i-1}$;
    case $d_i = d$: $\mathcal{M}_{d,i}$:=add new row f_i to $\widetilde{\mathcal{M}}_{d,i-1}$;
    case $d_i < d$: add new row $X^{(s,d-d_i)}f_i$ to $\widetilde{\mathcal{M}}_{d,i-1}$ for all monomials $X^{(s,d-d_i)} \in k[S_M^{(h)}]_{d-d_i}$ that are not in $\langle \mathrm{LM}(\mathcal{G}_{i-1})\rangle$;
    Compute the row echelon form $\widetilde{\mathcal{M}}_{d,i}$ of $\mathcal{M}_{d,i}$;
    Add to $\mathcal{G}_i$ all rows of $\widetilde{\mathcal{M}}_{d,i}$ not top reducible by $\mathcal{G}_i$;
return $\mathcal{G}_m$

A direct consequence of this lemma is:

Corollary 4.6. *Algorithm 1 is correct.*

Proof. With the notations of Algo. 1 A direct induction on d and i with Lemma 4.5 shows that the row span of $\mathcal{M}_{d,i}$ is equal to the row span of the Macaulay matrix in degree d of $(f_1, \ldots, f_i)$. DefProp. 4.4 concludes the proof. □

In practice, the choice of the parameter D in Algorithm 1 is driven by the explicit bounds on the witness degree that we shall derive in Section 5.

4.2 Sparse-FGLM algorithm

The `FGLM` algorithm and its variants might be seen as a tool to change the representation of a 0-dimensional ideal. It relies on the notion of *normal form* relative to an ideal I. A normal form relative to I is a k-linear map $\mathrm{NF} : k[S] \to k[S]$ whose kernel is $\ker(\mathrm{NF}) = I$. It sends every coset of I to the same representative, allowing effective computations in the ring $k[S]/I$. One important feature of a sparse Gröbner basis is that it provides a normal form and an algorithm to compute it by successive reductions of leading monomials.

Let $(p_1, \ldots, p_r)$ be the Hilbert basis of a semigroup $S \subset \mathbb{Z}^n$. Given new indeterminates $H = \{H_1, \ldots, H_r\}$, any monomial in $k[S]$ is the image of a monomial in $k[H]$ via the morphism $\varphi : k[H_1, \ldots, H_r] \to k[S]$ defined by $\varphi(H_i) = X^{p_i}$. Given an admissible monomial ordering $\prec_H$ on the ring $k[H_1, \ldots, H_r]$, an ideal $I \subset k[S]$ and a normal form relative to I (given for instance by a sparse Gröbner basis of I), Algorithm 2 computes a Gröbner basis of $\varphi^{-1}(I)$. Note that $\psi\left(Var(I) \cap (\overline{k}^*)^n\right) = Var\left(\varphi^{-1}(I)\right) \cap (\overline{k}^*)^r$, where $\psi : \overline{k}^n \to \overline{k}^r$ is the map $\mathbf{x} \mapsto (\mathbf{x}^{p_1}, \ldots, \mathbf{x}^{p_r})$. Also, we would like to point out that Algorithm 2 does not depend on the support of the input sparse system, but only on the ambient semigroup S_M.

The main principle of Algorithm 2 is similar to the original `FGLM` Algorithm [16]: we consider the monomials in $k[H_1, \ldots, H_r]$ in increasing order until we obtain sufficiently many linear relations between their normal forms. The only difference is that the computations of the normal forms are performed in $k[S]$ (using a previously computed sparse Gröbner basis) via the morphism φ. For solving sparse systems, we choose the *lexicographical ordering* for $\prec_H$.

Theorem 4.7. *Algorithm* `Sparse-FGLM` *is correct: it computes the reduced GB of the ideal* $\varphi^{-1}(I) \subset k[H_1, \ldots, H_r]$ *with respect to* $\prec_H$.

