Computing with Semi-Algebraic Sets: Relaxation Techniques and Effective Boundaries

Changbo Chen [a] James H. Davenport [b] Marc Moreno Maza [a]
Bican Xia [c] Rong Xiao [a]

[a] *University of Western Ontario London, Ontario, Canada N6A 5B7*

[b] *University of Bath, Bath BA2 7AY, United Kingdom*

[c] *Peking University, Beijing 100871, China*

Abstract

We discuss parametric polynomial systems, with algorithms for real root classification and triangular decomposition of semi-algebraic systems as our main applications. We exhibit new results in the theory of *border polynomials* of parametric semi-algebraic systems: in particular a geometric characterization of its "true boundary" (Definition 1). In order to optimize the corresponding decomposition algorithms, we also propose a technique, that we call *relaxation*, which can simplify the decomposition process and reduce the number of components in the output. This paper extends our earlier works [6, 7].

Key words: triangular decomposition, regular semi-algebraic system, border polynomial, effective boundary, relaxation.

1. Introduction

Triangular decompositions of semi-algebraic systems were introduced in [6] together with an algorithm for generating those decompositions. This algorithm can either be *eager*, computing the entire decomposition, or *lazy*, only computing a subset of it corresponding to the highest (complex) dimensional components, and deferring lower-dimensional components. While computing the entire decomposition is known to have a worst-case complexity which is doubly-exponential in the number of variables [4, 11], under plausible assumptions

⋆ This research was partly supported by NSERC, Maplesoft and MITACS of Canada. We are grateful to Professor Vorobjov for communicating [1].

Email addresses: cchen252@csd.uwo.ca (Changbo Chen), J.H.Davenport@bath.ac.uk (James H. Davenport), moreno@csd.uwo.ca (Marc Moreno Maza), xbc@math.pku.edu.cn (Bican Xia), rong@csd.uwo.ca (Rong Xiao).

the lazy variant has a singly-exponential complexity. Despite this encouraging complexity result, it is still desirable to improve the practical efficiency of both types of decomposition. In a subsequent paper [7], various research directions have been investigated in order to reach this goal. Two of those directions have lead us to new results which extend our work beyond the topic of triangular decomposition of semi-algebraic systems. We have chosen to dedicate this new article to these two directions while work extending the other materials of [7] will be reported elsewhere.

Solving polynomial systems of arbitrary dimension has natural connections with solving parametric polynomial systems. Algorithms for real root classification [20], cylindrical algebraic decomposition [9] and triangular decomposition of semi-algebraic systems [6] illustrate this observation. The two subjects that we discuss in this paper are based on the notion of a *border polynomial* introduced in [20] which is at the core of all algorithms solving parametric polynomial systems via triangular decomposition.

The contributions of the present paper enhance the tools supporting those algorithms in two ways. First, the notion of an *effective boundary* replaces that of a *border polynomial* with a more geometrical (and less algebraic) framework, leading to deeper results: Example 2 illustrates the difference between these two notions. Secondly, the technique of *relaxation* allows us to greatly simplify computations based on border polynomials, such as real root classification and triangular decomposition of semi-algebraic systems. To better understand the significance of these two contributions, we present informally our main results on effective boundaries. Then we illustrate the technique of relaxation with a detailed example.

One central question in the study of parametric polynomial systems is the dependence of the solutions on the parameter values. As discussed in [15], there are different ways to express the fact that the zeros of a parametric system "usually" depend *continuously* on the parameters in a neighborhood of a given parameter value. The notion of a border polynomial, introduced in [20], and the notion of a discriminant variety, introduced in [12] aims at capturing the parameter values at which this dependence is not continuous, i.e. where "things change qualitatively". Other similar but more restrictive notions like "generalized discriminant" and "generalized resultant" were introduced in [14].

Although the original definition of a border polynomial suggested in [20] is similar to the one of [15], thus based on topological considerations, the study of this notion in subsequent papers is primarily algebraic, which leads to some "unhelpful" results. Here's an example of those. For a squarefree regular chain T, regarded as a real parametric system in its free variables $\mathbf{u}$, the border polynomial $BP(T)$ encodes the locus of the $\mathbf{u}$-values at which T has lower rank or at which T is no longer a squarefree regular chain. (See §2 for the notions related to triangular decomposition and regular chains.) Consequently, for each connected component C of the complement of the real hypersurface defined by $BP(T)$ the number of real solutions of the regular chain T is constant at any point of C. However, $BP(T)$ is not an invariant of the variety $\overline{W(T)}$, [1] which is a bottleneck in designing better variety decomposition algorithms based on the notion of a border polynomial.

