Efficiently and Effectively Recognizing Toricity of Steady State Varieties

Dima Grigoriev, Alexandru Iosif, Hamid Rahkooy, Thomas Sturm and Andreas Weber

Abstract. We consider the problem of testing whether the points in a complex or real variety with non-zero coordinates form a multiplicative group or, more generally, a coset of a multiplicative group. For the coset case, we study the notion of shifted toric varieties which generalizes the notion of toric varieties. This requires a geometric view on the varieties rather than an algebraic view on the ideals. We present algorithms and computations on 129 models from the BioModels repository testing for group and coset structures over both the complex numbers and the real numbers. Our methods over the complex numbers are based on Gröbner basis techniques and binomiality tests. Over the real numbers we use first-order characterizations and employ real quantifier elimination. In combination with suitable prime decompositions and restrictions to subspaces it turns out that almost all models show coset structure. Beyond our practical computations, we give upper bounds on the asymptotic worst-case complexity of the corresponding problems by proposing single exponential algorithms that test complex or real varieties for toricity or shifted toricity. In the positive case, these algorithms produce generating binomials. In addition, we propose an asymptotically fast algorithm for testing membership in a binomial variety over the algebraic closure of the rational numbers.

Mathematics Subject Classification (2010). Primary 14Q20; Secondary 14P05, 92C42.

Keywords. Binomial ideals, chemical reaction networks, complexity bounds, complex balancing, Gröbner bases, mass-action kinetics, real quantifier elimination, scientific computing, symbolic computation, toric varieties.

1. Introduction

We are interested in situations where the points with non-zero coordinates in a given complex or real variety form a multiplicative group or, more generally, a coset. We illustrate this by means of a simple example. For $K \in \{\mathbb{C}, \mathbb{R}\}$, $(K^*)^n$ denotes the direct power of the multiplicative group of the respective field. Consider a family of ideals

$$I_k = \langle x^2 - ky^2 \rangle \tag{1}$$

with a rational parameter k. Let V_k be the complex variety of I_k, and let $V_k^* = V_k \cap (\mathbb{C}^*)^2$. Then V_1^* forms a *group*, but V_{-1}^* does not, because it does not contain $(1, 1)$. However, $V_{-1}^* = (1, i) \cdot V_1^*$ forms a *coset* of V_1^*. Over the reals, $V_1^* \cap (\mathbb{R}^*)^2$ is again a group, but $V_{-1}^* \cap (\mathbb{R}^*)^2 = \varnothing$ is not a coset of any group. Consider now $V_1 = V_{11} \cup V_{12}$, where V_{11} and V_{12} are given by $\langle x - y \rangle$ and $\langle x + y \rangle$, respectively. Notice that V_{11}^* is itself a group, and V_{12}^* is a coset of V_{11}^*. Both V_{11}^*

This work was supported by the bilateral project ANR-17-CE40-0036 and DFG-391322026 SYMBIONT..

and V_{12}^* are irreducible because their generating ideals are prime. Under the additional condition of irreducibility V_{11}^* forms a *torus* and V_{12}^* forms a *shifted torus*. We then call the varieties V_{11} and V_{12} *toric* and *shifted toric*, respectively.

Toric varieties are well established and have an important role in algebraic geometry [31, 23]. However, our principal motivation to study generalizations of toricity comes from the sciences, specifically *chemical reaction networks* such as the following:

$$2\,\mathrm{A} \underset{\frac{1}{2}}{\overset{\frac{k}{2}}{\rightleftharpoons}} 2\,\mathrm{B}.$$

Assuming, e.g., mass action kinetics one can derive a system of autonomous ordinary differential equations describing the development of concentrations of the species A and B as functions of time [28, Section 2.1.2]. For the given reaction network one obtains a polynomial vector field generating exactly our ideals I_k in (1). Our methods thus detect whether equilibrium points with non-zero coordinates form a group or a coset.

Detecting toricity of a variety in general, and of the steady state varieties of chemical reaction networks in particular, is a difficult problem. The first issue in this regard is finding suitable notions to describe the structure of the steady states. Existing work typically addresses algebraic properties of the steady state ideal, e.g., the existence of binomial Gröbner bases. In this article, in contrast, we take a geometric approach, focusing on varieties rather than ideals. We propose to study toricity and shifted toricity of varieties V over $K \in \{\mathbb{C}, \mathbb{R}\}$, which for irreducible varieties coincides with $V \cap (K^*)^n$ forming a multiplicative group or coset, respectively. It is noteworthy that chemical reaction network theory generally takes place in the interior of the first orthant of $\mathbb{R}^n$, i.e., all species concentrations and reaction rates are assumed to be strictly positive [28]. Our considering $(\mathbb{C}^*)^n$ in contrast to $\mathbb{C}^n$ is a first step in this direction, considering also $(\mathbb{R}^*)^n$ is another step.

Toric dynamical systems have been studied by Feinberg [26] and by Horn and Jackson [43]. Craciun et al. [16] showed that toric dynamical systems correspond to *complex balancing* [28]. Our generalized notions of toricity are inspired by Grigoriev and Milman's work on *binomial varieties* [39]. There are further definitions in the literature, where the use of the term "toric" is well motivated. Gatermann et al. considered *deformed toricity* for steady state ideals [32]. The exact relation between the principle of complex balancing and various definitions of toricity has obtained considerable attention in the last years [54, 33, 50].

Complex balancing itself generalizes *detailed balancing*, which has widely been used in the context of chemical reaction networks [27, 28, 43]. Gorban et al. [35, 34] related reversibility of chemical reactions in detailed balance to binomiality of the corresponding varieties. Historically, the principle of detailed balancing has attracted considerable attention in the sciences. It was used by Boltzmann in 1872 in order to prove his *H-theorem* [4], by Einstein in 1916 for his quantum theory of emission and absorption of radiation [22], and by Wegscheider [62] and Onsager [52] in the context of *chemical kinetics*, which lead to Onsager's Nobel prize in Chemistry in 1968. Pérez-Millán et al. [54] consider steady state ideals with binomial generators. They present a sufficient linear algebra condition on the *stoichiometry matrix* of a chemical reaction network in order to test whether the state ideal has binomial generators. Conradi and Kahle proposed a corresponding heuristic algorithm. They furthermore showed that the sufficient condition is even equivalent when the ideal is homogenous [14, 46, 45]. Based on the above-mentioned linear algebra condition, MESSI systems have been introduced in [53]. Recently, binomiality of steady states ideals was used to infer network structure of chemical reaction networks out of measurement data [61].

Besides its scientific adequacy as a generalization of complex balancing there are practical motivations for studying toricity. Relevant models are typically quite large. For instance, with our comprehensive computations in this article we will encounter systems up to 90 polynomials in dimension 71. This potentially takes symbolic computation to its limits. A possible approach to overcome this is to discover systematic occurrences of certain structural properties in the input, and to exploit those

structural properties towards more efficient algorithms. From this point of view, toricity and shifted toricity are interesting concepts because tools from toric geometry can be used as a complexity reduction step in the multistationarity problem. Grigoriev and Weber [40] gave a complexity analysis for solving binomial varieties, based on the computation of Smith normal forms. More interestingly, toric and shifted toric models are known to have scale invariant multistationarity in the space of linear conserved quantities, which further reduces the dimension of the multistationarity problem [13].

Our original contributions in this article are the following. Interested in the geometric structure of real and complex varieties V rather than the algebraic structure of the corresponding ideals, we study primarily $V^* = V \cap (K^*)^n$. We call V shifted toric when V is irreducible and V^* is a coset. This generalizes the notion of toric varieties V for groups V^*. Within this setting, we have two principal results:

- Relating our novel geometric view to the established algebraic view, we give a characterization in terms of Gröbner bases for V^* to be a group or coset. (Proposition 7).
- We show that Zariski closures of groups in $(\mathbb{C}^*)^n$ are binomial varieties (Proposition 6). The converse had been shown in [39].

We propose practical algorithms testing for given polynomial systems F whether their varieties contain group or coset structures.

- We consider over the complex numbers $V(F)^*$ (Algorithm 1) and $V(P_i)^*$ for prime components $\langle P_i \rangle$ of $\langle F \rangle$ (Algorithm 3).
- We consider the same over the real numbers, $V(F)^*$ (Algorithm 4) and $V(P_i)^*$ (Algorithm 6).
- With prime decomposition we find that for up to 98% of the prime components $V_K(P_i)^*$ is either empty or a coset.

Our algorithms are implemented in Maple[1] and Reduce[2] [57, 42, 41] and systematically applied to the steady state varieties of 129 models from the BioModels repository[3]. Our objective was to build on robust, off-the-shelf software, which has a chance to be accepted by scientific communities outside symbolic computation in the foreseeable future. As a consequence, our proposed algorithms must rely on existing implementations. Over $\mathbb{C}$ we use Gröbner bases [7, 24, 25], and over $\mathbb{R}$ we use real quantifier elimination techniques.

Gröbner bases and real quantifier elimination mentioned above come with high intrinsic complexity. The former are complete for exponential space [49]. The latter are double exponential [17, 37, 63]. From a more theoretical point of view we study the intrinsic complexity of the problems actually addressed. We follow Chistov–Grigoriev's complexity results for first-order quantifier elimination over algebraically closed fields [10], the algorithm constructing irreducible components of a variety [10, 36] and Grigoriev–Vorobjov's algorithm for solving polynomial system of inequalities [38]. These results are used to propose an algorithm to test within single exponential complexity bounds whether:

- a given complex variety is toric or shifted toric (Theorem 24);
- a given real variety is toric or shifted toric (Theorem 25);
- a given point belongs to a given binomial variety (Theorem 26).

The plan of the article is as follows. In Section 2 we present preliminaries from the literature and introduce our new notions and related results. In Section 3, we present new algorithms for group and coset tests over $\mathbb{C}$ and $\mathbb{R}$. As a first step towards irreducible varieties we also use prime decompositions over the coefficient field, i.e., rational numbers. For the sake of a concise discussion, the major part of our rather comprehensive computation results can be found in Appendix A. In

[1] Maple (2019). Maplesoft, a division of Waterloo Maple Inc., Waterloo, Ontario.
[2] `https://sourceforge.net/projects/reduce-algebra/`
[3] `https://www.ebi.ac.uk/biomodels/`

Section 4 we propose asymptotically fast algorithms for the practical computations in Section 3. The proposed algorithms induce upper complexity bounds on the corresponding problems. In Section 5 we summarize our findings and mention perspectives for future work.

2. Toric, Shifted Toric and Binomial Varieties

We use K to denote either $\mathbb{C}$ or $\mathbb{R}$ when definitions or results hold for both fields. The natural numbers $\mathbb{N}$ include 0. For positive $n \in \mathbb{N}$ and $X = (x_1, \ldots, x_n)$, the polynomial ring with coefficients in $\mathbb{Q}$ and variables $x_1, \ldots, x_n$ is written $\mathbb{Q}[X] = \mathbb{Q}[x_1, \ldots, x_n]$. For $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, $X^\alpha = x_1^{\alpha_1} \ldots x_n^{\alpha_n}$ is a monomial in $\mathbb{Q}[X]$. When mentioning Gröbner bases of ideals we always mean reduced Gröbner bases; when not mentioned explicitly the term order is not relevant. Given a polynomial $f \in \mathbb{Q}[X]$, the variety of a f over K is $V(f) = \{ z \in K \mid f(z) = 0 \} \subseteq K^n$; this naturally generalizes to sets F of polynomials and ideals I. Vice versa, given a variety $V \subseteq K^n$, we define the ideal of V to be $I(V) = \{ f \in \mathbb{Q}[X] \mid f(z) = 0 \text{ for all } z \in V \}$. Recall that over algebraically closed fields, $I(V(J)) = \sqrt{J}$, the radical of J.

Let K^* be the multiplicative group of K. A subgroup $G \subseteq (K^*)^n$ is called a *torus* over the direct product $(K^*)^n$, where multiplication is coordinate-wise, if there exists $m \in \mathbb{N}$ such that G is isomorphic to $(K^*)^m$. A variety $V \subseteq K^n$ is called *toric* if it is irreducible and there exists a torus $G \subseteq (K^*)^n$ such that $V = \overline{G}$, the Zariski closure of G [58]. It is noteworthy that there are alternative definitions that requires the variety to be normal as well [31].

For a variety $V \subseteq K^n$, by V^* we denote $V \cap (K^*)^n$, i.e., the points in V with non-zero coordinates. It is well-known that every torus is the (irreducible) zero set of a set of Laurent binomials of the form $X^\gamma - 1$ where $\gamma \in \mathbb{Z}$ [31]. We are going to make use of the following proposition, which is a consequence of results in [31, 58].

Proposition 1. *Let $V \subseteq \mathbb{C}^n$ be a variety. V is a toric variety if and only $I(V)$ is prime and the reduced Gröbner basis (with respect to any term order) of $I(V)$ contains only binomials of the form $X^\alpha - X^\beta$ where $\alpha, \beta \in \mathbb{N}^n$.*

By definition, V is toric if and only if V is irreducible and there exists a torus T such that $V = \overline{T}$. Assume that V is given by a set of generators of $I(V)$. Since V is irreducible, then V^* is irreducible, hence $T = V^*$, $V = \overline{V^*}$, and $I(V) = I(\overline{V^*})$. Therefore, it suffices to compute a Gröbner basis of $I(V)$ instead of $I(V^*)$ and use Proposition 1 in order to check if V^* is a group. Note that if a Gröbner basis of $I(V)$ is non-empty and its elements are of the form $X^\alpha - X^\beta$ and the ideal is prime, then V^* is clearly non-empty. For more detailed study of toric varieties refer to [31] and [58].

Not all subgroups of $(K^*)^n$ are reducible. For example if $K = \mathbb{C}$ and $V = V(x^2 - y^2) \subseteq \mathbb{C}^n$, one can check that $V^* = V \cap (\mathbb{C}^*)^2$ is a group, however it is not reducible, hence not a torus. Actually V can be decomposed into the torus $V_1 = V(x - y) \cap (\mathbb{C}^*)^2$ and $V_2 = V(x + y) \cap (\mathbb{C}^*)^2$, which is a coset of V_1. Varieties that admit a group structure have interesting properties. A class of such varieties, called binomial varieties are studied by Grigoriev and Milman [39], where the authors present a structure theorem for them and discuss the complexity of their Nash resolution. We remind ourselves of the definition and the structure theorem of binomial varieties, which will be used for classifying steady state ideals.

Definition 2 (Binomial Variety). A variety $V \subseteq \mathbb{C}^n$ is called binomial if $V^* := V \cap (\mathbb{C}^*)^n$ is the zero set of a finite set of binomials of the form $X^\alpha - 1 \in K[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ and $V = \overline{V^*}$.

Using Gröbner bases instead of Laurent polynomials, one can see that if a variety $V \subseteq \mathbb{C}^n$ is binomial then all elements of every Gröbner basis of $I(V)$ are binomials of the form $X^\alpha - X^\beta$, where $\alpha, \beta \in \mathbb{N} \setminus \{0\}$. The following theorem by Grigoriev and Milman shows the structure of the binomial varieties by precisely describing their irreducible components.

Theorem 3. [39, Theorem 3.7] *The irreducible components of a binomial variety $V \subseteq \mathbb{C}^n$ include exactly one toric variety $\overline{T}$, where $T \subseteq (\mathbb{C}^*)^n$ is a torus, and several varieties $V_1 = \overline{x_1 T}, \ldots, V_r = \overline{x_r T}$, where $x_1 T, \ldots, x_r T \subseteq (\mathbb{C}^*)^n$ are cosets of T as a group, with respect to $x_1, \ldots, x_r \in (\mathbb{C}^*)^n$, respectively.*

Later in this section we will show that the closure of any subgroup of $(\mathbb{C}^*)^n$ is a binomial variety. Proposition 1 gives the form of the polynomials in a Gröbner basis of the ideal describing the toric component of a binomial variety. For the components that are cosets of the torus, one can easily derive the form of the reduced Gröbner basis from the definition of the torus and Proposition 1. This is stated precisely in the following proposition. Intuitively, non-toric components of a binomial variety can be considered as the *shifts* of the toric component. This motivates us to define *shifted toric varieties*.

Definition 4 (Shifted Toric Variety). A shifted torus in $(K^*)^n$ is defined to be a coset of a torus in $(K^*)^n$. A variety $V \subseteq K^n$ is called shifted toric if it is the closure of a shifted torus.

Since every group is a coset of itself, every torus is a shifted torus, and therefore every toric variety is a shifted toric variety. However, a shifted toric variety is not necessarily a toric variety. Following the definition of shifted toric varieties and using Proposition 1, we show in the following proposition that ideals of shifted toric varieties have Gröbner bases of a specific form.

Proposition 5. *$V^* \subseteq \mathbb{C}^n$ is a shifted torus if and only if V^* is a zero set of Laurent binomials of the form $\left(\frac{X}{g}\right)^{\alpha} - 1$, where $g \in V^*$ and $\alpha \in \mathbb{Z}^n$. Similarly, V is shifted toric if and only if $I(V)$ is prime and the reduced Gröbner basis (with respect to any term order) of $I(V)$ contains only binomials of the form $X^{\alpha} + cX^{\beta}$ where $c \in \mathbb{C}^*$, $\alpha, \beta \in \mathbb{N}^n \setminus \{(0, \ldots, 0)\}$.*

Proof. V is shifted toric if and only if there exists a torus T and a coset C of T with respect to some $g \in K^n$ (i.e., $C = gT$), such that $V = \overline{C}$. Since V is irreducible, then $\overline{V^*} = \overline{C} = V$ and $I(V) = I(\overline{V^*}) = I(\overline{C})$, and this ideal is prime. Assume that V given by a set of generators of $I(V)$. V is shifted toric if and only if $I(\overline{V^*}) = I(V)$ is prime and V^* is a coset. This holds if and only if $I(V^*)$ is prime and $g^{-1}V^*$ is a group. Note that $\overline{g^{-1}V^*}$ is irreducible if and only if $\overline{V^*}$ is irreducible. This holds if and only if $I(V^*)$ is prime, or equivalently $I(V)$ is prime. Therefore, by Proposition 1, V is shifted toric if and only if $I(V)$ is prime and all the elements of every Gröbner basis of $I(\overline{g^{-1}V^*})$ is of the form $X^{\alpha} - X^{\beta}$, for $\alpha, \beta \in \mathbb{N}^n \setminus \{(0, \ldots, 0)\}$. Equivalently, all the elements of $I(V)$ are of the form $\left(\frac{X}{g}\right)^{\alpha} - \left(\frac{X}{g}\right)^{\beta}$. Cleaning the denominator, we have the desired form of the Gröbner basis elements. □

Proposition 5 along with the structure theorem for binomial ideals imply that the primary decomposition of an ideal generated by binomials of the form $X^{\alpha} - X^{\beta}$ include an ideal generated by binomials of the form $X^{\alpha} - X^{\beta}$ and several ideals generated by binomials of the form $X^{\alpha} - cX^{\beta}$.