Proof. Let $G = (g_1, \ldots, g_\mu)$ be the output of Algo. 2. Set $m_i = \mathrm{LM}(g_i)$. First, we prove that $G \subset \varphi^{-1}(I)$. Notice that each g_i is of the form $m_i - q$, where $\varphi(q) = \mathrm{NF}(\varphi(m_i))$. Consequently, $\mathrm{NF}(\varphi(g_i)) = 0$ and hence $g_i \in \varphi^{-1}(I)$. Next, let $h \in k[H]$ be a polynomial such that $\mathrm{LM}(h) \notin \langle \mathrm{LM}(G) \rangle$. Up to reducing its nonleading monomials by G, we can assume w.l.o.g. that all its monomials do not belong to $\langle \mathrm{LM}(G) \rangle$. Therefore, the normal forms of the images by φ of all the monomials in the support of h are linearly independent in $k[S]/I$ (otherwise the linear relation would have been detected by Algo. 2), which means that $\mathrm{NF}(\varphi(h)) \neq 0$ and hence $h \notin \varphi^{-1}(I)$, which proves that G is a Gröbner basis of $\varphi^{-1}(I)$. The proof that G is reduced is similar. □

Algorithm 2: `Sparse-FGLM`

Input : -a normal form NF: $k[S] \to k[S]$ of a 0-dim ideal I
-a monomial ordering $\prec_H$ on $k[H_1, \ldots, H_r]$
-a monomial map $\varphi : k[H_1, \ldots, H_r] \to k[S]$

Output: A Gröbner basis in $k[H_1, \ldots, H_r]$ w.r.t. $\prec_H$

$L := [1]$; `//list of monomials in` $k[H_1, \ldots, H_r]$
$E := [\,]$; `//staircase for the new ordering` $\prec_H$
$V := [\,]$; `//`$V = \mathrm{NF}(\varphi(S))$
$G := [\,]$; `//The Gröbner basis in` $k[H_1, \ldots, H_r]$
while $L \neq [\,]$ **do**
  $m := L[1]$; and Remove m from L;
  $v := \mathrm{NF}(\varphi(m))$; (1)
  $e := \#E$;
  if $v \in \mathrm{Span}_k(V)$ **then**
    $\exists\,(\lambda_i) \in k^e$ such that $v = \sum_{i=1}^{e} \lambda_i \cdot V_i$; (2)
    $G := G \cup \left[m - \sum_{i=1}^{s} \lambda_i \cdot E_i \right]$;
    Remove from L the elements top-reducible by G.
  else
    $E := E \cup [m]$; $V := \mathrm{RowEchelon}(V \cup [v])$; (3)
    $L := \mathrm{Sort}(L \cup [H_i\, m \mid i = 1, \ldots, r], \prec_H)$;
    Remove from L duplicate elements;
Return G;

5 Complexity

This section is devoted to the complexity of Algorithms 1 and 2 when the input is a homogeneous regular sequence. In the case of polytopal algebras, the complexity bounds of Theorems 5.3 and 5.4 depend mainly on the intrinsic combinatorial properties of the defining polytope.

Complexity model. All the complexity bounds count the number of arithmetic operations $\{+, \times, -, \div\}$ in k; each of them is counted with unit cost. It is not our goal to take into account operations in the semigroup S.

The first goal is to bound $\mathrm{d}_{\mathrm{wit}}$ (see DefProp. 4.4) via the Hilbert series of $k[S]/I$. For regular sequences, this Hilbert series can be computed by the following classical formula:

Proposition 5.1. *Let $\mathscr{P}$ be a normal lattice polytope, $f_1, \ldots, f_p \in k[\mathscr{P}]$ be a homogeneous regular sequence of homogeneous polynomials of respective degrees $(d_1, \ldots, d_p)$ and $I = \langle f_1, \ldots, f_p \rangle \subset k[\mathscr{P}]$. Then*

$$\mathrm{HS}_{k[\mathscr{P}]/I}(t) = \mathrm{HS}_{\mathscr{P}}(t) \cdot \prod_{i=1}^{p} (1 - t^{d_i}).$$

Proof. See *e.g.* [12, Exercise 21.17b]. □

The next lemma gives an explicit bound for the witness degree of regular sequences in a normal polytopal algebra:

Lemma 5.2. *Let $\mathscr{P} \subset \mathbb{R}^n$ be a normal lattice polytope and $f_1, \ldots, f_n$ be a homogeneous regular sequence in $k[\mathscr{P}]$ of degrees $(d_1, \ldots, d_n)$. Then any $\left[\mathrm{reg}(k[\mathscr{P}]) + 1 + \sum_{j=1}^{n}(d_j - 1)\right]$-sGB of the ideal $I = \langle f_1, \ldots, f_n \rangle$ is a sGB of I. In other words $\mathrm{d}_{\mathrm{wit}} \leq \mathrm{reg}(k[\mathscr{P}]) + 1 + \sum_{j=1}^{n}(d_j - 1)$.*

Proof. By Prop. 5.1 and with the notations of Prop. 2.6, the Hilbert series of $k[\mathscr{P}]/I$ is equal to

$$\begin{aligned} \mathrm{HS}_{\mathscr{P}}(t) \textstyle\prod_{i=1}^{n}(1 - t^{d_i}) &= \dfrac{Q(t) \prod_{i=1}^{n}(1 - t^{d_i})}{(1-t)^{n+1}} \\ &= \dfrac{Q(1) \prod_{i=1}^{n} d_i}{1 - t} + K(t) \end{aligned}$$

where $K(t) \in \mathbb{Z}[t]$ is a univariate polynomial with $\deg(K(t)) = \mathrm{reg}(k[\mathscr{P}]) - 1 + \sum_{i=1}^{p}(d_i - 1)$. Now, notice that the Hilbert series of $k[\mathscr{P}]/I$ is equal to that of $k[\mathscr{P}]/\mathrm{LM}(I)$. Therefore $\mathrm{HP}_{k[\mathscr{P}]/\mathrm{LM}(I)}(d)$ is constant for $d \geq \deg(K(t)) + 1$. Since $\ell < \ell'$ implies $\ell\mathscr{P} \subset \ell'\mathscr{P}$, we obtain

$$\begin{aligned} \max\{d \in \mathbb{N} \mid \exists X^{(s,d)} \notin \mathrm{LM}(I) \text{ s.t. } & s \in (d \cdot \mathscr{P}) \cap \mathbb{Z}^n \text{ and} \\ & s \notin ((d-1) \cdot \mathscr{P}) \cap \mathbb{Z}^n\} \\ = \deg(K(t)) + 1. \end{aligned}$$

Consequently, minimal generators of $\mathrm{LM}(I)$ and hence minimal homogeneous Gröbner bases of I have degree at most $\deg(K(t)) + 2 = \mathrm{reg}(k[\mathscr{P}]) + 1 + \sum_{j=1}^{n}(d_j - 1)$. □

Now that we have an upper bound for the witness degree, we can estimate the cost of computing a sGB by reducing the Macaulay matrix in degree $\mathrm{d}_{\mathrm{wit}}$ (although `sparse-MatrixF5` is a much faster way to compute a sGB in practice, it is not easy to bound precisely its complexity). Note that $\mathrm{reg}(k[\mathscr{P}])$ in the following theorem can be deduced from Prop. 2.7.

Theorem 5.3. *With the same notations as in Lemma 5.2, the complexity of computing a sGB of $\chi_{\mathscr{P}\cap\mathbb{Z}^n}(\langle f_1,\dots f_n\rangle) \subset k[S_{\mathscr{P}\cap\mathbb{Z}^n}]$ by reducing the Macaulay matrix in degree $\mathrm{d}_{\mathrm{wit}}$ is bounded above by*

$$O\left(n\,\mathrm{HP}_{\mathscr{P}}(\mathrm{d}_{\mathrm{wit}})^{\omega}\right),$$

where $\mathrm{d}_{\mathrm{wit}} \leq \mathrm{reg}(k[\mathscr{P}]) + 1 + \sum_{j=1}^{n}(d_j - 1)$ and ω is a feasible exponent for the matrix multiplication ($\omega < 2.373$ with [33]).