The definition and properties of a discriminant variety of a parametric system as presented in [12] are more intrinsic. However, this notion does not behave well under "splitting". More precisely, if the solution set of a parametric system S is the union of the solution sets of two parametric systems S_1 and S_2 (S, S_1, S_2 having of course the same parameters and unknowns), there is *a priori* no simple relationship between the minimal discriminant varieties of S, S_1 and S_2.

With the notion of an effective boundary, we aim at identifying a tool which is as independent as possible of algebraic considerations, while capturing the desirable geometrical properties of a border polynomial. In our ISSAC paper [7], the notion of an effective boundary was essentially restricted to squarefree regular chains. If T is such a set, we established in [7] that the effective boundary of T (regarded as a parametric system in its free variables) is an invariant of $\overline{W(T)}$ that is, unchanged when replacing T by T' as long as

[1] $\overline{W(T)}$ is the Zariski closure of the quasi-component of T, see Section 2

$\overline{W(T)} = \overline{W(T')}$ holds. In many ways, our notion of effective boundary is related to the "better projection" ideas in the classical [10, and many others] approach to cylindrical algebraic decomposition.

In Section 3, we study the notion of an *effective boundary* of a well-determined parametric system S, see Definition 1. Among the various results of this section, we highlight three of them below. The first one, see Theorem 1, states that the effective boundaries of S capture the topological properties of the border polynomials of S. The second, namely Theorem 3, states that transformations of S that are "generically" not significant leave the effective boundaries of S unchanged. Equivalently, this means that two well-determined parametric semi-algebraic systems S_1, S_2, which "generically" have the same zero set, have also the same effective boundaries. Definition 3 specifies the notion of genericity that we use here. The third of these main results, see Theorem 4, states that the effective boundaries are well-behaved under splitting. More precisely, if S, S_1, S_2 are three well-determined parametric semi-algebraic systems (with the same parameters and unknowns) such that a zero of S is generically either a zero of S_1 or a zero of S_2, and such that S_1 and S_2 have no common effective boundaries, then the set of effective boundaries of S is the union of those of S_1 and S_2. This type of result is clearly an important property in view of designing efficient decomposition algorithms. And again, neither border polynomials, nor discriminant varieties enjoy properties like Theorem 4.

Another important direction aiming at improving the practical efficiency of decomposition algorithms is to obtain criteria that prevent redundant computations. A well-known example of such techniques are Buchberger's criteria for the computation of Gröbner bases. Our technique of relaxation, presented in Section 4, was motivated by a similar purpose for the computation of triangular decomposition of semi-algebraic systems. Let us introduce this technique by an example. Consider the semi-algebraic system

$$S = [f = 0, x - b > 0], \quad \text{where} \quad f = ax^3 + bx - a \quad \text{and} \quad a < b < x.$$

The LazyRealTriangularize command [5] of the `RegularChains` library in MAPLE implements an algorithm of [6] which computes a triangular decomposition of S as follows. It starts by computing the border polynomial $B = \{a, b_1, b_2\}$ of the system S, regarded as parametric in (a, b), together with the fingerprint polynomial set [2] (FPS) $F = \{a, b_1, b_2, b, p_1, p_2, p_3\}$. where $b_1 = ab^3 + b^2 - a$, $b_2 = 27a^3 + 4b^3$, $p_1 = 2b^3 + 1$, $p_2 = b^3 - 4$ and $p_3 = b - 1$. Then, the LazyRealTriangularize command returns one regular semi-algebraic system, namely $S_1 = [Q_1, \{f = 0, x - b > 0\}]$, corresponding to the main component of S and seven un-evaluated recursive calls, corresponding to components of lower dimension. The regular semi-algebraic system S_1 consists of one equation, namely $f = 0$, one inequality, namely $x - b > 0$, and Q_1 which is a quantifier-free formula given below.