Using Proposition 5, one can design a randomised algorithm for testing shifted toricity of a variety $V = \overline{V^*}$. Let $g_1, \ldots, g_m$ be *generic points* in V^* and consider the set of Laurent binomials $G = \left\{ \left(\frac{X}{g_i}\right)^{\gamma_i} - 1 \;\middle|\; i = 1, \ldots, m \right\}$ with symbolic exponents $\gamma_i = (\gamma_{i1}, \ldots, \gamma_{in})$. Let

$$M = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1n} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \cdots & \gamma_{mn} \end{pmatrix} \tag{2}$$

be the matrix of exponents of the Laurent binomials and make it row reduced. This leads to a linear combination of rows with coefficients in $\mathbb{Z}$, say d_i. Then V is shifted toric if and only if

$$\sum_{i=1}^{m} d_i \gamma_i = 0, \tag{3}$$

which holds if and only if $\prod_{i=1}^{m} \left(g_{i1}^{d_{i1}} \dots g_{in}^{d_{in}} \right) = 1$. Solving the linear equations (3) will give us the exponents of the Laurent polynomials. Note that G obtained in this way is a reduced Gröbner basis.

One can see that a binomial variety is the closure of a group and furthermore, by Proposition 3, it can be decomposed into toric and shifted toric varieties as its irreducible components. A natural question is whether this property holds for every variety that is the closure of a group. The answer to this question is positive and indeed such varieties are precisely binomial varieties. This is explicitly formulated in a remark in [23, after Proposition 2.3].

Proposition 6. *Let W be a subgroup of $(\mathbb{C}^*)^n$. Then $\overline{W}$ is a binomial variety.*

Proof. By definition of binomial variety, we have to prove that $\overline{W}$ is the zero set of binomials of the form $X^\alpha - X^\beta$ and $\overline{(\overline{W})^*} = \overline{W}$. The equality $\overline{(\overline{W})^*} = \overline{W}$ directly comes from the definition of Zariski closure.

For proving that the generators of $I(\overline{W})$ have the desired form, we use the notations of [23]. By Proposition 2.3(a) in the latter reference, $\mathbb{C}[X^{\pm 1}]I(W) = I(\rho)$, for some partial character $\rho \in \mathrm{Hom}(\mathbb{Z}^n, \mathbb{C}^*)$. By Theorem 2.1(b) in the same reference, $I(\rho) = \langle x^{m_1} - \rho(m_1), \dots, x^{m_r} - \rho(m_r) \rangle$ where $m_1, \dots, m_r \in \mathbb{Z}^n$ is a basis of L_ρ. As W is a group, $(1, \dots, 1) \in W$; hence $\rho(m_1) = \dots = \rho(m_r) = 1$ and therefore $I(W)$ is generated by binomials of the form $X^\alpha - X^\beta$ where $\alpha, \beta \in \mathbb{N}$. Since $\overline{W} = V(I(W))$, we have proved the proposition. □

Proposition 6 can also be proved over the positive real numbers by considering the logarithm map on $(\mathbb{R}_{>0})^n$ acting coordinate-wise. The image of this map forms a linear space. A basis of this linear space provides a parametrization of a group.

From a computational point of view, a variety $V = V(I)$ is usually given by a set of generators of I and we would like to derive information about toricity, binomiality or coset property of V by computations over the generators of I. This can be done via Gröbner bases. Assume that G is a Gröbner basis of I, hence $V = V(G)$, and $V^* \neq \varnothing$. If all elements of G are of the form $X^\alpha - X^\beta$, then V^* is a subgroup of $(\mathbb{C}^*)^n$. If all elements of G are of the form $c_\alpha X^\alpha - c_\beta X^\beta$ where $c_\alpha \neq 0$ and $c_\beta \neq 0$, then V^* is a coset of a subgroup of $(\mathbb{C}^*)^n$. Note that the converse of the above does not hold. This is because $\overline{V^*}$ and V may not be equal and therefore $I(\overline{V^*})$ and $I(V)$ may not be equal, which means that a Gröbner basis of $I(V)$ does not give information about group or coset structure of V^*. In case V is irreducible, we have that $\overline{V^*} = V$, e.g., when V is toric or shifted toric. In order to solve this problem one needs to saturate I with the multiplication of the variables and then consider the radical of this saturation. Saturation removes the points that are in V but not in $\overline{V^*}$. The following proposition states this precisely and is the essence of this section for computations over complex numbers.

Proposition 7. *Let I be an ideal in $\mathbb{Q}[x_1, \dots, x_n]$, let $V := V(I) \subseteq \mathbb{C}^n$ be the variety of I, and let $G \subseteq \mathbb{Q}[x_1, \dots, x_n]$ be a reduced Gröbner basis of the radical of $I : \langle x_1 \dots x_n \rangle^\infty$. Then $V^* = \varnothing$ if and only if $G = \{1\}$. If $V^* \neq \varnothing$, then the following hold:*

(i) V^* is a subgroup of $(\mathbb{C}^*)^n$ if and only if all elements of G are of the form $X^\alpha - X^\beta$.
(ii) V^* is a coset of a subgroup of $(\mathbb{C}^*)^n$ if and only if all elements of G are of the form $c_\alpha X^\alpha - c_\beta X^\beta$ where $c_\alpha \neq 0$ and $c_\beta \neq 0$.

Proof. In order to prove the proposition we use [15, Chapter 4, Theorem 10 (iii)], which states that over an algebraically closed field, we have that $\overline{V(J_1) \setminus V(J_2)} = V(J_1 : J_2^\infty)$ for ideals J_1 and J_2. Set $J_1 = I$ and $J_2 = \langle x_1 \dots x_n \rangle$. Since $V^* = V(I) \setminus V(x_1 \dots x_n)$, and V is a variety over $\mathbb{C}$ which is algebraically closed, we have that $\overline{V^*} = V(I : \langle x_1 \dots x_n \rangle^\infty)$. Then

$$I(\overline{V^*}) = I(V(I : \langle x_1 \dots x_n \rangle^\infty)) = \sqrt{I : \langle x_1 \dots x_n \rangle^\infty}. \tag{4}$$

$V^* = \varnothing$ if and only if $V(I) \setminus V(x_1 \dots x_n) = \varnothing$, which is the case if and only if $V(I) \subseteq V(x_1 \dots x_n)$. This happens if and only if $I(V(I)) = I(V(x_1 \dots x_n))$, if and only if $\langle x_1, \dots, x_n \rangle \subseteq \sqrt{I}$, which is the case if and only if some product of the variables is in I, or equivalently $I : \langle x_1 \dots x_n \rangle^\infty = \langle 1 \rangle$, i.e., $G = 1$.

For proving (i), let V^* be a group. By Proposition 6, we have that all the elements of every Gröbner basis of $I(\overline{V^*})$ are of the form $X^\alpha - X^\beta$. But according to (4) and the assumption that G is a Gröbner basis of $\sqrt{I : \langle x_1 \dots x_n \rangle^\infty}$, this condition holds.

For the converse, let the elements of G have the desired form. Then $V(\sqrt{I : \langle x_1 \dots x_n \rangle^\infty})^*$ is obviously a group. But $V(\sqrt{I : \langle x_1 \dots x_n \rangle^\infty})^* = V(I : \langle x_1 \dots x_n \rangle^\infty)^*$ and therefore the latter is a group. Now using [15, Chapter 4, Theorem 10 (iii)], we have that $V(I : \langle x_1 \dots x_n \rangle^\infty)^* = (\overline{V(I) \setminus V(x_1 \dots x_n)})^*$. One can easily check that the latter is equal to $V(I)^*$. Hence $V^* = V(I)^*$ is a group and we are done.

The proof of part (ii) is analogous to that of part (i) above. □

For the rest of this section, we present the monomial parametrization of a torus and state propositions that allow one to find the cosets of a torus as irreducible components of a binomial variety using roots of unity. Readers primarily interested in our algorithms in Section 3 can safely skip this part.

We start with introducing the monomial parametrization of shifted toric varieties. Let $T \subseteq (K^*)^n$ be a torus of dimension m, hence $T \simeq (K^*)^m$, and let $x_0 \in (K^*)^n$. Following the monomial parametrization of a torus given in [23, Corollary 2.6], one can see that the coset x_0T of T can be seen as the image of the following monomial map, which is the monomial parametrization of x_0T.

$$\varphi_{(x_0,A)} : (K^*)^m \to (K^*)^n, \quad \varphi_{(x_0,A)}(t_1, \dots, t_m) = \left((x_0)_1 \prod_{i=1}^m t_i^{A_{i1}}, \dots, (x_0)_n \prod_{i=1}^m t_i^{A_{in}} \right),$$

where $A \in \mathbb{Z}^{d \times n}$is a rank m matrix. Note that while the matrix A is not unique, it only depends on T and not on x_0. In particular, T is its own coset with respect to the unity $\mathbf{1} := (1, \dots, 1) \in (K^*)^n$. Note that if $B \in \mathbb{Z}^{d \times n}$ is another matrix such that T equals the image of $\varphi_{(\mathbf{1},B)}$, then B corresponds to a re-parametrization of T.

Example 8. $V(xy - 1) \cap (\mathbb{C}^*)^2$ can be seen as the image of $\varphi_{(\mathbf{1},(1,-1))} : \mathbb{C}^* \to (\mathbb{C}^*)^2$ with $\varphi_{(\mathbf{1},(1,-1))}(t) = (t, t^{-1})$ or as the image of $\varphi_{(\mathbf{1},(-1,1))} : \mathbb{C}^* \to (\mathbb{C}^*)^2$ with $\varphi_{(\mathbf{1},(-1,1))}(t) = (t^{-1}, t)$.

Proposition 9. *If $G \subseteq (\mathbb{C}^*)^n$ is a group and reducible into $r \in \mathbb{N}$ cosets of a torus T, then there exist $y_1, \dots, y_r \in (\mathbb{C}^*)^n$ whose coordinates are roots of unity and $G = \bigcup_{i=1}^r y_i T$ is the irreducible decomposition of G.*

Proof. If $G = T$, then it is its own coset with respect to $\mathbf{1}$ and we are done. Otherwise, let $S_i = \tilde{y}_i T$ be a proper coset of T and suppose that there is no $\xi \in (\mathbb{C}^*)^n$ such that the coordinates of ξ are roots of unity and $S_i = \xi T$. This means that for all such ξ the image of $\varphi_{(\tilde{y}_i,T)}$ is different from the image of $\varphi_{(\xi,T)}$. Hence there exists $t \in (\mathbb{C}^*)^m$ such that for all $s \in (\mathbb{C}^*)^m$ one has $\varphi_{(\tilde{y}_i,T)}(t) \neq \varphi_{(\xi,T)}(s)$. In other words, there exists $t \in (\mathbb{C}^*)^m$ such that for all $s \in (\mathbb{C}^*)^m$ one has $\tilde{y}_i t^A \neq \xi s^A$ for some $i \in \{1, \dots, n\}$. As the coordinates of ξ are roots of unity, there is a natural number N such that $\xi^N = \mathbf{1}$. Therefore, for all $s \in (\mathbb{C}^*)^m$ one has $\tilde{y}_i^N t^{NA} \neq s^{NA}$ for some $i \in \{1, \dots, n\}$. As the image of φ is invariant under $A \mapsto NA$, the cosets $\tilde{y}_i^N T$ and T are distinct. By using a similar argument and induction, one can prove that $T, \tilde{y}_i^N T, \tilde{y}_i^{2N} T, \dots$ are distinct. As G is closed under multiplication, it contains all these cosets. However, this contradicts the assumption that G is reducible into a torus and a finite number of its cosets. □

Remark 10. Let S_1 and S_2 be two cosets of a torus $T \subseteq (\mathbb{C}^*)^n$. The coset S_1 is called the complex conjugate of the coset S_2, written $S_1 = S_2^+$, when every point of S_1 is the complex conjugate of a point of S_2 and every point of S_2 is the complex conjugate of a point of S_1. As the complex

conjugate is an automorphism of $(\mathbb{C}^*)^m$, S_1 is the complex conjugate of S_2 if and only if S_1 contains the complex conjugate of some point of S_2. A pair S_1, S_2 is called a pair of complex conjugates if $S_1 = S_2^+$. If $G \subseteq (\mathbb{C}^*)^n$ is reducible into a finite number of cosets of a torus then they come in pairs of complex conjugates. To see this, denote the toric component of G by T. If $G = T$, then clearly $G = G^+$. Suppose that G contains a proper coset S of T. By Lemma 9 there is a point $\xi \in S$ whose coordinates are roots of unity. Then $\xi\xi^+ = \mathbf{1}$. As G is a group, ξ^+ is an element of G. As $S = \xi T$ and $\xi^+ T = (\xi T)^+ = S^+$, we conclude that S^+ is contained in G.

Proposition 11. *Let $P = (X/\xi)^u - (X/\xi)^v \in \mathbb{C}[X]$ be a non-zero irreducible polynomial, where the coordinates of $\xi \in (\mathbb{C}^*)^n$ are roots of unity. Then for all i in $\mathbb{N}$ there exist g_i in $\mathbb{C}[X]$ and α_i, β_i in $\mathbb{N}^n$ such that $g_i P = X^{i\alpha_i} - X^{i\beta_i}$.*

Proof. Note that $P = (X^u - \xi^{u-v}X^v)/\xi^u$. As the coordinates of ξ are roots of unity, $\gamma = \xi^{u-v}$ is also a root of unity. Let m be the smallest positive integer such that $\gamma^m = 1$. As $\mathfrak{U}_m = \{\gamma, \gamma^2, \ldots, \gamma^m\}$ is a group of roots of unity of order m, it is clear that

$$(X^u - \gamma X^v)(X^u - \gamma^2 X^v) \cdots (X^u - \gamma^m X^v) = (X^{mu} - X^{mv}).$$

Hence one can take

$$g_1 = \xi^u (X^u - \gamma^2 X^v)(X^u - \gamma^3 X^v) \ldots (X^u - \gamma^m X^v).$$

Substituting $\mathfrak{U}_m$ with the group $\mathfrak{U}_{mi}$, i.e., the group of roots of unity of order mi, and following the steps for constructing g_1 accordingly, one can construct g_i for all $i \geq 2$. □

3. Algorithmic Classification of Biomodels

We want to apply our concept of shifted toricity to biomodels focusing on the BioModels[4] repository of mathematical models of biological and biomedical systems [8]. The BioModels repository uses the *Systems Biology Markup Language (SBML)* [44, 29]. SBML is a representation format, based on XML, for communicating and storing computational models of biological processes. It is a free and open standard with widespread software support and a community of users and developers. SBML models have been typically created in the context of numerical computations or simulations and must be processed carefully with symbolic computation. For instance, numerical values, like reaction rate constants, contained in the models are often represented as truncated fixed point floats, and the available SBML parsers possibly introduce further rounding errors when implicitly performing substitutions with those values. Such issues are addressed by ODEbase[5], which provides pre-processed versions of BioModels for use in symbolic computation. We consider here *all* models from ODEbase where the vector field of the ODE is polynomial over $\mathbb{Q}$ after application of certain SBML-specific rules and substitution of parameter values. This amounts to a total of 129 models considered in this article.

Following a convention often used in publications on chemical reaction network theory in the context of symbolic computation, ODEbase replaces names of species concentrations by more abstract names x_i using numbers as indices. With the application of SBML rules some of those x_i vanish in the ODEbase toolchain. We therefore consider, more abstractly, *ordered sets X of variables*, tacitly assuming that the order establishes a mapping between indeterminates in $\mathbb{Q}[X]$ and coordinates in $K^{|X|}$. As a matter of fact, the variables will also vanish during our own algorithms discussed throughout this section. The following example illustrates this.

[4] `https://www.ebi.ac.uk/biomodels/`
[5] `http://odebase.cs.uni-bonn.de/`

Example 12 (BIOMD0000000198). Consider the following system in $\mathbb{Q}[x_2, \ldots, x_{10}]$:

$$F = \{-350x_2 + 800x_3, 350x_2 - 1650x_3, 4250x_3 - 100x_4 + x_5, 100x_4 - x_5,$$
$$-350x_6 + 800x_7, 350x_6 - 1650x_7, 1700x_7 - 5x_8 + 50x_9, x_{10} + 125x_8 - 1330x_9, -x_{10} + 80x_9\}.$$

From its Gröbner basis $G = \{x_2, x_6, x_5 - 100x_4, 8x_8 - x_{10}, x_7, x_3, -x_{10} + 80x_9\}$ we can read off that for every point in $V_{\mathbb{C}}(F) \subseteq \mathbb{C}^9$, e.g., the x_2-coordinate must be 0. It follows that $V_{\mathbb{C}}(F)^* = \varnothing$. Geometrically, $V_{\mathbb{C}}(F)$ lives in $\mathbb{C}^5$ with coordinates x_4, x_5, x_8, x_9, x_{10}. Thinking about toricity as a geometric notion, it makes sense to study the variety as an object in that lower dimensional space. Hence, consider

$$\hat{G} = G \setminus \{x_2, x_3, x_6, x_7\} = \{x_5 - 100x_4, 8x_8 - x_{10}, -x_{10} + 80x_9\} \subseteq \mathbb{Q}[x_4, x_5, x_8, x_9, x_{10}].$$

It turns out that $V_{\mathbb{C}}(\hat{G})$ is shifted toric in $\mathbb{C}^5$.