Proof. Let $I \subset k[\mathscr{P}]$ be the ideal generated by $(f_1,\dots,f_n)$. The number of columns and rows of the Macaulay matrix in degree d are respectively

$$\begin{array}{rcl} \mathrm{nb}_{\mathrm{cols}} & = & \mathrm{HP}_{\mathscr{P}}(d), \\ \mathrm{nb}_{\mathrm{rows}} & = & \sum_{i=1}^{n}\mathrm{HP}_{\mathscr{P}}(d-\deg(f_i)) \leq n\,\mathrm{HP}_{\mathscr{P}}(d). \end{array}$$

Consequently, the row echelon form of such a matrix can be computed within $O(n\,\mathrm{HP}_{\mathscr{P}}(d)^{\omega})$ field operations [30, Prop. 2.11]. By Proposition 3.5 and Lemma 5.2, for $d = \mathrm{d}_{\mathrm{wit}} \leq \mathrm{reg}(k[\mathscr{P}]) + 1 + \sum_{j=1}^{n}(d_j - 1)$, this provides a sGB of $\chi_{\mathscr{P}\cap\mathbb{Z}^n}(I)$. □

We now investigate the complexity of Algorithm 2 when $I \subset k[S]$ is a 0-dim. ideal, and use the same notations as in Section 4.2. Notice that the map φ induces an isomorphism $\psi : k[H]/\varphi^{-1}(I) \to k[S]/I$ and therefore Algorithm 2 may be seen as a way to change the representation of $k[S]/I$.

Theorem 5.4. *Set $\delta = \dim_k(k[S]/I)$ and let r be the cardinality of the Hilbert basis of S. If the input normal form is computed via a reduced sGB of $I \subset k[S]$ (for some monomial ordering), S is a simplicial affine semigroup (see Def. 2.3) and $k[S]$ is Cohen-Macaulay, then Algorithm 2 computes the Gröbner basis G with at most $O(r\cdot\delta^3)$ operations in k.*

Proof. Once the r matrices of size $\delta\times\delta$ representing the multiplications by p_i in the canonical monomial basis of $k[S]/I$ are known, Step **(1)** in Algorithm 2 can be achieved in $O(\delta^2)$ as in the classical `FGLM` Algorithm [16]. Steps **(2)** and **(3)** are done by linear algebra as in [16], which leads to a total complexity of $O(r\cdot\delta^3)$ since the same analysis holds. It remains to prove that the multiplication matrices can be constructed in $O(r\cdot\delta^3)$ operations (this is a consequence of [16, Prop. 2.1] in the classical case). Since $k[S]$ is Cohen-Macaulay and S is simplicial, we obtain by [28, Thm. 1.1] that for any two distinct $p_i, p_j \in \mathrm{Hilb}(S)$ and for any $s \in S$, if $s - p_i$ and $s - p_j$ are in S then $s - p_i - p_j \in S$. With this extra property, the proof of [16, Prop. 2.1] extends to semigroup algebras. □

If the input system is a regular sequence of Laurent polynomials, then δ can be bounded by the mixed volume of their Newton polytopes by Kushnirenko-Bernstein's Theorem [3].

6 Dense, multi-homogeneous and overdetermined systems

In this section, we specialize Theorems 5.3 and 5.4 to several semigroups to obtain new results on the complexity of solving inhomogeneous systems with classical GB algorithms ($\mathscr{P}$ is the standard simplex), multi-homogeneous systems ($\mathscr{P}$ is a product of simplices) and we state a variant of Fröberg's conjecture for overdetermined sparse systems.