$$\begin{aligned}
Q_1 = {} & (b < 0 \wedge p_1 \neq 0 \wedge b_1 \neq 0 \wedge a \neq 0 \wedge b_2 \neq 0) \\
& \bigvee (p_1 > 0 \wedge b_1 > 0 \wedge a < 0 \wedge p_3 > 0 \wedge p_2 \neq 0 \wedge b_2 \neq 0) \\
& \bigvee (b > 0 \wedge p_1 > 0 \wedge b_1 \neq 0 \wedge a < 0 \wedge p_3 < 0 \wedge p_2 < 0 \wedge b_2 \neq 0) \\
& \bigvee (b > 0 \wedge p_1 > 0 \wedge b_1 < 0 \wedge a > 0 \wedge p_3 < 0 \wedge p_2 < 0 \wedge b_2 > 0).
\end{aligned}$$

The above mentioned seven recursive calls are all of the form LazyRealTriangularize($[p = 0, f = 0, x - b > 0]$), for each polynomial $p \in F$. The key observation is that some of these recursive calls can simply be avoided if some of the strict inequalities in Q_1 are relaxed, that is, replaced by non-strict inequalities. The results of §4, and in particular Theorem 7 provide criteria for this purpose. Returning to our example, when relaxation techniques are used, LazyRealTriangularize(S) will produce one regular semi-algebraic system $S_2 = [Q_2, \{f = 0, x - b > 0\}]$, and three un-evaluated recursive calls, where

[2] See Section 2 for a definition of this term

$$Q_2 = (b \leq 0 \wedge b_1 \neq 0 \wedge a \neq 0 \wedge b_2 \neq 0)$$
$$\bigvee (p_1 \geq 0 \wedge b_1 > 0 \wedge a < 0 \wedge p_3 \geq 0 \wedge b_2 \neq 0)$$
$$\bigvee (b \geq 0 \wedge p_1 \geq 0 \wedge b_1 \neq 0 \wedge a < 0 \wedge p_3 \leq 0 \wedge p_2 \leq 0 \wedge b_2 \neq 0)$$
$$\bigvee (b \geq 0 \wedge p_1 \geq 0 \wedge b_1 < 0 \wedge a > 0 \wedge p_3 \leq 0 \wedge p_2 \leq 0 \wedge b_2 > 0).$$

Moreover, it turns that these three un-evaluated recursive calls are of the form $\mathsf{LazyRealTriangularize}([p = 0, f = 0, x - b > 0])$, for $p \in B$. Continuing with that example, one can check that the full triangular decomposition of S produces 16 and 9 regular semi-algebraic systems, without and with relaxation techniques, respectively. Therefore, relaxation techniques can help simplify the output of our algorithms. More experimental evidence are provided in Section 4.3.

As we shall see in Section 4 with Example 13, relaxation techniques apply not only to triangular decomposition of semi-algebraic systems but also to real root classification. Moreover, with Theorem 7, they provide an algorithmic criterion for deciding whether a given semi-algebraic set, represented by a Tarski quantifier-free formula, is open or not.

This enhancement of Theorem 7 was not present in our ISSAC 2011, i.e. paper [7]. On the front of effective boundaries, many results are new developments w.r.t. to [7].:

- the more general definition of an effective boundary (Definition 1) which has lead us to revisit the proofs of Theorems 1 and 2,
- most of the properties of effective boundaries under splitting; in particular Theorems 3 and 4 which are completely new,
- the study of the effective boundaries of a parametric equation, in particular Theorem 6, which is also new.

2. Triangular decomposition

We summarize below the notions and notations introduced in [7]. For the definition of a triangular decomposition of an algebraic system we refer to [8, 18] and for that of an semi-algebraic system, we refer to [6].

Topology. We recall a few basic definitions for the reader's convenience. We refer to [16] for more advanced notions in general topology and to [2] for topological questions related to real algebraic geometry, such as Sard's Theorem. Let X be a topological space and S be a subset of X. The *interior* (resp. *closure*) of S, denoted by $\mathring{S}$ (resp. $\overline{S}$) consists of all points $x \in X$ such that there exists a neighborhood (resp. each neighborhood) of x contained in S (resp. contains a point of S). The *frontier* of S, denoted by ∂S consists of all points $x \in X$ such that every neighborhood of x contains at least one point of S and at least one point not of S. We have $\overline{S} = \mathring{S} \cup \partial S$. The subset S is said to be *connected* if it is connected under its subspace topology, that is, if it is not the union of two (or more) disjoint nonempty open sets of S. Assuming X not empty, the maximal connected subsets of X are called the *connected components* of X. Note that the connected components of X form a partition of X. In our proofs, arguments involving connectivity and dimension are always made for subsets of the Euclidean space $\mathbb{R}^n$, which is a fully normal topological space (while spaces endowed with Zariski Topology may not even be normal topological spaces). In particular, this context establishes an equivalence between the topological (or axiomatic) and algebraic notions of dimension.