Definition 13 (Maximal Projected Subset of Variables). Let $K \in \{\mathbb{R}, \mathbb{C}\}$, let X be an ordered set of variables, and let V_K be a variety in $K^{|X|}$. There exists a unique maximal subset $\hat{X} \subseteq X$, depending on K, such that for all $x \in \hat{X}$ there exists $a \in V_K(F)$ with x-coordinate different from 0. We call $\hat{X}$ the *maximal projected subset*.

The maximal projected subset $\hat{X}$ gives us coordinates for the highest dimensional subspace $K^{|\hat{X}|}$, where the variety V_K does not entirely vanish in any coordinate. We consider this a canonical choice for checking shifted toricity. When V_K is given as $V_K(F)$ for $F \subseteq \mathbb{Q}[X]$, then we define $\hat{F} = F \cap \mathbb{Q}[\hat{X}]$, as we did with the Gröbner basis G in Example 12. Note that $V_K(\hat{F})^*$ can still be empty in $K^{|\hat{X}|}$.

The principle domain of interest for us is $\mathbb{R}$-space, where, e.g., concentrations of species are located in the interior of the first orthant. In the literature there has been considerable attention to $\mathbb{C}$-space. We therefore start our algorithmic considerations over $\mathbb{C}$ in Subsection 3.1, and then turn to $\mathbb{R}$ in Subsection 3.2.

In Example 12 we could conclude that $V_{\mathbb{C}}(\hat{F})^*$ is shifted toric because $\hat{F}$ consists of binomials of the characteristic shape according to Proposition 7, and one can easily see from its linearity that it generates a prime ideal over $\mathbb{C}$. As prime ideal decomposition is related to polynomial factorization, decomposition or even primeness tests over our fields $\mathbb{C}$ and $\mathbb{R}$ of interest are not well supported in off-the-shelf computer algebra systems. In our algorithms we therefore limit ourselves to the properties "group" and "coset" rather than "toric" and "shifted toric". Nevertheless, we will consider prime decompositions over $\mathbb{Q}$, which are well supported in software and provide at least partial decompositions over $\mathbb{C}$ and $\mathbb{R}$. Note that for us the relevant notion is prime decomposition in contrast to primary decomposition, as the former corresponds to the irreducibility of the corresponding varieties.

Example 14 (Comparison of $V_{\mathbb{C}}^*$ with $V_{\mathbb{R}}^*$). For $F_1 = \langle x^2+2 \rangle$, $F_2 = \langle (x^2-1)(x^2+2) \rangle$ the following holds:

(i) $V_{\mathbb{C}}(F_1)^* = \{i\sqrt{2}, -i\sqrt{2}\}$ is a coset in $\mathbb{C}^*$, because $(-i\sqrt{2})^{-1} V_{\mathbb{C}}(F_1)^* = \{1, -1\}$ is a group. In contrast, $V_{\mathbb{R}}(F_1)^* = \varnothing$ is not a coset in $\mathbb{R}^*$.

(ii) $V_{\mathbb{C}}(F_2)^* = \{1, -1, i\sqrt{2}, -i\sqrt{2}\}$ is a coset in $\mathbb{C}^*$ if and only if it is group, due to $1 \in V_{\mathbb{C}}(F_2)^*$. This is not the case because it is not closed under multiplication: $(i\sqrt{2})(-i\sqrt{2}) = 2 \notin V_{\mathbb{C}}(F_2)^*$. In contrast, $V_{\mathbb{R}}(F_2)^* = \{1, -1\}$ is a group.

3.1. Classification over $\mathbb{C}$

Our methods over $\mathbb{C}$ are, naturally, based on Gröbner bases [7, 24, 25], for which we rely on the commercial computer algebra system Maple. We generally leave it to Maple to find a good term order. Our classifications hold over any algebraically closed extension field of the coefficient field $\mathbb{Q}$, including $\mathbb{C}$ as well as, e.g., the countable algebraic closure of $\mathbb{Q}$.

With the following discussion of Algorithm 1 we will introduce some textbook facts from commutative algebra as lemmas, together with short proofs. The algorithm recognizes for a given ideal basis F whether $V_{\mathbb{C}}(F)^*$ is a coset. In line 1 we compute a Gröbner basis G of F. Recall from

Algorithm 1 $\mathrm{ProjectAndClassify}_{\mathbb{C}}$

Input: 1. X, a finite ordered set of variables; 2. $F \subseteq \mathbb{Q}[X]$ finite and non-empty

Output: 1. $\hat{X} \subseteq X$; 2. $\gamma \in \{\mathtt{G}, \mathtt{C}, \mathtt{O}, \mathtt{X}, \mathtt{g}, \mathtt{c}, \mathtt{o}, \mathtt{x}\}$

$\hat{X}$ is the unique maximal subset of X with respect to $V_{\mathbb{C}}(F)$. The letter γ classifies $V_{\mathbb{C}}(F)^*$ in $\hat{X}$-space, using upper case when $\hat{X} = X$:

`G/g` – $V_{\mathbb{C}}(F)^*$ is a group; `C/c` – $V_{\mathbb{C}}(F)^*$ is a proper coset; `O/o` – $V_{\mathbb{C}}(F)^* = \varnothing$; `X/x` else.

1: $G := \mathrm{GroebnerBasis}(F)$
2: $X' := G \cap X$
3: $\hat{X} := X \setminus X'$
4: $\hat{G} := G \setminus X'$ ▷ $\langle \varnothing \rangle = \langle 0 \rangle$
5: $\tilde{G} := \mathrm{Radical}(\mathrm{Saturate}(\hat{G}, \prod \hat{X}))$ ▷ $\tilde{G}$ is a Gröbner basis
6: $\gamma := \mathrm{Classify}_{\mathbb{C}}(\hat{X}, \tilde{G})$
7: **if** $\hat{X} \neq X$ **then**
8: convert γ to a lower case letter
9: **end if**
10: **return** $\hat{X}, \gamma$

the previous section that we generally consider reduced Gröbner bases. We may safely assume that $\mathrm{vars}(G) \subseteq \mathrm{vars}(F)$. In line 2, the variables X' occurring as elements of G are exactly those that must be zero for all points in $V_{\mathbb{C}}(F)$. Removing X' from X in line 3 yields the unique maximal projected subset $\hat{X}$ of X according to Definition 13. Removing X' from G in line 4 is equivalent to plugging 0 into all X' in G, which in turn realizes the projection of $V_{\mathbb{C}}(F) = V_{\mathbb{C}}(G)$ into $\hat{X}$-space. Note that we follow the convention that the empty set is a generator of the trivial ideal [3, Definition 1.36]. In line 5 we obtain $\tilde{G}$ by saturating $\hat{G}$ and subsequently taking the radical. In line 6 we call Algorithm 2 in order to apply Proposition 7 with $I = \langle \hat{G} \rangle$ and $G = \tilde{G}$.

Algorithm 2 $\mathrm{Classify}_{\mathbb{C}}$

Input: 1. $\hat{X}$, a finite ordered set of variables; 2. $\tilde{G} \subseteq \mathbb{Q}[\hat{X}]$, a Gröbner basis of a saturated radical ideal

Output: $\gamma \in \{\mathtt{G}, \mathtt{C}, \mathtt{O}, \mathtt{X}\}$

1: **if** $\hat{X} = \varnothing$ **or** $\tilde{G} = \{1\}$ **then**
2: **return** `O`
3: **else if** all elements of $\tilde{G}$ are of the form $\hat{X}^\alpha - \hat{X}^\beta$ with $\alpha, \beta \in \mathbb{N}^m$ **then**
4: **return** `G`
5: **else if** all elements of $\tilde{G}$ are of the form $c_\alpha \hat{X}^\alpha - c_\beta \hat{X}^\beta$ with $\alpha, \beta \in \mathbb{N}^m$, $c_\alpha \neq 0$, $c_\beta \neq 0$ **then**
6: **return** `C`
7: **else**
8: **return** `X`
9: **end if**

In line 1 of Algorithm 2, if $\hat{X} = \varnothing$, then we are in zero-dimensional $\mathbb{C}$-space and certainly $V_{\mathbb{C}}(\hat{G})^* = \varnothing$. Otherwise $\tilde{G} = \{1\}$ is an equivalent criterion for $V_{\mathbb{C}}(\hat{G})^* = \varnothing$ by Proposition 7.

From line 3 on we know that $V_{\mathbb{C}}(\hat{G})^* \neq \varnothing$ and apply in line 3 and line 5 the criteria from part (i) and (ii) of Proposition 7, respectively. In the negative case we return `X` in line 8.

This takes us back to line 7 of Algorithm 1. For convenience, we patch the classification letter γ from upper case to lower case when proper projection has taken place. That information could be also reconstructed by comparing the variables X of the input F with the unique maximal subset $\hat{X}$ of X, which is returned in line 10 along with γ.

Example 15 (`BIOMD0000000519`). Consider $F = \{f_1, f_2, f_3\} \subseteq \mathbb{Q}[X]$, where $X = \{x_1, x_2, x_3\}$:

$$\begin{aligned}
f_1 = &-1105691950605246617909660 49x_1^2 - 1105691950605246617909660 49x_1x_2\\
&- 1105691950605246617909660 49x_1x_3 + 8268303407262959414915925880x_1,\\
f_2 = &-3934051960253477029254203 7060x_1^2 - 6471647090416070818162569 9581x_1x_2\\
&- 3934051960253477029254203 7060x_1x_3 + 4720862352304172435105044447200x_1\\
&- 2537595130162593788908366 2521x_2^2 - 2537595130162593788908366 2521x_2x_3\\
&+ 1783712878395505546690039502520x_2,\\
f_3 = &-4054220223364235403697211 2493x_1x_2 - 4054220223364235403697211 2493x_2^2\\
&- 4054220223364235403697211 2493x_2x_3 + 4865064268037082484436653499160x_2\\
&- 1101385347722460000000000000000x_3.
\end{aligned}$$

We obtain $\hat{X} = X$ and $\hat{G} = G = \{\hat{g}_1, \ldots, \hat{g}_4\}$. For space reasons, we present $\hat{g}_1, \ldots, \hat{g}_4$ with approximate coefficients here:

$$\begin{aligned}
\hat{g}_1 &\approx 5.72 \times 10^{41} x_3 - 1.05 \times 10^{42} x_2 + 1.47 \times 10^{42} x_1,\\
\hat{g}_2 &\approx 3.63 \times 10^{80} x_1^2 - 6.37 \times 10^{80} x_1,\\
\hat{g}_3 &\approx 8.89 \times 10^{67} x_1 x_2 - 2.44 \times 10^{69} x_1,\\
\hat{g}_4 &\approx 2.34 \times 10^{111} x_2^2 - 9.39 \times 10^{112} x_1 - 5.82 \times 10^{112} x_2.
\end{aligned}$$

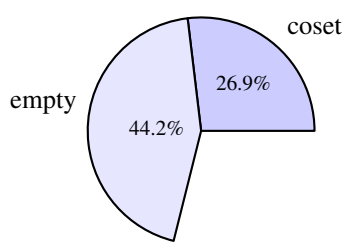
Notice that g_4 is not binomial. After saturation we obtain $\tilde{G} = \{\tilde{g}_1, \tilde{g}_2, \tilde{g}_3\}$ with

$$\tilde{g}_1 = \hat{g}_2, \quad \tilde{g}_2 = \hat{g}_3, \quad \tilde{g}_3 \approx 2.66 \times 10^{92} x_3 - 1.21 \times 10^{94},$$

which is classified as $\gamma =$ `C`.

It is important to understand that, although we are using tools from ideal theory, our results in Section 2 clarify that our classification solely depends on geometry. In particular, results are invariant with respect to the input ideal basis F in Algorithm 1.

In Appendix A.1 we discuss practical aspects of our implementation and give in Table 1 classification results from applying Algorithm 1 to the 129 models introduced at the beginning of this section. For our discussion here we note that our algorithm terminates within a time limit of 6 hours per model on 104 out of the 129 models. We obtain 2 `G`, 20 `C`, and 6 `c`, which can be summarized as $V_{\mathbb{C}}(\tilde{G})^*$ forming a coset. Furthermore we have 4 `O` and 42 `o`, i.e., $V_{\mathbb{C}}(\tilde{G})^* = \varnothing$. The rest is 29 `X` and one single `x`. In terms of percentages of the 104 successful computations this yields the following picture:



We are now going to turn to prime decompositions over $\mathbb{Q}$ of the generating ideals F of our varieties $V_{\mathbb{C}}(F)$. Recall that shifted toricity requires, in addition to the coset structure of $V_{\mathbb{C}}(F)^*$, irreducibility of $V_{\mathbb{C}}(F)$, which in turn corresponds to prime decompositions of F even over $\mathbb{C}$. From that point of view, our decompositions considered here are only a heuristic step into the right direction. On the other hand, the following example suggests that beyond the irreducibility issue, prime decompositions over $\mathbb{Q}$ can improve our hit rate on cosets.

Example 16 (BIOMD0000000359). Consider $F \subseteq \mathbb{Q}[X]$, where $X = \{x_1, \ldots, x_7, x_9\}$:

$$\begin{aligned} F = \{&-125x_1x_2 - 125x_1x_5 - 11x_1x_7 + 19250x_3 + 19250x_4, -5x_1x_2 + 20x_3x_7 + 770x_3,\\ &5x_1x_2 - 20x_3x_7 - 1190x_3, 250x_1x_5 - 300x_4x_6 + 21000x_3 - 38500x_4 + x_9,\\ &-2500x_1x_5 - 27x_5x_6 + 385000x_4 + 10x_7, -3000x_4x_6 - 27x_5x_6 + 10x_7 + 10x_9,\\ &-220x_1x_7 - 10000x_3x_7 + 27x_5x_6 - 10x_7, 11x_1x_7, 1000x_3x_7 + 300x_4x_6 - x_9\}. \end{aligned}$$

Applying Algorithm 1 to X and F yields $\hat{X} = \{x_1, x_2, x_4, \ldots, x_7\}$, $\hat{G} = \{x_1x_7, x_4x_7, x_1x_5 - 154x_4, x_1x_2, x_4x_6, 27x_5x_6 - 10x_7, x_2x_4\}$, and the saturated radical basis $\tilde{G} = \{1\}$. The classification result is $\hat{X}$ together with $\gamma = \mathtt{o}$.

The following is a prime decomposition of F over $\mathbb{Q}$:

$$\begin{aligned} \mathcal{P} = (&\{x_1, x_3, x_4, x_9, -27x_5x_6 + 10x_7\}, \{x_1, x_3, x_4, x_5, x_7, x_9\}, \{x_1, x_3, x_4, x_6, x_7, x_9\},\\ &\{x_2, x_3, x_4, x_5, x_7, x_9\}, \{x_2, x_3, x_6, x_7, x_9, x_1x_5 - 154x_4\}). \end{aligned}$$

Projection of every individual prime component yields respective maximal projected subsets of variables $\hat{\mathcal{X}} = (\{x_5, x_6, x_7\}, \varnothing, \varnothing, \varnothing, \{x_1, x_4, x_5\})$ and Gröbner bases

$$\tilde{\mathcal{G}} = \hat{\mathcal{G}} = (\{-27x_5x_6 + 10x_7\}, \{0\}, \{0\}, \{0\}, \{x_1x_5 - 154x_4\}),$$

which are already saturated. Application of Algorithm 2 to pairs of elements of $\hat{\mathcal{X}}$ and $\tilde{\mathcal{G}}$ yields $\Gamma = (\mathtt{c}, \mathtt{o}, \mathtt{o}, \mathtt{o}, \mathtt{c})$. This tells us that $V_{\mathbb{C}}(F)^*$ has two components, which live in different 3-dimensional subspaces of $\mathbb{C}^8$. Both of them are cosets.

Algorithm 3 formalizes the approach outlined in Example 16. We use the Weierstrass $\wp$ for power sets. The algorithm starts with the computation of a prime decomposition in line 1. We have

$$\bigcup_{i=1}^{k} V_{\mathbb{C}}(P_i) = V_{\mathbb{C}}(F) = V_{\mathbb{C}}(\sqrt{\langle F\rangle}) \quad \text{and} \quad \bigcap_{i=1}^{k} \langle P_i\rangle = \sqrt{\langle F\rangle}.$$

Note that the obtained prime ideals $\langle P_i\rangle$ are also radical.

Lemma 17. *Let $I \subseteq \mathbb{Q}[X]$ be a prime ideal. Then I is a radical ideal.*

Proof. Let $f^s \in I$. We show by induction on s that $f \in I$. If $s = 1$, then we are done. Otherwise consider $f^s = ff^{s-1} \in I$. Since I is prime, we have $f \in I$ or $f^{s-1} \in I$. In the latter case, $f \in I$ by the induction hypothesis. □

In lines 3–10, Algorithm 3 follows in a for-loop essentially Algorithm 1 for each prime component basis P_i. In line 5 we note that $\langle \hat{P}_i\rangle$ is prime by the following lemma.