Inhomogeneous dense systems. If $\mathscr{P} = \Delta_n$ is the standard simplex in $\mathbb{R}^n$, then computations of a sparse Gröbner basis in the cone over Δ_n correspond to classical Gröbner bases computations using the so-called "sugar strategy" introduced in [21]. Applying directly Theorems 5.3 and 5.4 with $\mathscr{P} = \Delta_n$ gives

Corollary 6.1. *Let $f_1, \ldots, f_n$ be a regular sequence of inhomogeneous polynomials of respective degrees $(d_1, \ldots, d_n)$ in $k[X_1, \ldots, X_n]$. Then the complexity of computing a classical Gröbner basis of $\langle f_1, \ldots, f_n \rangle$ with respect to a graded monomial ordering is bounded by*

$$O\left(n\binom{n+\mathrm{d_{wit}}}{n}^{\omega}\right),$$

where $\mathrm{d_{wit}} \leq 1 + \sum_{i=1}^{n}(d_i - 1)$.

This statement was already known under the assumption that the system of the homogeneous parts of highest degree $f_1^\infty, \ldots, f_n^\infty$ is also regular, see *e.g.* [1]. However, this condition is not verified for several systems appearing in applications. Up to our knowlegde, this is the first time that such complexity results are obtained for inhomogenous systems without any assumption on $f_1^\infty, \ldots, f_n^\infty$.

Multi-homogeneous systems. Another class of polynomials appearing frequently in applications are *multi-homogeneous systems*. A polynomial of multi-degree $(d_1, \ldots, d_\ell)$ w.r.t. a partition of the variables in blocks of sizes $(n_1, \ldots, n_\ell)$ is a polynomial whose Newton polytope is included in $d_1\Delta_{n_1} \times \cdots \times d_\ell\Delta_{n_\ell}$. In that case, the associated polytope is a product of simplices, which allows us to state the following complexity theorem:

Theorem 6.2. *Let $f_1, \ldots, f_n$ be a regular sequence of polynomials of multi-degree $(d_1, \ldots, d_\ell)$ w.r.t. a partition of the variables in blocks of sizes $(n_1, \ldots, n_\ell)$ (with $n_1 + \cdots + n_\ell = n$). Then the combined complexity of Steps (1) to (4) of the solving process in Section 3 is bounded by*

$$\begin{array}{rl} & O\left(n\,\mathrm{HP}_{\mathscr{P}}(\mathrm{d_{wit}})^{\omega} + n\,\mathrm{vol}(\mathscr{P})^3\right), \\ \textit{where} & \mathscr{P} = d_1\Delta_{n_1} \times \cdots \times d_\ell\Delta_{n_\ell}, \\ & \mathrm{d_{wit}} \leq n + 2 - \max_{i\in\{1,\ldots,\ell\}}(\lceil (n_i+1)/d_i \rceil), \\ & \mathrm{HP}_{\mathscr{P}}(\mathrm{d_{wit}}) = \binom{n_1+\mathrm{d_{wit}}\cdot d_1}{n_1} \cdots \binom{n_\ell+\mathrm{d_{wit}}\cdot d_\ell}{n_\ell}, \\ \textit{and} & \mathrm{vol}(\mathscr{P}) = \binom{n}{n_1,\ldots,n_\ell} \prod_{i=1}^{\ell} d_i^{n_i}. \end{array}$$

Proof. Applying Theorems 5.3 and 5.4 with $\mathscr{P}$ equal to $d_1\Delta_{n_1} \times \cdots \times d_\ell\Delta_{n_\ell}$ yields the complexity bound in terms of $\mathrm{d_{wit}}$, $\#\,\mathrm{Hilb}(S_{\mathscr{P}\cap\mathbb{Z}^n})$ and δ. First, notice that the semigroup