Zero sets. In this paper, we use "Z" to denote the zero set over $\mathbb{C}$ of a polynomial system, involving equations and inequations. We use "$Z_{\mathbb{R}}$" to denote the zero set over $\mathbb{R}$ of a semi-algebraic system. If a semi-algebraic set S is finite, we denote by $\#(S)$ the number of its distinct points. In $\mathbb{R}^n$, we use the Euclidean topology; in $\mathbb{C}^n$, we use the Zariski topology.

Polynomials. Throughout this paper, all polynomials are in $\mathbb{Q}[\mathbf{x}]$, with ordered variables $\mathbf{x} = x_1 < \cdots < x_n$. We order monomials of $\mathbb{Q}[\mathbf{x}]$ by the lexicographical ordering induced by $x_1 < \cdots < x_n$. Then, we require that the coefficient of the leading monomial of every polynomial in a regular chain or in a border polynomial set (defined hereafter) is equal to 1. Let p be a polynomial in $\mathbb{Q}[\mathbf{x}] \setminus \mathbb{Q}$. We denote by $\mathrm{mvar}(p)$, $\mathrm{init}(p)$, $\mathrm{mdeg}(p)$ and $\mathrm{der}(p)$ respectively the greatest variable appearing in p (called the *main variable* of p), the leading coefficient of p w.r.t. $\mathrm{mvar}(p)$ (called the *initial* of p), the degree of p w.r.t. $\mathrm{mvar}(p)$ (called the *main degree* of p) and the derivative of p w.r.t. $\mathrm{mvar}(p)$. Let $v \in \mathbf{x}$. Denote by $\mathrm{lc}(p, v)$, $\deg(p, v)$, $\mathrm{der}(p, v)$, $\mathrm{discrim}(p, v)$ respectively the leading coefficient, degree, derivative and discriminant of p w.r.t. v.

Triangular set. Let $T \subset \mathbb{Q}[\mathbf{x}]$ be a *triangular set*, that is, a set of non-constant polynomials with pairwise distinct main variables. Denote by $\mathrm{mvar}(T)$ the set of main variables of the polynomials in T. A variable v in $\mathbf{x}$ is called *algebraic* w.r.t. T if $v \in \mathrm{mvar}(T)$, otherwise it is said *free* w.r.t. T. If no confusion is possible, we shall always denote by $\mathbf{u} = u_1, \ldots, u_d$ and $\mathbf{y} = y_1, \ldots, y_m$ $(m + d = n)$ respectively the free and the main variables of T. When T is regarded as a *parametric system*, the free variables in T are its parameters.

Let h_T be the product of the initials of the polynomials in T. We denote by $\mathrm{sat}(T)$ the *saturated ideal* of T: if T is the empty triangular set, then $\mathrm{sat}(T)$ is defined as the trivial ideal $\langle 0 \rangle$, otherwise it is the colon ideal $\langle T \rangle : h_T^{\infty}$. The *quasi-component* $W(T)$ of T is defined as $V(T) \setminus V(h_T)$. Denote by $\overline{W(T)}$ the Zariski closure of $W(T)$, which is equal to $V(\mathrm{sat}(T))$. Denote by $W_{\mathbb{R}}(T)$ the set $Z_{\mathbb{R}}(T) \setminus Z_{\mathbb{R}}(h_T)$.

Iterated resultant. Let $p, q \in \mathbb{Q}[\mathbf{x}] \setminus \mathbb{Q}$. Let $v = \mathrm{mvar}(q)$. Denote by $\mathrm{res}(p, q, v)$ the resultant of p, q w.r.t. v. Let $T \subset \mathbb{Q}[\mathbf{x}]$ be a triangular set. We define $\mathrm{res}(p, T)$ inductively: if T is empty, then $\mathrm{res}(p, T) = p$; otherwise let v be the largest variable occurring in T, then $\mathrm{res}(p, T) = \mathrm{res}(\mathrm{res}(p, T_v, v), T_{<v})$, where T_v and $T_{<v}$ denote respectively the polynomials of T with main variables equal to and less than v.