Lemma 18. *Let $G \subseteq \mathbb{Q}[X]$ be a reduced Gröbner basis of a prime ideal $\langle G\rangle$, and let $x \in G \cap X$. Then $\langle G \setminus \{x\}\rangle$ is prime.*

Proof. Notice that $\langle G \setminus \{x\}\rangle$ is the elimination ideal $\langle G\rangle_x = \langle G\rangle \cap \mathbb{Q}[X \setminus \{x\}]$. Let $fg \in \langle G\rangle_x \subseteq \langle G\rangle$. Then w.l.o.g. $f \in \langle G\rangle$, because $\langle G\rangle$ is prime. Since x does not occur in fg, it does not occur in f either. Hence $f \in \langle G\rangle_x$. □

Algorithm 3 DecomposeProjectAndClassify$_{\mathbb{C}}$

Input: 1. X, a finite ordered set of variables; 2. $F \subseteq \mathbb{Q}[X]$ finite and non-empty

Output: 1. $\mathcal{P} \in \wp(\mathbb{Q}[X])^k$; 2. $\hat{\mathcal{X}} \in \wp(X)^k$; 3. $\Gamma \in \{\mathtt{G}, \mathtt{C}, \mathtt{O}, \mathtt{X}, \mathtt{g}, \mathtt{c}, \mathtt{o}, \mathtt{x}\}^k$

$\mathcal{P} = (P_1, \ldots, P_k)$ are Gröbner bases of a prime decomposition over $\mathbb{Q}$ of $\langle F \rangle$.
In $\hat{\mathcal{X}} = (\hat{X}_1, \ldots, \hat{X}_k)$, X_i is the unique maximal subset of X with respect to $V_{\mathbb{C}}(P_i)$.
In $\Gamma = (\gamma_1, \ldots, \gamma_k)$, the letter γ_i classifies $V_{\mathbb{C}}(P_i)^*$ in $\hat{X}_i$-space, using upper case when $\hat{X}_i = X$:
$\mathtt{G}/\mathtt{g}$ – $V_{\mathbb{C}}(P_i)^*$ is a group; $\mathtt{C}/\mathtt{c}$ – $V_{\mathbb{C}}(P_i)^*$ is a proper coset; $\mathtt{O}/\mathtt{o}$ – $V_{\mathbb{C}}(P_i)^* = \varnothing$; $\mathtt{X}/\mathtt{x}$ else.

1: $\mathcal{P} = (P_1, \ldots, P_k) := \mathrm{PrimeDecomposition}_{\mathbb{Q}}(F)$ ▷ $P_1, \ldots, P_k$ are Gröbner bases
2: **for** $i = 1, \ldots, k$ **do**
3: $\quad X_i' := P_i \cap X$
4: $\quad \hat{X}_i := X_i \setminus X_i'$
5: $\quad \hat{P}_i := P_i \setminus X_i'$ ▷ $\langle \varnothing \rangle = \langle 0 \rangle$
6: $\quad \tilde{P}_i := \mathrm{Saturate}(\hat{P}_i, \prod \hat{X}_i)$ ▷ $\tilde{P}_i$ is a Gröbner basis; the product runs over the set $\hat{X}_i$
7: $\quad \gamma_i := \mathrm{Classify}_{\mathbb{C}}(\hat{X}_i, \tilde{P}_i)$ ▷ call Algorithm 2
8: $\quad$ **if** $\hat{X}_i \neq X$ **then**
9: $\quad\quad$ convert γ_i to a lower case letter
10: $\quad$ **end if**
11: **end for**
12: **return** $\mathcal{P}, (\hat{X}_1, \ldots, \hat{X}_k), (\gamma_1, \ldots, \gamma_k)$

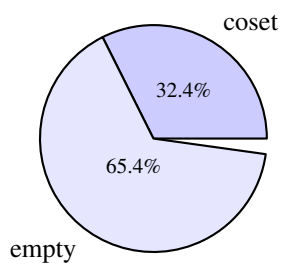
It follows that $\langle \hat{P}_i \rangle$ is also radical by Lemma 17. When computing $\tilde{P}_i$ in line 6 primality is again preserved, as the following lemma shows.

Lemma 19. *Let $I \subseteq \mathbb{Q}[X]$ be a prime ideal, and let $f \in \mathbb{Q}[X]$. Then $I : \langle f \rangle^\infty$ is a prime ideal.*

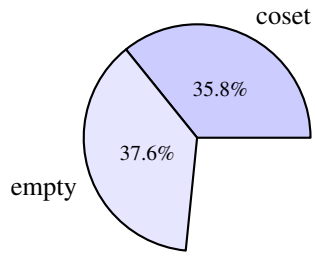
Proof. Recall that $I \subseteq I : \langle f \rangle^\infty$. Let $gh \in I : \langle f \rangle^\infty$. We must show that $g \in I : \langle f \rangle^\infty$ or $h \in I : \langle f \rangle^\infty$. By definition there is $s \in \mathbb{N}$ such that $f^s gh \in I$. If $f^s \in I$, then $I : \langle f \rangle^\infty = \langle 1 \rangle$, which is prime. Otherwise $gh \in I$ and therefore $g \in I \subseteq I : \langle f \rangle^\infty$ or $h \in I \subseteq I : \langle f \rangle^\infty$. □

We once more call Lemma 17 to obtain that $\langle \tilde{P}_i \rangle$ is also radical. Therefore, the radical ideal computation in line 5 of Algorithm 1 is not necessary here.

In Appendix A.2 we discuss practical aspects of our computations and give in Table 3 classification results using Algorithm 3 on the 129 models introduced at the beginning of this section. We succeed on 105 out of the 129 models within a time limit of 6 hours per model. This yields 3426 prime components to test altogether. We obtain 2 G, 22 C, and 1085 c, which can be summarized as $V_{\mathbb{C}}(\tilde{P}_i)^*$ forming a coset. Furthermore, we have 2242 o, i.e., $V_{\mathbb{C}}(\tilde{P}_i)^* = \varnothing$. The rest is only 34 X and 41 x. Again we visualize these results in terms of percentages of the total of 3426 prime components:



Recall that our selection from the BioModels repository presented here is essentially complete with respect to polynomial examples. This comes with the disadvantage that our data is somewhat dominated by `BIOMD0000000281`, which contributes 1008 `c` and 2136 `o`. We have verified that the ideal dimensions for the 1008 `c` components are positive, pointing at non-trivial coset structures in contrast to isolated points. For the sake of scientific rigor we also present the statistics without `BIOMD0000000281`:



3.2. Classification over $\mathbb{R}$

Our primary tool over $\mathbb{R}$ is real quantifier elimination [60, 11, 12, 63, 37, 64]. We use implementations by the fourth author and his students [20] in Redlog [19, 57], which is integrated with the open-source computer algebra system Reduce [41, 42]. Our strategy is to apply virtual substitution methods [64, 48, 47] for quantifier elimination within the relevant degree bounds and fall back into partial cylindrical algebraic decomposition [18, 55] with subproblems where this is not possible. Our results hold over any real-closed field, including $\mathbb{R}$ as well as, e.g, the countable field of real algebraic numbers.

Algorithm 4 is the real counterpart to Algorithm 1 in Subsection 3.1. In line 4 we construct for

Algorithm 4 $\mathrm{ProjectAndClassify}_{\mathbb{R}}$

Input: 1. X, a finite ordered set of variables; 2. $F \subseteq \mathbb{Q}[X]$ finite and non-empty

Output: 1. $\hat{X} \subseteq X$; 2. $\gamma \in \{\mathtt{G}, \mathtt{C}, \mathtt{O}, \mathtt{X}, \mathtt{g}, \mathtt{c}, \mathtt{o}, \mathtt{x}\}$

$\hat{X}$ is the unique maximal subset of X with respect to $V_{\mathbb{R}}(F)$. The letter γ classifies $V_{\mathbb{R}}(F)^*$ in $\hat{X}$-space, using upper case when $\hat{X} = X$:

`G/g` – $V_{\mathbb{R}}(F)^*$ is a group; `C/c` – $V_{\mathbb{R}}(F)^*$ is a proper coset; `O/o` – $V_{\mathbb{R}}(F)^* = \varnothing$; `X/x` else.

1: $\hat{X} := X$
2: $\hat{F} := F$
3: **for** $x_i \in X$ **do**
4: **if** $\mathbb{R} \models \underline{\forall}(\bigwedge_{f \in \hat{F}} f = 0 \longrightarrow x_i = 0)$ **then**
5: $\hat{X} := \hat{X} \setminus \{x_i\}$
6: $\hat{F} := \hat{F}[x_i/0]$
7: **end if**
8: **end for**
9: $\gamma := \mathrm{Classify}_{\mathbb{R}}(\hat{X}, \hat{F})$
10: **if** $\hat{X} \neq X$ **then**
11: convert γ to a lower case letter
12: **end if**
13: **return** $\hat{X}, \gamma$

each $x_i \in X$ the first-order L_{OR}-formula, where L_{OR} denotes the language of ordered rings:

$$\psi \doteq \underline{\forall}\Big(\bigwedge_{f \in \hat{F}} f = 0 \longrightarrow x_i = 0 \Big),$$

The underlined universal quantifier denotes the universal closure, which universally quantifies all variables freely occurring within its scope. Our formula ψ straightforwardly states that for all points in $V_{\mathbb{R}}(\hat{F})$ with coordinates $x_j \in X$, which occur as variables in the polynomials $f \in \hat{F}$, the specific coordinate x_i is zero. The if-condition $\mathbb{R} \models \psi$ expresses that $\mathbb{R}$ is a model of this formula, meaning that the formula holds in $\mathbb{R}$ or, equivalently, in the model class of real closed fields.

A *real quantifier elimination procedure* computes for any given first-order L_{OR}-formula φ an equivalent quantifier-free L_{OR}-formula φ', where the variables in φ' are a subset of the variables freely occurring in φ. Since ψ contains no free occurrences of variables, the corresponding ψ' will be variable-free and can be easily simplified to either true or false. In the former case the if-condition holds, in the latter case it does not.

When some x_i is identified to vanish in all points of $V_{\mathbb{R}}(\hat{F})$ it is removed from $\hat{X}$ in line 5. Accordingly, x_i is set to zero within $\hat{F}$ in line 6, where $[x_i/0]$ is a postfix operator substituting the term 0 for the variable x_i in its argument $\hat{F}$. From line 9 on, Algorithm 4 proceeds like its complex counterpart Algorithm 1 but using Algorithm 5 for real classification. In line 1 we define

Algorithm 5 $\text{Classify}_{\mathbb{R}}$

Input: 1. $\hat{X}$, a finite ordered set of variables, w.l.o.g. $\hat{X} = \{x_1, \ldots, x_n\}$; 2. $\hat{F} \subseteq \mathbb{Q}[\hat{X}]$ finite
Output: $\gamma \in \{\texttt{G}, \texttt{C}, \texttt{O}, \texttt{X}\}$

1: define operator $\Phi(t_1, \ldots, t_n) := (\bigwedge_{f \in \hat{F}} f = 0)[x_1/t_1, \ldots, x_n/t_n]$
2: **if** $\hat{X} = \varnothing$ **or** $\mathbb{R} \not\models \exists(\bigwedge_{i=1}^{n} x_i \neq 0 \wedge \Phi(x_1, \ldots, x_n))$ **then**
3: **return** O
4: **end if**
5: $\tau_{\text{inv}} := \underline{\forall}(\bigwedge_{i=1}^{n} g_i \neq 0 \wedge \bigwedge_{i=1}^{n} x_i \neq 0 \wedge$
$\Phi(g_1, \ldots, g_n) \wedge \Phi(g_1 x_1, \ldots, g_n x_n) \longrightarrow \Phi(g_1 x_1^{-1}, \ldots, g_n x_n^{-1})$
6: **if** $\mathbb{R} \not\models \tau_{\text{inv}}$ **then**
7: **return** X
8: **end if**
9: $\tau_{\text{mult}} := \underline{\forall}(\bigwedge_{i=1}^{n} g_i \neq 0 \wedge \bigwedge_{i=1}^{n} x_i \neq 0 \wedge \bigwedge_{y=1}^{n} y_i \neq 0 \wedge \Phi(g_1, \ldots, g_n) \wedge$
$\Phi(g_1 x_1, \ldots, g_n x_n) \wedge \Phi(g_1 y_1, \ldots, g_n y_n) \longrightarrow \Phi(g_1 x_1 y_1, \ldots, g_n x_n y_n)$
10: **if** $\mathbb{R} \not\models \tau_{\text{mult}}$ **then**
11: **return** X
12: **end if**
13: $\tau_{\text{group}} := \Phi(1, \ldots, 1)$
14: **if** $\mathbb{R} \models \tau_{\text{group}}$ **then**
15: **return** G
16: **else**
17: **return** C
18: **end if**

$\Phi(t_1, \ldots, t_n)$ to generate a first order L_{OR}-formula which states that $(t_1, \ldots, t_n) \in V_{\mathbb{R}}(\hat{F})$, where the t_i are L_{OR}-terms. In lines 2–4 we handle the case $V_{\mathbb{R}}(\hat{F})^* = \varnothing$. Hence in line 5 we know $V_{\mathbb{R}}(\hat{F})^* \neq \varnothing$. We are going to use the following characterization of cosets.

Proposition 20. *Let K^* be a multiplicative group. Let $C \subseteq (K^*)^n$, $C \neq \varnothing$. Then the following are equivalent:*

(i) C *is a coset;*
(ii) there exists $g_0 \in (K^*)^n$ *such that* $g_0^{-1}C$ *is a group;*
(iii) there exists $g_0 \in C$ *such that* $g_0^{-1}C$ *is a group;*
(iv) for all $g \in C$ *we have that* $g^{-1}C$ *is a group.*

Proof. The equivalence between (i) and (ii) is the definition of a coset. When (ii) holds, then $C = g_0G$ for a group G, hence $g_0 \cdot 1 \in C$, which shows (iii). The implication from (iii) to (ii) is obvious, and so is the implication from (iv) to (iii). It remains to be shown that (iii) implies (iv).

Assume that $g_0 \in C$ and $G = g_0^{-1}C$ is a group; equivalently $C = g_0G$. Let $g \in C$. Then there is $y \in G$ such that $g = g_0y$. It follows that $g^{-1}C = (g_0y)^{-1}C = y^{-1}g_0^{-1}C = y^{-1}G = G$. □

Proposition 20(iv) yields a first-order characterization for $V_{\mathbb{R}}(\hat{F})^*$ to be a coset, which could be informally stated as follows:

$$\forall g, x, y \in (\mathbb{R}^*)^n \colon g \in V_{\mathbb{R}}(\hat{F}) \wedge gx \in V_{\mathbb{R}}(\hat{F}) \wedge gy \in V_{\mathbb{R}}(\hat{F}) \Rightarrow gx^{-1} \in V_{\mathbb{R}}(\hat{F}) \wedge gxy \in V_{\mathbb{R}}(\hat{F}). \quad (5)$$

As a first-order L_{OR}-formula this yields:

$$\tau \doteq \underline{\forall}\Bigg(\bigwedge_{i=1}^{n} g_i \neq 0 \wedge \bigwedge_{i=1}^{n} x_i \neq 0 \wedge \bigwedge_{y=1}^{n} y_i \neq 0 \wedge \Phi(g_1, \ldots, g_n) \wedge \Phi(g_1x_1, \ldots, g_nx_n) \wedge \Phi(g_1y_1, \ldots, g_ny_n) \longrightarrow \Phi(g_1x_1^{-1}, \ldots, g_nx_n^{-1}) \wedge \Phi(g_1x_1y_1, \ldots, g_nx_ny_n) \Bigg).$$

In the equations originating from $\Phi(g_1x_1^{-1}, \ldots, g_nx_n^{-1})$ principal denominators containing variables from $x_1, \ldots, x_n$ can be equivalently dropped, because the left hand side of the implication requires those variables to be different from zero. The first-order L_{OR}-formula τ can be equivalently transformed into $\tau_{\text{inv}} \wedge \tau_{\text{mult}}$ with τ_{inv} and τ_{mult} as in line 5 and line 9 of Algorithm 5, respectively. Therefore it is correct to exit with $\gamma = \texttt{X}$ in line 7 or 11 when either part does not hold. This splitting into subproblems has two advantages. First, separate quantifier eliminations on smaller problems are more efficient. Second, when τ_{inv} does not hold in line 6, then τ_{mult} need not be considered at all.

When reaching line 13, we know that $V_{\mathbb{R}}(\hat{F})^*$ is a coset and apply the following corollary, which concludes our discussion of Algorithm 5.

Corollary 21. *Let* C *be a coset. Then* C *is group if and only if* $1 \in C$.

Proof. If $1 \in C$, then $C = 1^{-1}C$ is a group by Proposition 20(iv). The converse implication follows from the definition of a group. □

Remark 22. As an alternative to (5), Proposition 20(iii) yields the following characterization of cosets, which might appear more natural because it is closer to the original definition of cosets:

$$\exists g \in (\mathbb{R}^*)^n \; \forall x, y \in (\mathbb{R}^*)^n \colon gx \in V_{\mathbb{R}}(\hat{F}) \wedge gy \in V_{\mathbb{R}}(\hat{F}) \Rightarrow gx^{-1} \in V_{\mathbb{R}}(\hat{F}) \wedge gxy \in V_{\mathbb{R}}(\hat{F}). \quad (6)$$

The first difference to observe is that in (6) in contrast to (5) there is quantifier alternation from $\exists$ to $\forall$. The number of quantifier alternations is known to be a critical parameter for asymptotic complexity of the real quantifier elimination problem [37, 63]. Furthermore, in the presence of the leading existential quantifier prohibits our splitting into two independent smaller problems. Experimental computations on the complete dataset considered here have confirmed that formulation (5) is clearly preferable.

Example 23 (`BIOMD0000000159`). Consider $F \subseteq \mathbb{Q}[X]$, where $X = \{x_1, \ldots, x_3\}$:

$$F = \{-32x_1x_2 + 3, -x_2 + x_3, 4x_1 - x_3\}.$$

In lines 3–8 of Algorithm 4 we consecutively apply real quantifier elimination to the following formulas:

$$\forall x_1 \forall x_2 \forall x_3 (-32x_1x_2 + 3 = 0 \wedge -x_2 + x_3 = 0 \wedge 4x_1 - x_3 = 0 \longrightarrow x_i = 0), \quad i = 1, \ldots, 3.$$

Neither of them holds in $\mathbb{R}$ so that in line 9 we enter Algorithm 5 with $\hat{X} = X$ and $\hat{F} = F$.