generated by $\mathscr{P} \cap \mathbb{Z}^n$ is $\mathbb{N}^n$, and hence $\#\operatorname{Hilb}(S_{\mathscr{P}\cap\mathbb{Z}^n}) = n$. Next, $\beta(d_1\Delta_{n_1} \times \cdots \times d_\ell\Delta_{n_\ell})$ has an interior lattice point if and only if for all i, $\beta d_i\Delta_{n_i}$ has an interior lattice point, *i.e.* $\beta d_i > n_i$. The smallest β that verifies this condition is $\max(\lceil (n_1+1)/d_1\rceil, \ldots, \lceil (n_\ell+1)/d_\ell\rceil)$. By Prop. 2.7, $\operatorname{reg}(k[\mathscr{P}]) = n+1-\max(\lceil (n_1+1)/d_1\rceil, \ldots, \lceil (n_\ell+1)/d_\ell\rceil)$. Since the $f_1, \ldots, f_n$ have degree 1 in $k[\mathscr{P}]$, we get $\mathrm{d_{wit}} \leq \operatorname{reg}(k[\mathscr{P}]) + 1$. Finally, notice that the unnormalized volume of $d\Delta_q \in \mathbb{R}^q$ is $d^q/q!$. Consequently, the unnormalized volume of $\mathscr{P}$ is $\prod_{i=1}^{\ell} d_i^{n_i}/n_i!$. Normalizing the volume amounts to multiplying this value by $n!$, which yields the formula for $\operatorname{vol}(\mathscr{P})$ and equals the multi-homogeneous Bézout number. The number of solutions (counted with multiplicity) is classically bounded by this value and hence $\delta \leq \operatorname{vol}(\mathscr{P})$. □

Finally, we state a variant of Fröberg's conjecture [19] in the sparse framework, leading to a notion of "sparse semi-regularity". It provides a bound on the witness degree of generic overdetermined sparse systems: this conjecture can be used to adjust the parameter D of Algorithm 1.

Conjecture 6.3. *Let $\mathscr{P} \subset \mathbb{R}^n$ be a normal lattice polytope, $(d_1, \ldots, d_m) \in \mathbb{N}^m$ be a sequence of integers with $m > n$. If $f_1, \ldots, f_m \in \mathbb{C}[\mathscr{P}]$ are generic homogeneous polynomials of respective degrees $(d_1, \ldots, d_m)$, then*

$$\mathrm{HS}_{\mathbb{C}[\mathscr{P}]/\langle f_1,\ldots,f_m\rangle}(t) = \left[\mathrm{HS}_{\mathscr{P}}(t)\prod_{i=1}^{m}(1-t^{d_i})\right]_+,$$

where $[\ ]_+$ means truncating the series expansion at its first nonpositive coefficient. Systems for which this equality holds are called semi-regular*. The witness degree of a semi-regular sequence is bounded above by the index of the first zero coefficient in the series expansion of $\mathrm{HS}_{\mathbb{C}[\mathscr{P}]/\langle f_1,\ldots,f_m\rangle}(t)$.*

7 Experimental results

In this section, we estimate the speed-up that one can expect for solving sparse systems or systems of Laurent polynomials via sparse Gröbner bases computations, compared to classical Gröbner bases algorithms. The same linear algebra routines are used in the compared implementations. Consequently, the speed-up reflects the differences between the characteristics (size, sparseness,...) of the matrices that have to be reduced.

Workstation. All experiments have been conducted on a 2.6GHz `IntelCore i7`.

We compare `sparse-MatrixF5` (abbreviated `sp-MatrixF5`) with the implementation of the F_5 algorithm in the FGb library. We report more detailed experimental results on a benchmarks' webpage[1]. In all these experiments, the base field k is the finite field GF(65521). All tests are done with overdetermined systems with one rational solution in $\mathrm{GF}(65521)^n$. The goal is to recover this solution. In that case, the `FGLM` algorithm is not necessary since the sparse Gröbner basis describes explicitly the image of the solution by a monomial map. In several settings, we report the speed-up obtained with our prototype implementation.

[1] `http://www-polsys.lip6.fr/~jcf/Software/benchssparse.html`

(n_x, n_y, m)	sp-MatrixF5	FGb-F5	Speed-up
(2,29,40)	0.12s	5.2s	43
(2,39,53)	0.49s	36.7s	74
(2,49,65)	1.53s	298.5s	195
(2,59,78)	4.63s	852.3s	184
(6,19,52)	1.10s	25.2s	22
(6,21,56)	2.13s	51.5s	24
(6,27,71)	7.07s	236.0s	33