Regular chain. A triangular set $T \subset \mathbb{Q}[\mathbf{x}]$ is called a *regular chain* if: either T is empty; or (letting t be the polynomial in T with maximum main variable), $T \setminus \{t\}$ is a regular chain, and the initial of t is regular [3] w.r.t. $\mathrm{sat}(T \setminus \{t\})$. Let $H \subset \mathbb{Q}[\mathbf{x}]$. The pair $[T, H]$ is a *regular system* if each polynomial in H is regular modulo $\mathrm{sat}(T)$. A regular chain T or a regular system $[T, H]$, is *squarefree* if for all $t \in T$, the polynomial $\mathrm{der}(t)$ is regular w.r.t. $\mathrm{sat}(T)$. Given $u \in \mathbb{R}^d$, we say that a squarefree regular system $[T, H]$ *specializes well* at u if $h_T(u) \neq 0$ and $[T(u), H(u)]$ is a squarefree regular system. A regular chain is called *d-dimensional* if it has exactly d free variables, that is, if its saturated ideal has dimension d.

Semi-algebraic system. Consider four finite polynomial sets $F = \{f_1, \ldots, f_s\}$, $N = \{n_1, \ldots, n_k\}$, $P = \{p_1, \ldots, p_e\}$, and $H = \{h_1, \ldots, h_\ell\}$ of $\mathbb{Q}[\mathbf{x}]$. Let $N_{\geq}$ denote the set of non-negative inequalities $\{n_1 \geq 0, \ldots, n_k \geq 0\}$. Let $P_>$ denote the set of positive inequalities $\{p_1 > 0, \ldots, p_e > 0\}$. Let $H_{\neq}$ denote the set of inequations $\{h_1 \neq 0, \ldots, h_\ell \neq 0\}$. We denote by $S = [F, N_{\geq}, P_>, H_{\neq}]$ the *semi-algebraic system* (SAS) defined as the conjunction of the constraints $f_1 = \cdots = f_s = 0$, $N_{\geq}$, $P_>$, $H_{\neq}$. When $N_{\geq}$, $H_{\neq}$ are both empty, the system S is called a *basic semi-algebraic system* and denoted by $[F, P_>]$.

Triangular semi-algebraic systems. Let $T \subset \mathbb{Q}[\mathbf{x}]$ be a regular chain with $\mathbf{u} = u_1, \ldots, u_d$ as free variables and $\mathbf{y} = y_1 < \cdots < y_m$ as main variables, so that $\mathbf{x} \setminus \mathbf{y} = \mathbf{u}$. Let H and P be as above. The pair $[T, P_>]$ (resp. $[T, H_{\neq}]$) is called a *squarefree triangular semi-algebraic system*, STSAS for short, if $[T, P]$ (resp. $[T, H]$) forms a squarefree regular system. More generally, the triple $[T, H_{\neq}, P_>]$ is called an STSAS if $[T, H \cup P]$ is a squarefree regular system. A point of $\mathbb{R}^{d+m}$ is a zero of $[T, H_{\neq}, P_>]$ if it is a zero of $[T, H_{\neq}]$ making each polynomial of $P_>$ strictly positive. Unless specified differently, we will regard $[T, P_>]$, $[T, H_{\neq}]$, $[T, H_{\neq}, P_>]$ as parametric semi-algebraic systems with $\mathbf{u}$ as parameters and $\mathbf{y}$ as unknowns.

Regular semi-algebraic system. Let T and P be as above. let $\mathcal{Q}$ be a quantifier-free formula of $\mathbb{Q}[\mathbf{u}]$. We say that $R := [\mathcal{Q}, T, P_>]$ is a *regular semi-algebraic system* if

[3] We say that a polynomial $p \in \mathbb{Q}[\mathbf{x}]$ is regular w.r.t. an ideal $\mathcal{I} \subset \mathbb{Q}[\mathbf{x}]$ if p is neither null nor a zero-divisor modulo $\mathcal{I}$

(*i*) $\mathcal{Q}$ defines a non-empty open semi-algebraic set S in $\mathbb{R}^d$;
(*ii*) $[T, P]$ specializes well at every point of S,
(*iii*) at each $u \in S$, the specialized system $[T(u), P(u)_>]$ has at least one real zero.
Regular semi-algebraic systems play the role for semi-algebraic systems that regular systems play from algebraic systems, that is, the zero set of any semi-algebraic systems can be decomposed into finitely many zero sets of regular semi-algebraic systems. See [6] for details.