In line 2 of Algorithm 5 we test

$$\exists x_1 \exists x_2 \exists x_3 (x_1 \neq 0 \wedge x_2 \neq 0 \wedge x_3 \neq 0 \wedge -32x_1x_2 + 3 = 0 \wedge -x_2 + x_3 = 0 \wedge 4x_1 - x_3 = 0). \quad (7)$$

Real quantifier elimination [47] confirms that (7) holds in $\mathbb{R}$, and extended quantifier elimination [48] even gives us a witness

$$(x_1, x_2, x_3) = \left(\tfrac{\sqrt{3}}{8\sqrt{2}}, \tfrac{\sqrt{3}}{2\sqrt{2}}, \tfrac{\sqrt{3}}{2\sqrt{2}}\right) \in V_{\mathbb{R}}(F)^*. \quad (8)$$

Therefore we set up τ_{inv} in line 5 as follows:

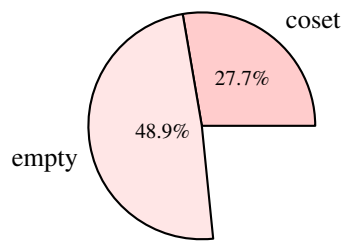
$$\begin{aligned}\tau_{\text{inv}} \doteq \forall g_1 \forall g_2 \forall g_3 \forall x_1 \forall x_2 \forall x_3 (g_1 = 0 \wedge g_2 = 0 \wedge g_3 = 0 \wedge x_1 = 0 \wedge x_2 = 0 \wedge x_3 = 0 \\ \wedge -32g_1g_2 + 3 = 0 \wedge -g_2 + g_3 = 0 \wedge 4g_1 - g_3 = 0 \\ \wedge -32g_1x_1g_2x_2 + 3 = 0 \wedge -g_2x_2 + g_3x_3 = 0 \wedge 4g_1x_1 - g_3x_3 = 0 \\ \longrightarrow -32g_1g_2 + 3x_1x_2 = 0 \wedge -g_2x_3 + g_3x_2 = 0 \wedge 4g_1x_3 - g_3x_1 = 0). \quad (9)\end{aligned}$$

Notice that in the three equations in last line of (9) we have equivalently dropped denominators x_1x_2, x_2x_3, and x_1x_3, respectively. The inequalities for $x_1, \ldots, x_3$ in first line of (9) ensure that those denominators do not vanish. In line 10, quantifier elimination confirms that τ_{inv} holds in $\mathbb{R}$, and so does τ_{mult} in line 10. Thus we reach line 13 and set up the following formula to test whether our coset $\hat{F}$ is even a group:

$$-32 + 3 = 0 \wedge -1 + 1 = 0 \wedge 4 - 1 = 0.$$

This is obviously not the case. Algorithm 5 returns '`C`' in line 17, and Algorithm 4 finally returns $\{x_1, x_2, x_3\}$ and '`C`' in line 13.

In Appendix A.1 we discuss practical aspects of our implementation and give in Table 1 classification results from applying Algorithm 4 to the 129 models introduced at the beginning of this section. For our discussion here we note that our algorithm terminates within a time limit of 6 hours per model on 94 out of the 129 models. We obtain 20 `C` and 6 `c`, which can be summarized as $V_{\mathbb{R}}(\tilde{G})^*$ forming a coset. Furthermore we have 4 `O` and 42 `o`, i.e., $V_{\mathbb{R}}(\tilde{G})^* = \varnothing$. The rest is 21 `X` and one single `x`. In terms of percentages of the 94 successful computations this gives the following picture:



In analogy to Algorithm 3 in Subsection 3.1, Algorithm 6 applies prime decompositions over $\mathbb{Q}$ also in the real case. It starts with the computation of a prime decomposition in line 1. We then have

$$\bigcup_{i=1}^{k} V_{\mathbb{R}}(P_i) = V_{\mathbb{R}}(F) \quad \text{and} \quad \mathbb{R} \models \bigvee_{i=1}^{k} \bigwedge_{p \in P_i} p = 0 \longleftrightarrow \bigwedge_{f \in F} f = 0.$$

In lines 3–4 we apply Algorithm 4 to each component and collect the results.

In Appendix A.2 we discuss practical aspects of our computations and give in Table 3 classification results using Algorithm 6 on the 129 models introduced at the beginning of this section. We succeed on 88 out of the 129 models within a time limit of 6 hours per model. This yields 3390 prime components to test altogether. We obtain 2 `G`, 22 `C`, and 1083 `c`, which can be summarized as $V_{\mathbb{R}}(\tilde{P}_i)^*$ forming a coset. Furthermore, we have 7 `O` and 2232 `o`, i.e., $V_{\mathbb{R}}(\tilde{P}_i)^* = \varnothing$. The rest is

Algorithm 6 $\mathrm{DecomposeProjectAndClassify}_{\mathbb{R}}$

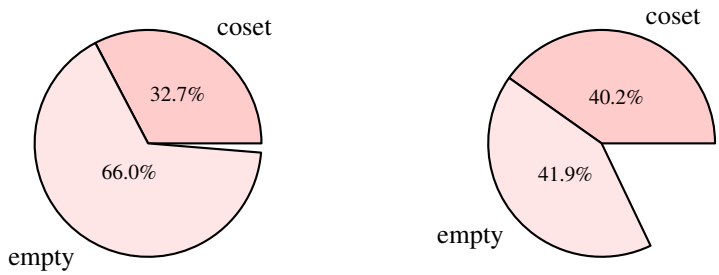
Input: 1. X, a finite ordered set of variables; 2. $F \subseteq \mathbb{Q}[X]$ finite and non-empty

Output: 1. $\mathcal{P} \in \wp(\mathbb{Q}[X])^k$; 2. $\hat{\mathcal{X}} \in \wp(X)^k$; 3. $\Gamma \in \{\mathtt{G}, \mathtt{C}, \mathtt{O}, \mathtt{X}, \mathtt{g}, \mathtt{c}, \mathtt{o}, \mathtt{x}\}^k$

$\mathcal{P} = (P_1, \ldots, P_k)$ are Gröbner bases of a prime decomposition over $\mathbb{Q}$ of $\langle F \rangle$.
In $\hat{\mathcal{X}} = (\hat{X}_1, \ldots, \hat{X}_k)$, X_i is the unique maximal subset of X with respect to $V_{\mathbb{R}}(P_i)$.
In $\Gamma = (\gamma_1, \ldots, \gamma_k)$, the letter γ_i classifies $V_{\mathbb{R}}(P_i)^*$ in $\hat{X}_i$-space, using upper case when $\hat{X}_i = X$:
$\mathtt{G/g}$ – $V_{\mathbb{R}}(P_i)^*$ is a group; $\mathtt{C/c}$ – $V_{\mathbb{R}}(P_i)^*$ is a proper coset; $\mathtt{O/o}$ – $V_{\mathbb{R}}(P_i)^* = \varnothing$; $\mathtt{X/x}$ else.

1: $\mathcal{P} = (P_1, \ldots, P_k) := \mathrm{PrimeDecomposition}_{\mathbb{Q}}(F)$
2: **for** $i = 1, \ldots, k$ **do**
3: $\quad \hat{X}_i, \gamma_i := \mathrm{ProjectAndClassify}_{\mathbb{R}}(X, P_i)$
4: **end for**
5: **return** $\mathcal{P}, (\hat{X}_1, \ldots, \hat{X}_k), (\gamma_1, \ldots, \gamma_k)$

only 18 `X` and 26 `x`. In left hand side picture below, we visualize these results in terms of percentages of the total of 3390 prime components. In the right hand side picture, we see the corresponding statistics without `BIOMD0000000281`:



Recall from the discussion at the end of Subsection 3.1 that we consider the left hand side picture more adequate and add the right hand side one for the sake of scientific rigor.

4. Upper Complexity Bounds

In this section we give asymptotic upper bounds on the worst case complexity of problems addressed in this paper. In Subsection 4.1 we derive bounds for recognizing toric and shifted toric varieties over algebraically closed fields of characteristic zero. In Subsection 4.2 we derive corresponding bounds for toric varieties over real closed fields. Subsection 4.3 finally gives bounds for the membership problem in subgroups of $(\overline{\mathbb{Q}}^*)^n$, which correspond to binomial varieties.

4.1. Toricity over Algebraically Closed Fields of Characteristic Zero

As mentioned in Section 2, a torus G can be represented as the set of solutions of binomials of the form $x_1^{a_{1i}} \cdots x_n^{a_{ni}} = 1$, $1 \leq i \leq n - m$, where $a_{ji} \in \mathbb{Z}$ and m is equal to the dimension of G and every Gröbner basis of a toric variety $\overline{G}$ consists only of binomials of the form $x_1^{b_{1i}} \cdots x_n^{b_{ni}} - x_1^{c_{1i}} \cdots x_n^{c_{ni}}$ where integers b_{ji}, c_{ji} are non-negative [58].

Let $f_1, \ldots, f_k \in \mathbb{Z}[x_1, \ldots, x_n]$ and $V \subseteq \mathbb{C}^n$ be the algebraic variety of common zeroes of $f_1, \ldots, f_k$. We design an algorithm which recognizes whether V is toric. Note that $f_1, \ldots, f_k$ are not necessary binomials. To estimate the complexity of the algorithm we suppose that $\deg(f_i) \leq d$, $1 \leq i \leq k$ and that the bit-size of each integer coefficient of $f_1, \ldots, f_k$ does not exceed L. Invoking [9, 36] we first verify whether V is irreducible and $\overline{V^*} = V$ (if this is not true then V is not toric).

The complexity of the algorithms from [9, 36] can be bounded by $(Ld^{n^2})^{O(1)}$. Then we verify that V^* is a group. This holds if and only if we have the following first-order formula in the theory of algebraically closed fields of characteristic zero:

$$\forall x \forall y (x \neq 0 \wedge y \neq 0 \wedge \bigwedge_i f_i(x) = 0 \wedge \bigwedge_i f_i(y) = 0 \longrightarrow \bigwedge_i f_i(xy) = 0 \wedge \bigwedge_i f_i(1/x) = 0). \quad (10)$$

This can be verified via the algorithm in [10]. The complexity of this step is bounded by $(Ld^{n^2})^{O(1)}$ as well.

So far, the algorithm has verified whether V is toric. Now we show how to find a system of binomial equations determining V^*. In order to do so, in this subsection we find a set of Laurent binomials determining V^* instead of binomials in $\mathbb{C}[X]$ whose set of solutions is V. This is because as it has been shown by Mayr and Meyer in their seminal work, the number of binomials and their degrees in a Gröbner basis of $I(V)$ can be double-exponential [49]. Using the algorithms from [9, 36] one can produce $m = \dim V$ coordinates among $x_1, \ldots, x_n$ which form a transcendental basis of V^*. Without loss of generality, assume that $\{x_1, \ldots, x_m\}$ be a transcendental basis. Fix an integer j, $m < j \leq n$ and project V^* on the space generated by $x_1, \ldots, x_m, x_j$, which is isomorphic to $(\mathbb{C}^*)^{m+1}$ invoking again [10]. Let $W \subseteq (\mathbb{C}^*)^{m+1}$ be image of the projection. Due to the choice of the transcendental basis, we have that $\dim W = m$ and W is a hypersurface. Therefore, W can be determined by a single polynomial $h := h_j \in \mathbb{Z}[x_1, \ldots, x_m, x_j]$ (c.f. e.g., [56]). Moreover, $\deg W \leq \deg V^* \leq d^n$; the latter follows from Bezout inequality [56]. The algorithm from [10] constructs h and a generic point of W within the complexity $(Ld^{n^2})^{O(1)}$.

Observe that W is also a group, hence h can be rewritten as a binomial of the form $x_1^{q_{j1}} \cdots x_m^{q_{jm}} \cdot x_j^{q_j} - 1 \in \mathbb{Z}[x_1^{\pm}, \ldots, x_n^{\pm}]$ for suitable relatively prime integers $q_{j1}, \ldots, q_{jm}, q_j \in \mathbb{Z}$. Doing so for every j, $m < j \leq n$, the algorithm yields polynomials $h_j \in \mathbb{Z}[x_1, \ldots, x_n]$. Denote by $H \subseteq (\mathbb{C}^*)^n$ the variety given by equations h_j, $m < j \leq n$. Clearly, $V^* \subseteq H$, $\dim H = m$ and H is a group, therefore, V^* is an irreducible component of H. Moreover, H is a binomial variety and hence, V^* is its subgroup (of a finite index). In particular, $(1, \ldots, 1) \in V^*$, and every irreducible component H_1 of H has the form $H_1 = gV^*$ for an arbitrary element $g \in H_1$. Moreover, one can choose g such that its coordinates are roots of unity [39, Remark 3.1, Remark 5.2].

In addition, in order to obtain the Laurent binomials defining V^*, the algorithm finds a $\mathbb{Z}$-basis of the intersection $\mathbb{Q}(Q_{m+1}, \ldots, Q_n) \cap \mathbb{Z}^n$ of the $\mathbb{Q}$-linear space generated by vectors $Q_j = (q_{j1}, \ldots, q_{jm}, 0, \ldots, 0, q_j, 0 \ldots, 0)$, $m < j \leq n$ with the grid $\mathbb{Z}^n$ [39, Remark 3.1]. To find that $\mathbb{Z}$-basis, the algorithm first applies [30, 21] to produce (within polynomial complexity) a $\mathbb{Z}$-basis Z of the space of integer solutions of the linear system with rows Q_j, $m < j \leq n$, and subsequently applies [30, 21] to construct a $\mathbb{Z}$-basis of the linear system with the rows from Z.

Recall from the Definition 4 that for a torus $V^* \subseteq (\mathbb{C}^*)^n$ and a point $g \in (\mathbb{C}^*)^n$ we call gV^* a shifted torus, and $\overline{gV^*} \subseteq \mathbb{C}^n$ a shifted toric variety. In particular, every irreducible component of a binomial variety in $(\mathbb{C}^*)^n$ is a shifted torus (Proposition 3). One can modify the described algorithm to test whether an input variety $V \subseteq \mathbb{C}^n$ is shifted toric. To this end, pick an arbitrary point $g = (g_1, \ldots, g_n) \in V^*$ and test whether $g^{-1}V^*$ is a torus. If the latter holds, the algorithm produces binomial equations for $g^{-1}V^*$. Clearly a binomial equation $x_1^{s_1} \cdots x_n^{s_n} = 1$ vanishes on $g^{-1}V^*$ if and only if $x_1^{s_1} \cdots x_n^{s_n} = g_1^{s_1} \cdots g_n^{s_n}$ vanishes on V^* .

The following theorem summarizes our algorithm.

Theorem 24. *Let $f_1, \ldots, f_k \in \mathbb{Z}[x_1, \ldots, x_n]$, $\deg(f_i) \leq d$, $1 \leq i \leq k$ with bit-sizes of integer coefficients of $f_1, \ldots, f_k$ at most L. One can design an algorithm which tests whether the variety $V \subseteq \mathbb{C}^n$ determined by $f_1, \ldots, f_k$ is (shifted) toric. In the positive case, the algorithm yields a transcendental basis $x_{i_1}, \ldots, x_{i_m}$, $m = \dim V$ of V. It furthermore yields binomial equations defining $V^* = V \cap (\mathbb{C}^*)^n$. Each binomial equation has the form $x_1^{s_1} \cdots x_n^{s_n} = g_1^{s_1} \cdots g_n^{s_n}$ for all $(g_1, \ldots, g_n) \in V^*$ and for suitable integers $s_1, \ldots, s_n \in \mathbb{Z}$ satisfying $|s_i| < O(d^{n^2})$. The complexity of the designed algorithm does not exceed $(Ld^{n^2})^{O(1)}$.*

One can extend the algorithm in Theorem 24 so that it takes a reducible variety V and decomposes V into irreducible components and then tests whether its irreducible components are toric or shifted toric (following the lines of the algorithm for each irreducible component of V^* separately). Moreover, if the irreducible components of V are exclusively toric and shifted toric varieties and V^* is a group, then the extended algorithm yields the described representation of V^* by means of binomials. The complexity of the extended algorithm is still $(Ld^{n^2})^{O(1)}$.

Our discussion here can be straightforwardly generalized to any algebraically closed field with characteristic zero. Independently, the algorithm can be generalized to coefficients from a finite field extension of $\mathbb{Q}$ [9, 36].

4.2. Toricity over Real Closed Fields

In this subsection we design an algorithm that recognizes toricity of a semi-algebraic set over $\mathbb{R}$. We refer to [2] for the algorithms in real algebraic geometry. Let $\mathbb{R}_{>0} := \{ z \in \mathbb{R} \mid z > 0 \}$ denote the positive orthant. Keeping the notations from Subsection 4.1 consider the semi-algebraic set $T := \{ x \in (\mathbb{R}_{>0})^n \mid f_i(x) \geq 0, 1 \leq i \leq k \}$. Modify (10) replacing $\mathbb{C}^*$ by $\mathbb{R}_{>0}$ and equalities $f_i = 0$ by inequalities $f_i \geq 0$, respectively. Also keep the first-order formula (10). This formula can be verified by applying the algorithms from [38]. Clearly, (10) is true if and only if T is a group (a torus). Thus, assume that (10) is true. Then the image of the coordinate-wise logarithm map $\log(T) \subseteq \mathbb{R}^n$ is a linear subspace. Hence, in particular, T is connected.

First, compute $m := \dim(T)$ and produce m coordinates such that the projection U of T on m-dimensional space with these coordinates has the full dimension m. Without loss of generality assume that these coordinates are $x_1, \ldots, x_m$. Fix $j, m < j \leq n$ and denote by U_j the projection of T on the $(m+1)$-dimensional space with coordinates $x_1, \ldots, x_m, x_j$. Denote by p the projection map of the latter space along x_j onto m-dimensional space with the coordinates $x_1, \ldots, x_m$. Then $p(U_j) = U$. Since $\dim(U_j) = \dim(U) = m$, we have that $p(\log(U_j)) = \log(U) = \mathbb{R}^m$, and therefore we conclude that any point of U has a unique preimage of p in U_j. Moreover, U_j is determined by a single binomial-type (analytic) equation of the form

$$x_1^{t_{j1}} \cdots x_m^{t_{jm}} x_j^{t_j} = c_j \tag{11}$$

for some reals $t_{j1}, \ldots, t_{jm}, t_j, c_j$. Since U_j is a group we get that $c_j = 1$.