Table 1: Overdetermined bilinear systems in (n_x, n_y) variables and m equations

(n_x, n_y, m)	sp-MatrixF5	FGb-F5	Speed-up
(1,34,36)	0.2s	395.1s	1975
(1,39,41)	0.45s	1641s	3646
(1,44,46)	0.75s	3168.8s	4225
(2,15,25)	0.09s	410.1s	4556
(2,17,27)	0.15s	1894.7s	12631
(2,19,30)	0.4s	5866.1s	14665
(3,10,24)	0.15s	2937.7s	19584
(10,4,50)	23.1s	1687.3s	73
(11,5,66)	155.1s	6265.8s	40
(12,6,86)	872.2s	27093.3s	31

Table 2: Systems in (n_x, n_y) variables of bidegree $(2, 1)$ and m equations

Bilinear systems. In Table 1, we focus on overdetermined bilinear systems. For $(n_x, n_y, m) \in \mathbb{N}^3$, we generate a system of m polynomials with support $\Delta_{n_x} \times \Delta_{n_y}$ uniformly at random in the set of such systems which have at least one solution in $\mathrm{GF}(65521)^{n_x+n_y}$.

Systems of bidegree $(2, 1)$**.** In Table 2, we report the performances on overdetermined systems with support $2\Delta_{n_x} \times \Delta_{n_y}$. Note that we obtain important speed-ups when $n_x < n_y$ (more than 19000 for $(n_x, n_y, m) = (3, 10, 24)$).

Fewnomial systems. In Table 3, we report performances on fewnomial systems. The complexity analysis in Section 5 do not apply to this context because the semigroup algebra in which we compute is not normal. However, the correctness of the algorithms still holds. The systems are generated as follows: for $(n, t, m) \in \mathbb{N}^3$ we pick t monomials of degree 2 in n variables uniformly at random and we generate a system of m polynomials with this support in $\mathrm{GF}(65521)[X_1, \ldots, X_n]$ with random coefficients such that there is at least one solution in $\mathrm{GF}(65521)^n$. The computations are done w.r.t. the semigroup generated by the t monomials. Note that for some specific instances, the speed-up factor can be as high as 16800.

(n, t, m)	`sp-MatrixF5`	`FGb-F5`	Speed-up
(80,240,221)	0.10s	54.5s	545
(80, 240, 223)	0.08s	16.3s	203
(150, 450, 434)	0.24s	161.2s	671
(300, 900, 881)	4.56s	11301.0s	2478
(120, 240, 233)	0.01s	16.8s	16800
(40, 160, 128)	0.21s	5.93s	28
(60, 240, 211)	0.55s	29.04s	52

Table 3: Fewnomials systems

References

[1] G. Ars, J.-C. Faugère, H. Imai, M. Kawazoe, and M. Sugita. Comparison between XL and Gröbner basis algorithms. In *ASIACRYPT 2004*, LNCS, pages 157–167, 2004.

[2] M. Bardet, J.-C. Faugère, and B. Salvy. On the complexity of the F5 Gröbner basis algorithm. *arXiv*, 1312.1655, 2013.

[3] D. Bernstein. The number of roots of a system of equations. *Funct. Anal. and its Appli.*, 9(3):183–185, 1975.

[4] M. Brickenstein. Slimgb: Gröbner bases with slim polynomials. *Revista Matemática Complutense*, 23(2):453–466, 2010.

[5] M. P. Brodmann and R. Y. Sharp. *Local cohomology: an algebraic introduction with geometric applications*. Cambridge University Press, 1998.

[6] W. Bruns, J. Gubeladze, and N. V. Trung. Normal polytopes, triangulations, and Koszul algebras. *J. fur die reine und angewandte Mathematik*, 485:123–160, 1997.

[7] J. F. Canny and I. Z. Emiris. An efficient algorithm for the sparse mixed resultant. In *Applied Algebra, Algebraic Algo. and Error-correcting Codes*, pages 89–104. Springer, 1993.

[8] J. F. Canny and I. Z. Emiris. A subdivision-based algorithm for the sparse resultant. *J. ACM*, 47:417–451, 1999.