Well-determined parametric polynomial system. Let S be a semi-algebraic system defined by polynomials in $\mathbb{Q}[\mathbf{u}, \mathbf{y}]$ where $\mathbf{u} = u_1, \ldots, u_d$ are regarded as parameters and $\mathbf{y} = y_1 < \cdots < y_m$ as unknowns. Since we are solving parametrically over the reals, we consider, the canonical projection Π_{U}:

$$\Pi_{\mathrm{U}} : Z_{\mathbb{R}}(S) \subset \mathbb{R}^{d+m} \mapsto \mathbb{R}^d$$

$$\Pi_{\mathrm{U}}(u_1, \ldots, u_d, y_1, \ldots, y_m) = (u_1, \ldots, u_d)$$

Let us denote by F (resp. I) the set of the polynomials of S defining equations (resp. inequations or strict inequalities). The ideal $\langle F \rangle : (\prod_{h \in I} h)^{\infty}$ is called the *ideal associated with* S. We say that S is *well-determined* if the set $\mathbf{u}$ is an $\subseteq$-maximal algebraic independent variable set modulo the ideal associated with S. Note that the notion of " well-determined" is more general than the notion of "well-behaved" used in [12], in the sense that it is less restrictive for F. Indeed, the polynomial set F is not required to have exactly m elements, nor to generate a radical ideal in $\mathbb{Q}(\mathbf{u})[\mathbf{y}]$.

Example 1. Consider the following semi-algebraic system

$$S = \{x(x^2 + ay + b) = x(y^2 + bx + a) = 0, x > 0\}$$

with parameters a, b. The ideal generated by the polynomials defining the equations of S is

$$\langle x \rangle \cap \langle x^2 + ay + b, y^2 + bx + a \rangle.$$

The polynomial system $S' = \{x(x^2+ay+b) = x(y^2+bx+a) = 0\}$ with parameters a, b is not well-determined, since $\{a, b\}$ is not a maximal algebraic independent set modulo $\langle x \rangle$. However, the ideal associated to S is $\mathcal{I} := \langle x^2+ay+b, y^2+bx+a \rangle$, and $\{a, b\}$ is a maximal algebraic independent variable set modulo $\mathcal{I}$. Therefore, S is a well-determined parametric semi-algebraic system.

Border polynomial [20, 21, 6, 15]. Let S be a well-determined parametric polynomial system of $\mathbb{Q}[\mathbf{u}, \mathbf{x}]$. Let $\alpha \in \mathbb{C}^d$ (resp. $\alpha \in \mathbb{R}^d$). We say that S is *Z-continuous* (resp. *$Z_{\mathbb{R}}$-continuous*) at α if there exists a neighborhood O_α of α such that for any two parameter values $\alpha_1, \alpha_2 \in O_\alpha$, we have $\#(Z(S(\alpha_1)) = \#(Z(S(\alpha_2))$ (resp. $\#(Z_{\mathbb{R}}(S(\alpha_1)) = \#(Z_{\mathbb{R}}(S(\alpha_2)))$. A polynomial b in $\mathbb{Q}[\mathbf{u}]$ is called a *border polynomial* of the parametric polynomial system S over $\mathbb{C}$ (resp. over $\mathbb{R}$) if the zero set $Z(b)$ (resp. $Z_{\mathbb{R}}(b)$) contains all the points at which S is not Z-continuous (resp. not $Z_{\mathbb{R}}$-continuous). In this paper, unless specified differently, all border polynomials are considered over $\mathbb{R}$. Let R be either a squarefree regular chain T, or a squarefree regular system $[T, H]$, or an STSAS $[T, P_>]$ in $\mathbb{Q}[\mathbf{x}]$. We denote by $B_{sep}(T)$, $B_{ini}(T)$, $B_{ineqs}([T, P])$ the set of the irreducible factors of $\prod_{t \in T} \mathrm{res}(\mathrm{discrim}(t, \mathrm{mvar}(t)), T)$, $\prod_{t \in T} \mathrm{res}(\mathrm{init}(t), T)$, and $\prod_{f \in P} \mathrm{res}(f, T)$, respectively. The set $\mathrm{BP}(R)$ defined as $B_{sep}(T) \cup B_{ini}(T) \cup B_{ineqs}([T, P])$ is called a *border polynomial set* of S. Lemma 1 justifies the terminology while Lemma 2 states the fundamental property of border polynomials.

Lemma 1. With the above notations, the polynomial $\prod_{f \in \mathrm{BP}(R)} f$ is a border polynomial of R.