On the other hand, applying the algorithm from [2], one can construct the projection $p : U_j \to U$ and conclude that equation (11) is algebraic, thus $t_{j1}, \ldots, t_{jm}, t_j \in \mathbb{Z}$. The algorithm also yields $t_{j1}, \ldots, t_{jm}, t_j$. Note that without loss of generality one can assume that they are relatively prime, otherwise divide by their greatest common divisor. The complexity of the algorithm is again $(Ld^{n^2})^{O(1)}$. Doing so for each j, $m < j \leq n$, the algorithm yields binomial equations of the form (11) which determine T uniquely. Similar to subsection 4.1 one can produce a $\mathbb{Z}$-basis of the intersection of $\mathbb{Q}$-linear space generated by vectors $(t_{j1}, \ldots, t_{jm}, 0, \ldots, 0, t_j, 0 \ldots, 0)$ with the grid $\mathbb{Z}^n$. Then any vector $(s_1, \ldots, s_n)$ from this basis provides a binomial $X_1^{s_1} \cdots X_n^{s_n} - 1$ that vanishes on T. The above $\mathbb{Z}$-basis need not be constructed, since binomials of the form (11) already determine T uniquely. We summarize the described algorithm in the following theorem.

Theorem 25. *Let $f_1, \ldots, f_k \in \mathbb{Z}[x_1, \ldots, x_n]$, $\deg(f_i) \leq d$, $1 \leq i \leq k$ with bit-sizes of integer coefficients of $f_1, \ldots, f_k$ at most L. One can design an algorithm which tests whether the semi-algebraic set $T := \{ x \in (\mathbb{R}_{>0})^n \mid f_i(x) \geq 0, 1 \leq i \leq k \}$ is a group (a torus). In the positive case, the algorithm yields coordinates $x_{i_1}, \ldots, x_{i_m}$, where $m = \dim(T)$, such that the dimension of the projection of T on the m-dimensional space with the coordinates $x_{i_1}, \ldots, x_{i_m}$ equals m. It furthermore yields for each $j \notin \{i_1, \ldots, i_m\}$ a binomial equation of the form $x_1^{t_{j1}} \cdots x_m^{t_{jm}} \cdot x_j^{t_j} = 1$ which vanishes on T, with relatively prime $t_{j1}, \ldots, t_{jm}, t_j \in \mathbb{Z}$, where $|t_{j1}| + \cdots + |t_{jm}| + |t_j| \leq d^{O(n)}$. The complexity of the algorithm does not exceed $(Ld^{n^2})^{O(1)}$.*

Similar to Subsection 4.1, our results here can be generalized to arbitrary real closed field.

4.3. Membership in Binomial Varieties

Let a group $G \subset (\overline{\mathbb{Q}}^*)^n$ be given by binomial equations $x_1^{a_{i,1}} \cdots x_n^{a_{i,n}} = 1$, $1 \leq i \leq k$, where $a_{i,j} \in \mathbb{Z}$, $|a_{i,j}| \leq d$. Let $v = (v_1, \ldots, v_n) \in \mathbb{Q}^n$ be a point such that the absolute values of the numerators and denominators of $v_1, \ldots, v_n$ do not exceed M. We design an algorithm which tests whether $v \in \overline{G}$. Recall that $\overline{G}$ is a binomial variety, and it is a toric variety when G is irreducible.

Theorem 26. *There is an algorithm which tests whether a point v belongs to a binomial variety $\overline{G}$ with its complexity bounded by*

(i) $(k \cdot \log M \cdot (dn)^n)^{O(1)}$ *and by*
(ii) $(k \cdot M \cdot n \cdot \log d)^{O(1)}$.

Proof. Permuting the coordinates, assume without loss of generality that $v = (0, \ldots, 0, v_{s+1}, \ldots, v_n)$, where $v_{s+1} \cdots v_n \neq 0$. Due to Claim 5.3 in [39], $v \in \overline{G}$ if and only if there exist a point $u = (u_1, \ldots, u_s, v_{s+1}, \ldots, v_n) \in G$ and positive $0 < b_1, \ldots, b_s \in \mathbb{Z}$ such that

$$v = \lim_{t \to 0} (u_1 \cdot t^{b_1}, \ldots, u_s \cdot t^{b_s}, v_{s+1}, \ldots, v_n) \tag{12}$$

and for all $t \neq 0$, i.e., a shift of a one-parametric subgroup, we have that $(u_1 \cdot t^{b_1}, \ldots, u_s \cdot t^{b_s}, v_{s+1}, \ldots, v_n) \in G$. Then (12) is equivalent to the existence of $u \in G$ and a one-parametric subgroup $\{ (t^{b_1}, \ldots, t^{b_s}, 1, \ldots, 1) \mid t \neq 0 \} \subseteq G$. The existence of the latter is equivalent to the existence of a non-negative vector $(b_1, \ldots, b_s, 0, \ldots, 0)$ orthogonal to the vectors $(a_{i,1}, \ldots, a_{i,n})$, $1 \leq i \leq k$. This can be checked by means of linear programming.

The existence of a point $u \in G$ satisfying (12) is equivalent to the existence of non-zero u_1, $\ldots$, $u_s \in \overline{\mathbb{Q}}^*$ satisfying the binomial equations

$$u_1^{a_{i,1}} \cdots u_s^{a_{i,s}} = v_{s+1}^{-a_{i,s+1}} \cdots v_n^{-a_{i,n}}, \quad 1 \leq i \leq k. \tag{13}$$

One can apply the algorithm in [40] to this system and transform the $k \times s$ submatrix $A = (a_{i,j})$, $1 \leq i \leq k$, $1 \leq j \leq s$ to its Smith form. Then solvability of (13) is equivalent to that the right-hand side of (13) fulfils (at most k) relations of the form

$$\prod_{1 \leq i \leq k} (v_{s+1}^{-a_{i,s+1}} \cdots v_n^{-a_{i,n}})^{c_i} = 1 \tag{14}$$

for some integers c_i being suitable minors of matrix A, hence $|c_i| \leq (ds)^{O(s)}$ by Hadamard's inequality.

One can verify relations (14) using the binary form of the numerators and denominators of $v_{s+1}, \ldots, v_n$. This leads to the complexity bound (i). Alternatively, one can factorize the numerators and denominators of $v_{s+1}, \ldots, v_n$ and execute calculations in terms of exponents of their prime factors which leads to the complexity bound (ii). This completes the verification of (12) and the description of the algorithm. □

One can extend the complexity bound (i) for $v_j \in \overline{\mathbb{Q}}$, $1 \leq j \leq n$ being algebraic numbers. In this case $\log M$ plays the role of the bit-size of the representation of v.

5. Conclusions and Future Work

We have taken a geometric approach to studying steady state varieties, which—besides significant theoretical results—generated comprehensive empirical data from computations on 129 networks from BioModels repository. We are not aware of any comparable systematic large scale symbolic computations on those data in the literature. We were indeed surprised by the success rate of Gröbner basis and real quantifier elimination techniques with input sizes up to 71 variables. We find this most

encouraging and believe that robust and supported software tools for systems biology and medicine that include symbolic computation components are not out of reach.

It was important to learn that real methods do not significantly fall behind complex methods efficiency-wise. After all, chemical reaction network theory takes place in the interior of the positive first orthant. In a way, our consideration of V^* in favor of V here marks a first step in that direction by looking at points in the variety rather than polynomials in the ideal. Only real methods allow to go further and, e.g., identify prime components whose varieties reach into the first orthant.

Our work here gives a number of quite concrete challenges to be considered in subsequent work. To start with, one must complete the step from "coset" to "shifted toric" in practical software by producing suitable prime decompositions beyond decompositions over $\mathbb{Q}$. As mentioned above, real methods allow to explicitly refer to the interior of the first orthant, and our framework and the first-order descriptions we use should be refined in this direction. On the complex side, one should also test suitable elimination methods with the logic descriptions we developed here.

It is noteworthy that all quantifier elimination problems considered here were decision problems, even without quantifier alternation. On the one hand, this allows the application of methods and tools from Satisfiability Modulo Theories Solving [51, 1]. On the other hand, it shows that we are not yet using the full power of quantifier elimination methods. One could, e.g., leave a subset of reaction rates parametric and study invariance of shifted toricity under variation of those reaction rates.

Our input of 129 models considered here is in a way complete: We took all available models from BioModels/ODEbase for which we could straightforwardly produce polynomial vector fields. So far we did not consider systems with rational vector fields. Such systems come with interesting challenges on the algebraic side: While the variety V is blind for the presence or non-presence of polynomials factors from the denominators, those factors can still affect shifted toricity V^*.

References

[1] Erika Ábrahám, John Abbott, Bernd Becker, Anna M. Bigatti, Martin Brain, Bruno Buchberger, Alessandro Cimatti, James H. Davenport, Matthew England, Pascal Fontaine, Stephen Forrest, Alberto Griggio, Daniel Kroening, Werner M. Seiler, and Thomas Sturm. SC^2: Satisfiability checking meets symbolic computation. In Michael Kohlhase, Moa Johansson, Bruce Miller, Leonardo de Moura, and Frank Tompa, editors, *CICM 2016*, volume 9791 of *LNCS*, pages 28–43. Springer, 2016.

[2] Saugata Basu, Richard Pollack, and Marie-Francoise Roy. *Algorithms in Real Algebraic Geometry*, volume 10 of *Algorithms and Computation in Mathematics*. Springer, 2006.

[3] Thomas Becker, Volker Weispfenning, and Heinz Kredel. *Gröbner Bases, a Computational Approach to Commutative Algebra*, volume 141 of *Graduate Texts in Mathematics*. Springer, 1993.

[4] Ludwig Boltzmann. *Lectures on Gas Theory*. University of California Press, Berkeley and Los Angeles, CA, 1964.

[5] F. Boulier, F. Fages, O. Radulescu, S. Samal, A. Schuppert, W. Seiler, T. Sturm, S. Walcher, and A. Weber. The SYMBIONT project: Symbolic methods for biological networks. *ACM Communications in Computer Algebra*, 52(3):67–70, December 2018.

[6] F. Boulier, F. Fages, O. Radulescu, S. Samal, A. Schuppert, W. Seiler, T. Sturm, S. Walcher, and A. Weber. The SYMBIONT project: Symbolic methods for biological networks. *F1000Research*, 7(1341), 2018.

[7] Bruno Buchberger. *Ein Algorithmus zum Auffinden der Basiselemente des Restklassenringes nach einem nulldimensionalen Polynomideal*. Doctoral dissertation, Mathematical Institute, University of Innsbruck, Innsbruck, Austria, 1965.

[8] Vijayalakshmi Chelliah, Nick Juty, Ishan Ajmera, Raza Ali, Marine Dumousseau, Mihai Glont, Michael Hucka, Gaël Jalowicki, Sarah Keating, Vincent Knight-Schrijver, Audald Lloret-Villas, Kedar Nath Natarajan, Jean-Baptiste Pettit, Nicolas Rodriguez, Michael Schubert, Sarala M. Wimalaratne,

Yangyang Zhao, Henning Hermjakob, Nicolas Le Novère, and Camille Laibe. BioModels: Ten-year anniversary. *Nucl. Acids Res.*, 2015.

[9] A. L. Chistov. Algorithm of polynomial complexity for factoring polynomials and finding the components of varieties in subexponential time. *Journal of Soviet Mathematics*, 34(4):1838–1882, August 1986.

[10] A. L. Chistov and D. Yu. Grigoriev. Complexity of quantifier elimination in the theory of algebraically closed fields. In *Mathematical Foundations of Computer Science*, volume 176 of *LNCS*, pages 17–31. Springer, 1984.

[11] George E. Collins. Quantifier elimination for the elementary theory of real closed fields by cylindrical algebraic decomposition. In H. Brakhage, editor, *Automata Theory and Formal Languages. 2nd GI Conference*, volume 33 of *LNCS*, pages 134–183. Springer, 1975.

[12] George E. Collins and Hoon Hong. Partial cylindrical algebraic decomposition for quantifier elimination. *J. Symb. Comput.*, 12(3):299–328, September 1991.

[13] Carsten Conradi, Alexandru Iosif, and Thomas Kahle. Multistationarity in the space of total concentrations for systems that admit a monomial parametrization. *Bulletin of Mathematical Biology*, Jul 2019.

[14] Carsten Conradi and Thomas Kahle. Detecting binomiality. *Advances in Applied Mathematics*, 71:52–67, 2015.

[15] David Cox, John Little, and Donald O'Shea. *Ideals, Varieties and Algorithms*. Undergraduate Texts in Mathematics. Springer, 1992.

[16] Gheorghe Craciun, Alicia Dickenstein, Anne Shiu, and Bernd Sturmfels. Toric dynamical systems. *J. Symb. Comput.*, 44(11):1551–1565, November 2009.

[17] James H. Davenport and Joos Heintz. Real quantifier elimination is doubly exponential. *J. Symb. Comput.*, 5(1–2):29–35, February–April 1988.

[18] Andreas Dolzmann, Andreas Seidl, and Thomas Sturm. Efficient projection orders for CAD. In *Proceedings of the ISSAC 2004*, pages 111–118. ACM Press, New York, NY, 2004.

[19] Andreas Dolzmann and Thomas Sturm. Redlog: Computer algebra meets computer logic. *ACM SIGSAM Bulletin*, 31(2):2–9, June 1997.

[20] Andreas Dolzmann and Thomas Sturm. Simplification of quantifier-free formulae over ordered fields. *J. Symb. Comput.*, 24(2):209–231, August 1997.

[21] Jean-Guillaume Dumas, B. David Saunders, and Gilles Villard. On efficient sparse integer matrix Smith normal form computations. *J. Symb. Comput.*, 32(1–2):71–99, July 2001.

[22] Albert Einstein. Strahlungs-emission und -absorption nach der Quantentheorie. *Verh. Dtsch. Phys. Ges.*, 18:318–323, 1916.

[23] David Eisenbud and Bernd Sturmfels. Binomial ideals. *Duke Math. J.*, 84(1):1–45, 1996.

[24] Jean-Charles Faugère. A new efficient algorithm for computing Gröbner bases (F4). *J. Pure Appl. Algebra*, 139(1):61–88, 1999.

[25] Jean-Charles Faugère. A new efficient algorithm for computing gröbner bases without reduction to zero (F5). In *Proceedings of the ISSAC 2002*, pages 75–83. ACM, New York, NY, 2002.

[26] Martin Feinberg. Complex balancing in general kinetic systems. *Arch. Ration. Mech. An.*, 49(3):187–194, January 1972.

[27] Martin Feinberg. Stability of complex isothermal reactors—I. The deficiency zero and deficiency one theorems. *Chem. Eng. Sci.*, 42(10):2229–2268, 1987.

[28] Martin Feinberg. *Foundations of Chemical Reaction Network Theory*, volume 202 of *Applied Mathematical Sciences*. Springer, 2019.

[29] A. Finney and M. Hucka. Systems biology markup language: Level 2 and beyond. *Biochemical Society Transactions*, 31(6):1472–1473, December 2003.

[30] M. A. Frumkin. An application of modular arithmetic to the construction of algorithms for solving systems of linear equations. *Soviet Math. Dokl*, 229(5):1067–1070, 1976.

[31] William Fulton. *Introduction to Toric Varieties*, volume 131 of *Annals of Mathematics Studies*. Princeton University Press, 1993.

[32] Karin Gatermann. Counting stable solutions of sparse polynomial systems in chemistry. In *Symbolic Computation: Solving Equations in Algebra, Geometry, and Engineering*, volume 286 of *Contemporary Mathematics*, pages 53–69. AMS, Providence, RI, 2001.

[33] Karin Gatermann and Matthias Wolfrum. Bernstein's second theorem and Viro's method for sparse polynomial systems in chemistry. *Adv. Appl. Math.*, 34(2):252–294, February 2005.

[34] A. N Gorban, E. M. Mirkes, and G. S. Yablonski. Thermodynamics in the limit of irreversible reactions. *Physica A*, 392:1318–1335, 2013.

[35] A. N Gorban and G. S. Yablonski. Extended detailed balance for systems with irreversible reactions. *Chem. Eng. Sci.*, 66:5388–5399, 2011.

[36] D. Yu. Grigoriev. Factorization of polynomials over a finite field and the solution of systems of algebraic equations. *Journal of Soviet Mathematics*, 34(4):1762–1803, August 1986.

[37] D. Yu. Grigoriev. Complexity of deciding Tarski algebra. *J. Symb. Comput.*, 5(1–2):65–108, February–April 1988.

[38] D. Yu. Grigoriev and N. N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *J. Symb. Comput.*, 5(1-2):37–64, February–April 1988.

[39] Dima Grigoriev and Pierre D. Milman. Nash resolution for binomial varieties as Euclidean division. A priori termination bound, polynomial complexity in essential dimension 2. *Adv. Math.*, 231(6):3389–3428, 2012.

[40] Dima Grigoriev and Andreas Weber. Complexity of solving systems with few independent monomials and applications to mass-action kinetics. In *Proceedings of the CASC 2012*, volume 7442 of *LNCS*, pages 143–154. Springer, 2012.

[41] Anthony C. Hearn. Reduce—A user-oriented system for algebraic simplification. *ACM SIGSAM Bulletin*, 1(6):50–51, May 1967.

[42] Anthony C. Hearn. Reduce: The first forty years. In *Proceedings of the A3L 2005*. BOD, Norderstedt, Germany, 2005.

[43] F. Horn and R. Jackson. General mass action kinetics. *Arch. Ration. Mech. An.*, 47(2):81–116, January 1972.

[44] M. Hucka, A. Finney, H. M. Sauro, H. Bolouri, J. C. Doyle, H. Kitano, A. P. Arkin, B. J. Bornstein, D. Bray, A. Cornish-Bowden, A. A. Cuellar, S. Dronov, E. D. Gilles, M. Ginkel, V. Gor, I. I. Goryanin, W. J. Hedley, T. C. Hodgman, J.-H. Hofmeyr, P. J. Hunter, N. S. Juty, J. L. Kasberger, A. Kremling, U. Kummer, N. Le Novère, L. M. Loew, D. Lucio, P. Mendes, E. Minch, E. D. Mjolsness, Y. Nakayama, M. R. Nelson, P. F. Nielsen, T. Sakurada, J. C. Schaff, B. E. Shapiro, T. S. Shimizu, H. D. Spence, J. Stelling, K. Takahashi, M. Tomita, J. Wagner, and J. Wang. The systems biology markup language (SBML): a medium for representation and exchange of biochemical network models. *Bioinformatics*, 19(4):524–531, March 2003.