[9] D. A. Cox, J. B. Little, and H. K. Schenck. *Toric varieties*. AMS, 2011.

[10] C. Eder and J.-C. Faugère. A survey on signature-based Gröbner basis computations. *arXiv*, 1404.1774, 2014.

[11] E. Ehrhart. Sur les polyèdres rationnels homothétiques à n dimensions. *CR Acad. Sci. Paris*, 254:616–618, 1962.

[12] D. Eisenbud. *Commutative Algebra: with a view toward algebraic geometry*, volume 150. Springer, 1995.

[13] I. Z. Emiris. Toric resultants and applications to geometric modelling. In *Solving polynomial equations*, pages 269–300. Springer, 2005.

[14] I. Z. Emiris and V. Y. Pan. Symbolic and numeric methods for exploiting structure in constructing resultant matrices. *J. of Symbolic Computation*, 33(4):393–413, 2002.

[15] J.-C. Faugère. A new efficient algorithm for computing Gröbner bases without reduction to zero (F5). In *ISSAC 2002*, pages 75–83. ACM, 2002.

[16] J.-C. Faugère, P. Gianni, D. Lazard, and T. Mora. Efficient Computation of Zero-dimensional Gröbner Bases by Change of Ordering. *J. of Symb. Computation*, 16(4):329–344, 1993.

[17] J.-C. Faugère and S. Rahmany. Solving systems of polynomial equations with symmetries using SAGBI-Gröbner bases. In *ISSAC '09*, pages 151–158. ACM, 2009.

[18] J.-C. Faugère, M. Safey El Din, and P.-J. Spaenlehauer. Gröbner bases of bihomogeneous ideals generated by polynomials of bidegree (1, 1): Algorithms and complexity. *J. of Symbolic Computation*, 46(4):406–437, 2011.

[19] R. Fröberg. An inequality for Hilbert series of graded algebras. *Mathematica Scandinavica*, 56:117–144, 1985.

[20] W. Fulton. *Introduction to Toric Varieties*. Princeton University Press, 1993.

[21] A. Giovini, T. Mora, G. Niesi, L. Robbiano, and C. Traverso. "One sugar cube, please" or selection strategies in the Buchberger algorithm. In *ISSAC '91*, pages 49–54. ACM, 1991.

[22] M. Hochster. Rings of invariants of tori, Cohen-Macaulay rings generated by monomials, and polytopes. *The Annals of Mathematics*, 96(2):318–337, 1972.

[23] D. Lazard. Gröbner bases, Gaussian elimination and resolution of systems of algebraic equations. In *Computer algebra*, pages 146–156. Springer, 1983.

[24] I. MacDonald. Polynomials associated with finite cell-complexes. *J. London Math. Soc.*, 4:181–192, 1971.

[25] E. Miller and B. Sturmfels. *Combinatorial commutative algebra*, volume 227. Springer Verlag, 2005.

[26] T. Oda. *Convex bodies and algebraic geometry*. Springer, 1988.

[27] P. Pedersen and B. Sturmfels. Mixed monomial bases. In *Algorithms in algebraic geometry and applications*, pages 307–316. Springer, 1995.

[28] J. Rosales and P. A. Garcia-Sanchez. On Cohen-Macaulay and Gorenstein simplicial affine semigroups. *Proceedings of the Edinburgh Mathematical Society*, 41(3):517–538, 1998.

[29] R. P. Stanley. Decompositions of rational convex polytopes. *Ann. Discrete Math. v6*, pages 333–342, 1980.

[30] A. Storjohann. Algorithms for matrix canonical forms. *Ph.D. thesis*, 2000.

[31] B. Sturmfels. Sparse elimination theory. In *Proc. Comp. Algebraic Geom. and Commut. Algebra*, pages 377–396. Cambridge Univ. Press, 1991.

[32] B. Sturmfels. *Gröbner bases and convex polytopes*, volume 8. AMS, 1996.

[33] V. V. Williams. Multiplying matrices faster than Coppersmith-Winograd. In *Proc. of STOC'12*, pages 887–898. ACM, 2012.