Lemma 2. Let $b \in \mathbb{Q}[\mathbf{u}]$ be a border polynomial of S. Then, for any connected component C of $Z_{\mathbb{R}}(b \neq 0)$ and for any two parameter values u_1, u_2 in C, we have $\#Z_{\mathbb{R}}(R(u_1)) = \#Z_{\mathbb{R}}(R(u_2))$.

Fingerprint polynomial set. Let T, P be as above and let $B \subset \mathbb{Q}[\mathbf{u}]$ be finite. The polynomial system $R = [B_{\neq}, T, P_{>}]$ is called a *pre-regular semi-algebraic system*, if each $p \in \mathrm{BP}([T, P_{>}])$ is a factor of some polynomial in B. Suppose R is a pre-regular semi-algebraic system. A polynomial set $D \subset \mathbb{Q}[\mathbf{u}]$ is called a *fingerprint polynomial set* (FPS) of R if:

(i) we have $Z_{\mathbb{R}}(D_{\neq}) \subseteq Z_{\mathbb{R}}(B_{\neq})$ and,

(ii) for all $\alpha, \beta \in Z_{\mathbb{R}}(D_{\neq})$ with $\alpha \neq \beta$, if the signs of $p(\alpha)$ and $p(\beta)$ are the same for all $p \in D$, then $R(\alpha)$ has real solutions if and only if $R(\beta)$ does.

Open CAD operator [17, 3, 6]. Let $\mathbf{u} = u_1 < \cdots < u_d$ be ordered variables. For a polynomial $p \in \mathbb{Q}[\mathbf{u}]$, denote by $\mathrm{factor}(p)$ the set of the non-constant irreducible factors of p; for $A \subset \mathbb{Q}[\mathbf{u}]$, define $\mathrm{factor}(A) = \cup_{p \in A} \mathrm{factor}(p)$. For a squarefree polynomial p, the *open projection operator* (oproj) w.r.t. a variable $v \in \mathbf{u}$ is defined as below:

$$\mathrm{oproj}(p, v) := \mathrm{factor}(\mathrm{discrim}(p, v)\,\mathrm{lc}(p, v)).$$

If p is not squarefree, then we define $\mathrm{oproj}(p, v) := \mathrm{oproj}(p^*, v)$, where p^* is the squarefree part of p; then for a polynomial set A, we define $\mathrm{oproj}(A, v) := \mathrm{oproj}(\Pi_{f \in A} f, v)$.

Given $A \subset \mathbb{Q}[\mathbf{u}]$ and $x \in \{u_1, \ldots, u_d\}$, denote by $\mathrm{der}(A, x)$ the *derivative closure* of A w.r.t. x. The *open augmented projected factors* of A, denoted by $\mathrm{oaf}(A)$, is defined as follows. Let k be the smallest positive integer such that $A \subset \mathbb{Q}[u_1, \ldots, u_k]$ holds. Let $C = \mathrm{factor}(\mathrm{der}(A, u_k))$. Then, we have

(1) if $k = 1$, then $\mathrm{oaf}(A) := C$,

(2) if $k > 1$, then $\mathrm{oaf}(A) := C \cup \mathrm{oaf}(\mathrm{oproj}(C, u_k))$.

3. Effective Boundaries

Throughout this section, we consider a well-determined parametric semi-algebraic system S with parameters $\mathbf{u} = u_1, u_2, \ldots, u_d$ and unknowns $\mathbf{y} = y_1 < \cdots < y_m$, with $d \geq 1$ and $m \geq 1$.

Definition 1 formalizes the notion of an effective boundary for S and thus extends the definition we introduced in [7] for the restricted case of squarefree triangular semi-algebraic systems. Moreover Definition 1 is somewhat simpler than that of [7] and relies on the following observation. Let $\mathbf{h} \subset \mathbb{R}^d$ be a real hypersurface defined by $f(\mathbf{u}) = 0$, for $f \in \mathbb{Q}[\mathbf{u}]$. Let α be a point of $\mathbf{h}$ at which $\mathbf{h}$ is not singular. Then there exists an open ball O centered at α such that $O \setminus \mathbf{h}$ admits two connected components (in the topological space induced by O).

This result could be derived from the implicit function theorem. It turns out that it is also an immediate consequence of Lemma 2.15 in [1] by S. Basu, A. Gabrielov, and N. Vorobjov. This lemma is a generalization of the Jordan–Brouwer Theorem that suits our above observation. A similar result appearing in the literature is another generalization of the Jordan–Brouwer Theorem by E.L. Lima [13], see the remark at the end of this latter paper.