[45] Thomas Kahle. Decompositions of binomial ideals. *Ann. I. Stat. Math.*, 62(4):727–745, August 2010.

[46] Thomas Kahle. Decompositions of binomial ideals. *Journal of Software for Algebra and Geometry*, 4(1):1–5, 2012.

[47] Marek Košta. *New Concepts for Real Quantifier Elimination by Virtual Substitution*. Doctoral dissertation, Saarland University, Germany, December 2016.

[48] Marek Košta, Thomas Sturm, and Andreas Dolzmann. Better answers to real questions. *J. Symb. Comput.*, 74:255–275, May–June 2016.

[49] Ernst W. Mayr and Albert R. Meyer. The complexity of the word problems for commutative semigroups and polynomial ideals. *Adv. Math.*, 46(3):305–329, 1982.

[50] Stefan Müller, Elisenda Feliu, Georg Regensburger, Carsten Conradi, Anne Shiu, and Alicia Dickenstein. Sign conditions for injectivity of generalized polynomial maps with applications to chemical reaction networks and real algebraic geometry. *Found. Comput. Math.*, 16(1):69–97, February 2016.

[51] Robert Nieuwenhuis, Albert Oliveras, and Cesare Tinelli. Solving SAT and SAT modulo theories: From an abstract Davis–Putnam–Logemann–Loveland procedure to DPLL(T). *Journal of the ACM*, 53(6):937–977, 2006.

[52] Lars Onsager. Reciprocal relations in irreversible processes. I. *Phys. Rev.*, 37(4):405, 1931.

[53] Mercedes Pérez Millán and Alicia Dickenstein. The structure of MESSI biological systems. *SIAM J. Appl. Dyn. Syst.*, 17(2):1650–1682, 2018.

[54] Mercedes Pérez Millán, Alicia Dickenstein, Anne Shiu, and Carsten Conradi. Chemical reaction systems with toric steady states. *Bull. Math. Biol.*, 74(5):1027–1065, October 2012.

[55] Andreas Seidl. *Cylindrical Decomposition Under Application-Oriented Paradigms*. PhD thesis, University of Passau, Germany, March 2006.

[56] Igor R. Shafarevich. *Basic Algebraic Geometry*. Springer, 1974.

[57] Thomas Sturm. Redlog online resources for applied quantifier elimination. *Acta Acad. Ab.*, 67(2):177–191, 2007.

[58] Bernd Sturmfels. *Gröbner bases and convex polytopes*, volume 8 of *University Lecture Series*. AMS, Providence, RI, 1996.

[59] Ole Tange. GNU Parallel: The command-line power tool. *login: The USENIX Magazine*, 36(1):42–47, February 2011.

[60] Alfred Tarski. A decision method for elementary algebra and geometry. Prepared for publication by J. C. C. McKinsey. RAND Report R109, August 1, 1948, Revised May 1951, Second Edition, RAND, Santa Monica, CA, 1957.

[61] Shu Wang, Jia-Ren Lin, Eduardo D. Sontag, and Peter K. Sorger. Inferring reaction network structure from single-cell, multiplex data, using toric systems theory. *bioRxiv*, page 731018, 2019.

[62] Rudolf Wegscheider. Über simultane Gleichgewichte und die Beziehungen zwischen Thermodynamik und Reactionskinetik homogener Systeme. *Monatsh. Chem. Verw. Tl.*, 22(8):849–906, August 1901.

[63] Volker Weispfenning. The complexity of linear problems in fields. *J. Symb. Comput.*, 5(1–2):3–27, February–April 1988.

[64] Volker Weispfenning. Quantifier elimination for real algebra—The quadratic case and beyond. *Appl. Algebr. Eng. Comm.*, 8(2):85–101, February 1997.

Acknowledgments

This work has been supported by the bilateral project ANR-17-CE40-0036/DFG-391322026 SYMBIONT [5, 6]. The first author is grateful to grant RSF-16-11-10075 and to MCCME for the inspiring atmosphere. Stephen Forrest at Maplesoft and Arthur Norman at the University of Cambridge greatly helped with Maple- and Reduce-related technical issues, respectively. Christoph Lüders at the University of Bonn greatly supported the project with discussions around and generation of suitable input data.

Appendix A. Computations on 129 Models from the BioModels Repository

We conducted our computations on a 2.40 GHz Intel Xeon E5-4640 with 512 GB RAM and 32 physical cores providing 64 CPUs via hyper-threading. For parallelization of the jobs for the individual models we used GNU Parallel [59]. Results are stored in an Sqlite3 database file. That database contains significantly more information than can be presented here in print.[6] Beyond our own computations the database imports data from BioModels and ODEbase, so that our models and the history of our input can be reliably tracked. For instance, we store the mappings between our variables and the original species names and even the original SBML file with each model.

Among the information imported from ODEbase there is a binary flag indicating whether or not a model has mass action kinetics according to the following criterion [28, Section 2.1.2].

Definition 27 (Mass Action Test). A system is considered a mass action system when the kinetic law is made up of the product of the concentrations of the reactant species to the power of their respective stoichiometry times a constant

[6] It is planned to make the database available with the publication of this article.

A.1. Classifications of the Original Systems

We start with the classification over $\mathbb{C}$ and $\mathbb{R}$ of the original, not decomposed, systems using algorithms from Subsection 3.1 and Subsection 3.2, respectively. For comments on the tables from a theoretical point of view compare those sections. Beyond the data presented here we generally save the sets X, $\hat{X}$. Furthermore, over $\mathbb{C}$ we save the computed Gröbner bases G, $\hat{G}$, $\tilde{G}$ and their term orders, and over $\mathbb{R}$ we save witnesses for $V_{\mathbb{R}}^* \neq \varnothing$.

TABLE 1: Applying $\mathrm{ProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$. Model numbers *nnn* stand for `BIOMD0000000`*nnn*. "m/a" indicates mass action kinetics (Definition 27). $|X|$ and $|\hat{X}|$ are numbers of variables before and after projection, respectively. γ is `G` for group, `C` fo coset, `O` for empty set, and `X` else; lower case letters indicate projection. Time columns give total CPU times in seconds or "$\perp$" for a timeout with a limit of 6 hours per model.

		Algorithm 1 ($\mathbb{C}$)				Algorithm 4 ($\mathbb{R}$)			
model	m/a	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)
001	1	12	12	C	10.32	12	12	C	3.37
002	1				$\perp$				$\perp$
009	0	22	22	C	85.96	22	22	C	21.26
011	0	22	22	C	150.21	22	22	C	11.20
026	1	11	11	X	3.99	11	11	X	1.12
028	1	16	16	X	66.60				$\perp$
030	1	18	18	X	57.84				$\perp$
035	0	9	9	X	21.11	9	9	X	0.25
038	0				$\perp$				$\perp$
040	0	3	3	X	4.65	3	3	X	0.05
046	0				$\perp$				$\perp$
050	0	9	0	o	0.09	9	0	o	0.02
052	0	6	0	o	0.09	6	0	o	0.01
057	0	6	6	C	1.53	6	6	C	0.11
069	0	10	10	X	14.09				$\perp$
072	0	7	3	o	0.63	7	3	o	0.03
077	0	7	7	C	5.82	7	7	C	0.14
080	0	10	8	o	3.72	10	8	o	0.09
082	0	10	8	o	2.44	10	8	o	0.14
085	0				$\perp$				$\perp$
086	0				$\perp$				$\perp$
091	0	14	0	o	0.05	14	0	o	0.03
092	0	3	3	C	1.87	3	3	C	0.04
099	0	7	7	C	2.48	7	7	C	0.34
101	1	6	6	X	1.82	6	6	X	0.09
102	0	13	13	X	226.67				$\perp$
103	0	17	17	X	20238.89				$\perp$
104	0	4	4	O	1.36	4	4	O	0.01
105	0	26	0	o	0.24	26	0	o	0.51
108	0				$\perp$				$\perp$
122	0	12	12	X	8639.19				$\perp$
123	0				$\perp$				$\perp$
125	0	5	5	X	2.26	5	5	X	0.05
137	0	21	20	x	50.16	21	20	x	2.91

TABLE 1: Applying $\mathrm{ProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$ (continued)

		Algorithm 1 ($\mathbb{C}$)				Algorithm 4 ($\mathbb{R}$)			
model	m/a	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)
147	0				$\perp$				$\perp$
150	0	4	4	C	3.43	4	4	C	0.04
152	0				$\perp$				$\perp$
153	0				$\perp$				$\perp$
156	0	3	3	C	0.96	3	3	C	0.02
158	0	3	3	X	1.00	3	3	X	0.02
159	0	3	3	C	2.49	3	3	C	0.02
163	0	16	16	X	9.09	16	16	X	1.75
173	1				$\perp$				$\perp$
178	0	4	0	o	0.06	4	0	o	0.00
186	1	10	10	O	4.63	10	10	O	0.14
187	1	10	10	O	9.85	10	10	O	0.12
188	1	10	0	o	0.10	10	0	o	0.02
189	1	7	0	o	0.03	7	0	o	0.01
193	1	8	8	X	7.11	8	8	X	0.13
194	1	5	5	X	7.07	5	5	X	0.03
197	0	5	5	X	9.93	5	5	X	0.15
198	1	9	5	c	0.79	9	5	c	0.06
199	0	8	8	C	2.79	8	8	C	0.18
200	0				$\perp$				$\perp$
205	0				$\perp$				$\perp$
220	0	56	46	o	5.27	56	46	o	88.15
226	0	14	14	X	14.19				$\perp$
227	0	39	0	o	0.14	39	0	o	0.19
229	0	7	7	C	1.33	7	7	C	0.20
230	0	24	23	o	47.72				$\perp$
233	0	2	2	X	1.03	2	2	X	0.01
243	0	19	12	o	1.32	19	11	o	5.21
257	1	8	8	X	2.65				$\perp$
259	0	16	0	o	0.38	16	0	o	0.17
260	0	16	0	o	0.09	16	0	o	0.16
261	0	16	0	o	0.09	16	0	o	0.17
262	0	9	1	c	3.18	9	1	c	0.05
263	0	9	1	c	1.11	9	1	c	0.05
264	0	11	9	c	11.45	11	9	c	0.37
267	0	3	0	o	0.09	3	0	o	0.00
270	0				$\perp$				$\perp$
271	1	4	1	c	0.20	4	1	c	0.01
272	1	4	1	c	0.27	4	1	c	0.01
281	0	32	31	o	1.20	32	31	o	24.50
282	1	3	3	O	0.16	3	3	O	0.01
283	1	3	2	o	0.28	3	2	o	0.01
286	0				$\perp$				$\perp$
287	1	20	20	X	33.76				$\perp$
289	0	4	4	X	0.84	4	0	o	0.01
292	0	2	0	o	0.16	2	0	o	0.01

TABLE 1: Applying $\mathrm{ProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$ (continued)

		Algorithm 1 ($\mathbb{C}$)				Algorithm 4 ($\mathbb{R}$)			
model	m/a	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)
306	0	2	2	C	0.25	2	2	C	0.01
307	0	2	0	o	0.14	2	0	o	0.01
310	0	1	0	o	0.02	1	0	o	0.00
311	0	1	0	o	0.02	1	0	o	0.00
312	0	2	0	o	0.02	2	0	o	0.00
314	0	10	7	o	0.73	10	7	o	0.07
315	1				$\perp$				$\perp$
321	0	3	0	o	0.02	3	0	o	0.00
332	0				$\perp$	70	64	o	395.97
333	0	49	43	o	573.10	49	43	o	120.98
334	0				$\perp$	69	63	o	271.40
335	1	29	28	o	97.52				$\perp$
344	0				$\perp$				$\perp$
357	1	8	4	o	0.20	8	4	o	0.10
359	0	8	6	o	0.21	8	6	o	0.08
360	0	8	6	o	0.93	8	6	o	0.07
361	0	8	7	o	1.04	8	7	o	0.04
362	1	29	28	o	1304.71	29	28	o	429.42
363	0	3	0	o	0.07	3	0	o	0.00
364	1	12	10	o	1.03	12	10	o	0.54
365	1	30	24	o	212.01				$\perp$
407	0				$\perp$				$\perp$
413	1	5	5	X	1.05	5	5	X	0.06
416	0	32	32	X	25.87	32	32	X	4.17
430	0				$\perp$	23	23	X	13.99
431	0				$\perp$	27	27	X	31.90
439	0	20	20	X	52.51	20	20	X	15.42
459	0	3	3	C	0.50	3	3	C	0.09
460	0	3	3	X	0.52	3	3	X	0.04
475	0	22	20	o	3.26	22	20	o	7.92
478	0	29	28	o	33.38	29	28	o	16.85
479	0				$\perp$				$\perp$
483	0	6	6	X	1.26	6	6	X	0.11
484	0	1	1	C	0.14	1	1	C	0.01
485	0	1	1	X	1.22	1	1	X	0.12
486	1	2	2	C	0.63	2	2	C	0.02
487	1	6	6	C	0.69	6	6	C	0.16
491	1	57	57	G	12.21				$\perp$
492	1	52	52	G	17.96				$\perp$
504	0				$\perp$				$\perp$
519	0	3	3	C	4.71	3	3	C	0.18
546	0	3	0	o	0.07	3	0	o	0.00
559	0	71	0	o	0.32	71	0	o	0.91
581	0				$\perp$	25	25	X	6.00
584	0	9	9	C	0.60	9	9	C	0.13
619	1	8	0	o	0.07	8	0	o	0.01

TABLE 1: Applying $\mathrm{ProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$ (continued)

		Algorithm 1 ($\mathbb{C}$)				Algorithm 4 ($\mathbb{R}$)			
model	m/a	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)	$\lvert X\rvert$	$\lvert\hat{X}\rvert$	γ	time (s)
629	0	5	5	C	0.41	5	5	C	0.08
637	1	12	12	X	492.76				$\perp$
647	1	11	11	X	6.03	11	11	X	0.37

There are 15 models where the complex classification in Table 1 succeeded but the real classification timed out: 028, 030, 069, 102, 103, 122, 226, 230, 257, 287, 335, 365, 491, 492, 637. Among those, models 491 and 492 have classification G and models 230, 335, and 265 have classification o over $\mathbb{C}$, from which we can conclude that they have the same classification over $\mathbb{R}$, respectively. Vice versa, there are 5 models where real classification succeeded but complex classification timed out: 332, 334, 430, 431, 581. There is one single model where we succeeded over both $\mathbb{C}$ and $\mathbb{R}$ but obtained different classifications: model 289 has X over $\mathbb{C}$ but o over $\mathbb{R}$.

Table 2 collects some statistical information about the computations. Figure 1 provides some analysis of the computation times. Notice that many computations finish quite quickly.

TABLE 2. Statistical information about the computations in Table 1.

	Algorithm 1 ($\mathbb{C}$)	Algorithm 4 ($\mathbb{C}$)
time limit	6 h	6 h
# models	129	129
# successful computations	104	94
success rate	80.62%	72.87%
median(time)	1.33 s	0.09 s

A.2. Classifications of Rational Prime Decompositions

Recall that Algorithm 3 and Algorithm 6 compute prime decompositions $\mathcal{P}$ over $\mathbb{Q}$ and then apply our classification approach to each prime component individually. This yields lists $\mathcal{X}$, $\hat{\mathcal{X}}$ containing in turn lists X and $\hat{X}$ of variables before and after projection, respectively, as well as a list Γ of classifications γ. We have $|\mathcal{X}| = |\hat{\mathcal{X}}| = |\Gamma| = |\mathcal{P}|$, and elements can be matched by position. Since this information is too comprehensive to be displayed in a table, we give only $|\mathcal{P}|$ and summarize the numbers of occurrences of the various classifications in Γ. Our database, of course, stores the complete information.

TABLE 3: Applying $\mathrm{DecomposeProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$. Model numbers *nnn* stand for BIOMD0000000*nnn*. "m/a" indicates mass action kinetics (Definition 27). $|\mathcal{P}|$ is the number of prime components over $\mathbb{Q}$. Γ_{summary} summarizes the classification of the components using G for group, C fo coset, O for empty set, and X else; lower case letters indicate projection. Time columns give total CPU times in seconds or "$\perp$" for a timeout with a limit of 6 hours per model.

			Algorithm 3 ($\mathbb{C}$)		Algorithm 6 ($\mathbb{R}$)	
model	m/a	$\lvert\mathcal{P}\rvert$	time (s)	Γ_{summary}	time (s)	Γ_{summary}	time (s)
001	1	1	4.49	C	12.64	C	5.00
002	1	1	192.42	X	318.37		$\perp$
009	0	28	199.03	C + 13c + 14o	238.71	C + 13c + 14o	255.23
011	0	20	167.45	C + 9c + 10o	197.90	C + 9c + 10o	212.07

TABLE 3: Applying $\mathrm{DecomposeProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$ (continued)

				Algorithm 3 ($\mathbb{C}$)		Algorithm 6 ($\mathbb{R}$)	
model	m/a	$\lvert\mathcal{P}\rvert$	time (s)	Γ_{summary}	time (s)	Γ_{summary}	time (s)
026	1	2	4.74	$\mathtt{o}+\mathtt{X}$	5.95	$\mathtt{o}+\mathtt{X}$	5.36
028	1	2	135.05	$\mathtt{o}+\mathtt{X}$	162.90		$\bot$
030	1	2	157.89	$\mathtt{o}+\mathtt{X}$	175.57		$\bot$
035	0	1	12.31	$\mathtt{X}$	18.26	$\mathtt{X}$	12.76
038	0	1	391.05	$\mathtt{X}$	463.38		$\bot$
040	0	2	1.64	$\mathtt{o}+\mathtt{X}$	2.14	$\mathtt{o}+\mathtt{X}$	1.67
046	0	2	764.98	$\mathtt{X}+\mathtt{x}$	817.90		$\bot$
050	0	1	0.59	$\mathtt{o}$	0.61	$\mathtt{o}$	0.60
052	0	1	0.95	$\mathtt{o}$	0.98	$\mathtt{o}$	0.95
057	0	1	0.63	$\mathtt{C}$	5.44	$\mathtt{C}$	0.70
069	0	2	56.34	$2\mathtt{x}$	61.34		$\bot$
072	0	2	0.26	$2\mathtt{c}$	0.55	$2\mathtt{c}$	0.27
077	0	1	1.05	$\mathtt{C}$	6.26	$\mathtt{C}$	1.15
080	0	7	9.43	$3\mathtt{c}+4\mathtt{o}$	11.43	$3\mathtt{c}+4\mathtt{o}$	9.62
082	0	7	8.45	$3\mathtt{c}+4\mathtt{o}$	13.70	$3\mathtt{c}+4\mathtt{o}$	8.64
085	0		$\bot$		$\bot$		$\bot$
086	0		$\bot$		$\bot$		$\bot$
091	0	1	0.02	$\mathtt{o}$	0.06	$\mathtt{0}$	0.02
092	0	2	0.75	$\mathtt{C}+\mathtt{o}$	1.36	$\mathtt{C}+\mathtt{o}$	0.77
099	0	1	0.26	$\mathtt{C}$	0.92	$\mathtt{C}$	0.47
101	1	1	0.36	$\mathtt{X}$	5.90	$\mathtt{X}$	0.43
102	0	1	38.46	$\mathtt{X}$	249.72		$\bot$
103	0	1	3090.38	$\mathtt{X}$	20628.00		$\bot$
104	0	4	0.63	$4\mathtt{o}$	0.64	$4\mathtt{o}$	0.63
105	0	1	0.37	$\mathtt{o}$	0.55	$\mathtt{0}$	0.37
108	0	2	40.86	$2\mathtt{X}$	68.74		$\bot$
122	0	1	326.19	$\mathtt{X}$	363.45		$\bot$
123	0		$\bot$		$\bot$		$\bot$
125	0	1	0.18	$\mathtt{X}$	1.22	$\mathtt{X}$	0.21
137	0	5	180.68	$5\mathtt{x}$	220.13	$5\mathtt{x}$	190.34
147	0		$\bot$		$\bot$		$\bot$
150	0	1	1.12	$\mathtt{C}$	2.86	$\mathtt{C}$	1.15
152	0		$\bot$		$\bot$		$\bot$
153	0		$\bot$		$\bot$		$\bot$
156	0	2	0.53	$\mathtt{C}+\mathtt{o}$	0.77	$\mathtt{C}+\mathtt{o}$	0.54
158	0	1	0.22	$\mathtt{X}$	1.06	$\mathtt{X}$	0.24
159	0	1	0.20	$\mathtt{C}$	0.60	$\mathtt{C}$	0.23
163	0	1	9.69	$\mathtt{X}$	19.50	$\mathtt{X}$	10.36
173	1	1	295.07	$\mathtt{x}$	1719.63		$\bot$
178	0	1	3.66	$\mathtt{o}$	3.67	$\mathtt{o}$	3.66
186	1	5	4.73	$2\mathtt{c}+3\mathtt{x}$	7.78	$2\mathtt{c}+3\mathtt{x}$	5.09
187	1	5	18.41	$2\mathtt{c}+3\mathtt{x}$	29.25	$2\mathtt{c}+3\mathtt{x}$	18.79
188	1	1	0.27	$\mathtt{o}$	0.29	$\mathtt{0}$	0.27
189	1	1	0.02	$\mathtt{o}$	0.04	$\mathtt{0}$	0.02
193	1	1	6.50	$\mathtt{X}$	7.12	$\mathtt{X}$	6.62
194	1	1	1.38	$\mathtt{X}$	1.98	$\mathtt{X}$	1.41

TABLE 3: Applying $\mathrm{DecomposeProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$ (continued)

				Algorithm 3 ($\mathbb{C}$)		Algorithm 6 ($\mathbb{R}$)	
model	m/a	$\lvert\mathcal{P}\rvert$	time (s)	Γ_{summary}	time (s)	Γ_{summary}	time (s)
197	0	1	4.18	$\mathtt{X}$	6.96	$\mathtt{X}$	4.24
198	1	1	0.69	$\mathtt{c}$	3.16	$\mathtt{c}$	0.72
199	0	1	0.35	$\mathtt{C}$	2.63	$\mathtt{C}$	0.50
200	0		$\perp$		$\perp$		$\perp$
205	0		$\perp$		$\perp$		$\perp$
220	0		$\perp$		$\perp$		$\perp$
226	0	1	17.10	$\mathtt{X}$	25.93		$\perp$
227	0	1	0.03	$\mathtt{o}$	0.15	$\mathtt{0}$	0.03
229	0	2	1.60	$\mathtt{C}+\mathtt{c}$	6.77	$\mathtt{C}+\mathtt{c}$	1.73
230	0	5	181.14	$5\mathtt{x}$	249.78		$\perp$
233	0	3	0.14	$2\mathtt{C}+\mathtt{o}$	2.46	$2\mathtt{C}+\mathtt{o}$	0.15
243	0	8	9.13	$8\mathtt{o}$	9.19	$8\mathtt{o}$	9.18
257	1	2	0.63	$\mathtt{o}+\mathtt{X}$	2.29	$\mathtt{o}+\mathtt{X}$	590.61
259	0	1	2.93	$\mathtt{o}$	2.98	$\mathtt{o}$	2.95
260	0	1	1.41	$\mathtt{o}$	1.46	$\mathtt{o}$	1.42
261	0	1	0.84	$\mathtt{o}$	0.87	$\mathtt{o}$	0.86
262	0	1	1.83	$\mathtt{c}$	4.43	$\mathtt{c}$	1.85
263	0	1	2.48	$\mathtt{c}$	2.66	$\mathtt{c}$	2.49
264	0	1	3.12	$\mathtt{c}$	7.28	$\mathtt{c}$	3.31
267	0	1	0.10	$\mathtt{o}$	0.11	$\mathtt{o}$	0.10
270	0	1	14511.34	$\mathtt{X}$	15819.74		$\perp$
271	1	1	3.87	$\mathtt{c}$	5.31	$\mathtt{c}$	3.88
272	1	1	0.12	$\mathtt{c}$	1.11	$\mathtt{c}$	0.12
281	0	3144	8264.69	$1008\mathtt{c}+2136\mathtt{o}$	8317.80	$1008\mathtt{c}+2136\mathtt{o}$	8655.75
282	1	2	0.36	$2\mathtt{o}$	0.37	$2\mathtt{o}$	0.37
283	1	2	0.10	$2\mathtt{o}$	0.11	$2\mathtt{o}$	0.11
286	0		$\perp$		$\perp$		$\perp$
287	1	1	12.78	$\mathtt{X}$	27.29		$\perp$
289	0	2	1.15	$\mathtt{o}+\mathtt{X}$	2.84	$2\mathtt{o}$	1.15
292	0	1	1.20	$\mathtt{o}$	1.20	$\mathtt{0}$	1.20
306	0	2	0.53	$\mathtt{C}+\mathtt{o}$	1.04	$\mathtt{C}+\mathtt{o}$	0.54
307	0	1	0.02	$\mathtt{o}$	0.03	$\mathtt{o}$	0.02
310	0	1	0.02	$\mathtt{o}$	0.03	$\mathtt{o}$	0.02
311	0	1	0.02	$\mathtt{o}$	0.02	$\mathtt{o}$	0.02
312	0	1	0.04	$\mathtt{o}$	0.05	$\mathtt{o}$	0.04
314	0	3	0.94	$3\mathtt{c}$	1.11	$3\mathtt{c}$	0.99
315	1		$\perp$		$\perp$		$\perp$
321	0	1	0.07	$\mathtt{o}$	0.08	$\mathtt{o}$	0.07
332	0		$\perp$		$\perp$		$\perp$
333	0		$\perp$		$\perp$		$\perp$
334	0		$\perp$		$\perp$		$\perp$
335	1		$\perp$		$\perp$		$\perp$
344	0		$\perp$		$\perp$		$\perp$
357	1	2	0.23	$2\mathtt{o}$	0.25	$2\mathtt{o}$	0.24
359	0	5	0.96	$2\mathtt{c}+3\mathtt{o}$	2.80	$2\mathtt{c}+3\mathtt{o}$	1.02
360	0	4	0.72	$2\mathtt{c}+2\mathtt{o}$	1.48	$2\mathtt{c}+2\mathtt{o}$	0.77

TABLE 3: Applying $\mathrm{DecomposeProjectAndClassify}_K$ over $\mathbb{C}$ and $\mathbb{R}$ (continued)

				Algorithm 3 ($\mathbb{C}$)		Algorithm 6 ($\mathbb{R}$)	
model	m/a	$\lvert\mathcal{P}\rvert$	time (s)	Γ_{summary}	time (s)	Γ_{summary}	time (s)
361	0	2	0.48	$2\mathtt{c}$	1.01	$2\mathtt{c}$	0.60
362	1		$\bot$		$\bot$		$\bot$
363	0	1	0.26	$\mathtt{o}$	0.27	$\mathtt{o}$	0.26
364	1	4	2.23	$2\mathtt{c}+2\mathtt{o}$	3.22	$2\mathtt{c}+2\mathtt{o}$	2.37
365	1		$\bot$		$\bot$		$\bot$
407	0		$\bot$		$\bot$		$\bot$
413	1	1	1.07	$\mathtt{X}$	1.59	$\mathtt{X}$	1.11
416	0	3	36.24	$\mathtt{X}+2\mathtt{x}$	43.95	$\mathtt{X}+2\mathtt{x}$	41.76
430	0	20	112.47	$8\mathtt{c}+10\mathtt{o}+\mathtt{X}+\mathtt{x}$	221.05	$8\mathtt{c}+10\mathtt{o}+\mathtt{X}+\mathtt{x}$	12709.47
431	0		$\bot$		$\bot$		$\bot$
439	0	9	32.38	$2\mathtt{c}+2\mathtt{o}+\mathtt{X}+4\mathtt{x}$	40.10		$\bot$
459	0	1	0.24	$\mathtt{C}$	0.50	$\mathtt{C}$	0.25
460	0	1	1.27	$\mathtt{X}$	1.49	$\mathtt{X}$	1.30
475	0	30	30.46	$14\mathtt{c}+4\mathtt{o}+12\mathtt{x}$	37.62	$14\mathtt{c}+4\mathtt{o}+12\mathtt{x}$	43.66
478	0		$\bot$		$\bot$		$\bot$
479	0		$\bot$		$\bot$		$\bot$
483	0	1	0.23	$\mathtt{X}$	0.59	$\mathtt{X}$	0.28
484	0	1	0.10	$\mathtt{C}$	0.14	$\mathtt{C}$	0.11
485	0	1	0.16	$\mathtt{X}$	0.23	$\mathtt{X}$	0.23
486	1	1	0.15	$\mathtt{C}$	0.27	$\mathtt{C}$	0.16
487	1	1	1.40	$\mathtt{C}$	1.78	$\mathtt{C}$	1.49
491	1	1	2.55	$\mathtt{G}$	22.49	$\mathtt{G}$	47.84
492	1	1	1.59	$\mathtt{G}$	10.04	$\mathtt{G}$	35.49
504	0		$\bot$		$\bot$		$\bot$
519	0	3	1.60	$\mathtt{C}+\mathtt{c}+\mathtt{o}$	4.15	$\mathtt{C}+\mathtt{c}+\mathtt{o}$	1.62
546	0	1	0.13	$\mathtt{o}$	0.14	$\mathtt{o}$	0.13
559	0	1	0.98	$\mathtt{o}$	1.18	$\mathtt{0}$	0.99
581	0		$\bot$		$\bot$		$\bot$
584	0	1	0.26	$\mathtt{C}$	2.36	$\mathtt{C}$	0.38
619	1	1	0.86	$\mathtt{o}$	0.88	$\mathtt{o}$	0.88
629	0	1	1.03	$\mathtt{C}$	1.70	$\mathtt{C}$	1.06
637	1	3	8383.02	$\mathtt{X}+2\mathtt{x}$	8464.64		$\bot$
647	1	1	5.95	$\mathtt{X}$	6.70	$\mathtt{X}$	6.23

There are 17 models where the complex classification in Table 3 succeeded but the real classification timed out: 002, 028, 030, 038, 046, 069, 102, 103, 108, 122, 173, 226, 230, 270, 287, 439, 637. Vice versa, there are no models where the classification succeeded over $\mathbb{R}$ but not over $\mathbb{C}$. There are 8 models where we succeeded over both $\mathbb{C}$ and $\mathbb{R}$ but obtained different classifications: 091, 105, 188, 189, 227, 289, 292, 559. All those differences are visible in the summaries Γ_{summary} in Table 3. Model 289 is $\mathtt{o}+\mathtt{X}$ over $\mathbb{C}$ but $2\mathtt{o}$ over $\mathbb{R}$. We have addressed this difference already with the computations in Appendix A.1. With all other models listed above the difference is $\mathtt{o}$ over $\mathbb{C}$ in contrast to $\mathtt{0}$ over $\mathbb{R}$.

Table 4 collects some statistical information about the computations. Figure 2 provides some analysis of the computation times. Notice that many computations finish quite quickly.

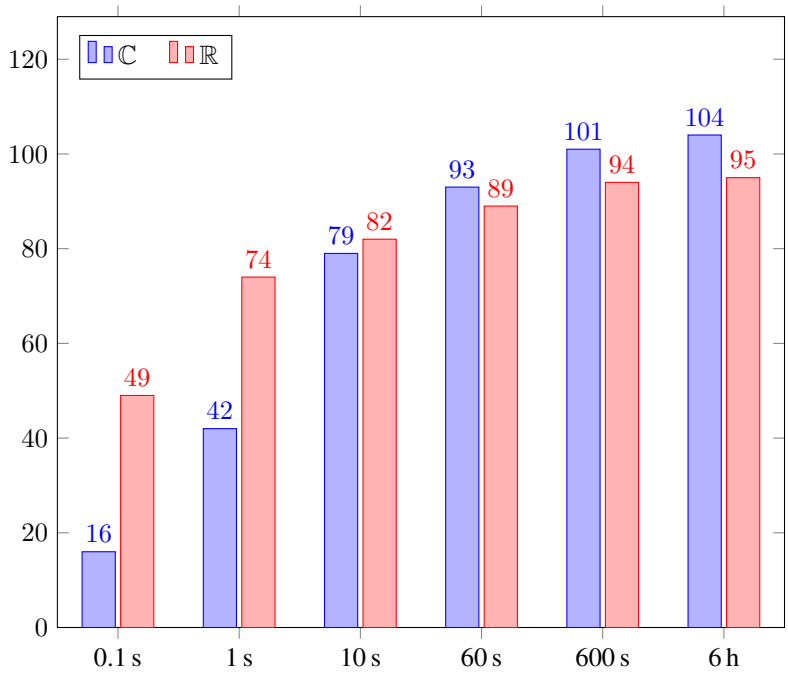


FIGURE 1. Numbers of problems solved within certain time limits by Algorithm 1 over $\mathbb{C}$ (left) and Algorithm 4 over $\mathbb{R}$ (right). The total number of problems is 129.

TABLE 4. Statistical information about the computations in Table 3.

	Algorithm 3 ($\mathbb{C}$)	Algorithm 6 ($\mathbb{C}$)
time limit	6 h	6 h
# models	129	129
# successful computations	105	88
success rate	81.40%	68.22%
median(time)	2.80 sec	0.99 sec

Dima Grigoriev
CNRS and the University of Lille, Lille, France
e-mail: dmitry.grigoryev@math.univ-lille1.fr

Alexandru Iosif
RWTH Aachen University, Aachen, Germany
e-mail: iosif@aices.rwth-aachen.de

Hamid Rahkooy
CNRS, Inria, and the University of Lorraine, Nancy, France
e-mail: hamid.rahkooy@inria.fr

Thomas Sturm
CNRS, Inria, and the University of Lorraine, Nancy, France
MPI Informatics and Saarland University, Saarbrücken, Germany
e-mail: thomas.sturm@loria.fr, sturm@mpi-inf.mpg.de

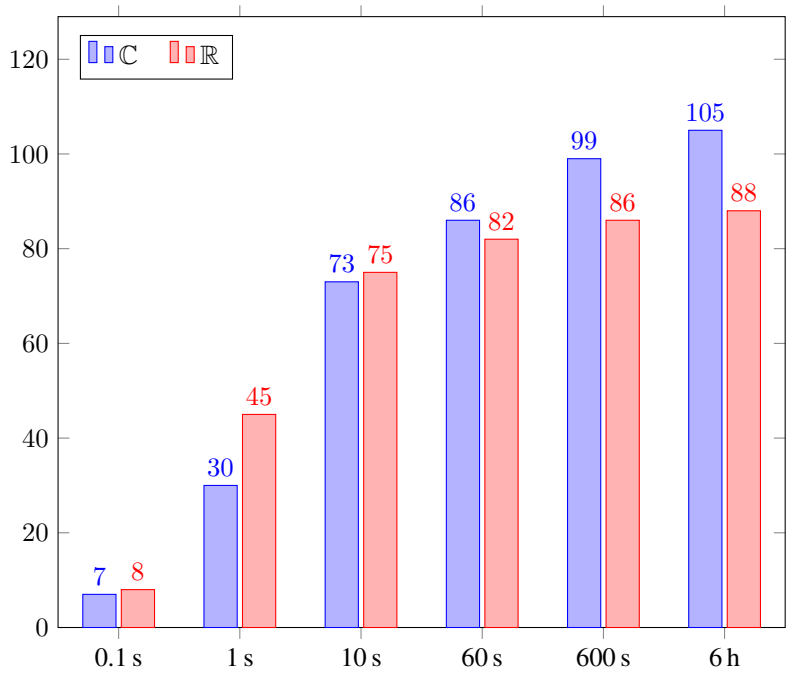


FIGURE 2. Numbers of problems solved within certain time limits by Algorithm 3 over $\mathbb{C}$ (left) and Algorithm 6 over $\mathbb{R}$ (right). The total number of problems is 129.

Andreas Weber
University of Bonn, Bonn, Germany
e-mail: weber@informatik.uni-bonn.de