Definition 1 (Effective boundary)**.** Let $\mathbf{h} \subset \mathbb{R}^d$ be a hypersurface defined by an irreducible polynomial $p \in \mathbb{Q}[\mathbf{u}]$. We say that $\mathbf{h}$ is an *irreducible effective boundary* for S if there exists an open ball $O \subset \mathbb{R}^d$ satisfying the following three conditions:

(a) $O \setminus \mathbf{h}$ consists of two connected components O_1, O_2,

(b) for $i = 1, 2$ and any two points $\alpha_1, \alpha_2 \in O_i$ we have $\#Z(S(\alpha_1)) = \#Z(S(\alpha_2))$,

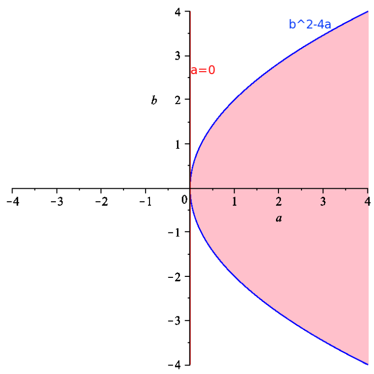
(c) for any $\beta_1 \in O_1, \beta_2 \in O_2$ we have $\#Z(S(\beta_1)) \neq \#Z(S(\beta_2))$.

When the above holds, we say that O is a *witness ball* for $\mathbf{h}$ and we say that p is an *irreducible effective border factor*. The union of all irreducible effective boundaries of S is called the *effective boundary* of S, denoted by $\mathrm{eb}(S)$; the set of all irreducible border factors is denoted $\mathrm{ebf}(S)$.

The following example illustrates the notion of an effective boundary.

Example 2. Consider the semi-algebraic system $S = \{ax^2 + bx + 1 = 0\}$ with parameters a, b. One border polynomial of S is $a(b^2 - 4a)$. One can verify from Figure 1 that $Z_{\mathbb{R}}(b^2 - 4a = 0)$ is an irreducible effective boundary of S. Meanwhile, $Z_{\mathbb{R}}(a = 0)$ is not, as the behavior on both sides of this hypersurface is the same, even though the behavior *on* the line itself is different. Indeed, all (a, b)-values in the unfilled area (note that $a = 0$ is filled) will specialize S to have 2 real solutions while all (a, b)-values in the filled region will specialize S to have no real solutions.

Fig. 1. Effective and non-effective boundary



The following lemma is a basic property of witness balls that we often use in the proof of other results of this section.

Lemma 3. With the notations of Definition 1, let O be a witness ball of the hypersurface $\mathbf{h}$ defined by $Z_{\mathbb{R}}(p = 0)$. Recall that $O \setminus \mathbf{h}$ consists of two connected components O_1, O_2. Then $O \cap Z_{\mathbb{R}}(p = 0)$ has dimension $d - 1$ and we have

$$O \cap \partial O_1 \cap \partial O_2 = O \cap Z_{\mathbb{R}}(p = 0).$$

Proof. Observe that $\{O_1, O_2, O \cap Z_{\mathbb{R}}(p = 0)\}$ forms a partition of O. Observe also that O, O_1, O_2 have dimension d. On one hand this implies that any open ball centered at a point of $O \cap Z_{\mathbb{R}}(p = 0)$ meets both O_1, O_2. Thus we have $O \cap Z_{\mathbb{R}}(p = 0) \subseteq O \cap \partial O_1 \cap \partial O_2$. On the other hand, both $O \cap \partial O_1$ and $O \cap \partial O_2$ are clearly contained in $O \cap Z_{\mathbb{R}}(p = 0)$. Thus $\partial O_1 \cap \partial O_2 \supseteq O \cap Z_{\mathbb{R}}(p = 0)$ also holds. Moreover, it is clear that $O \cap Z_{\mathbb{R}}(p = 0)$ has dimension $d - 1$. $\square$

The following proposition shows that we have much flexibility to choose witness balls for a given irreducible effective boundary.

Proposition 1. Let p be an irreducible effective border factor for the effective boundary $\mathbf{h}$ of S. Let b be a non-zero squarefree polynomial in $\mathbb{Q}[u]$. Then there exists a witness ball O_b of $\mathbf{h}$ satisfying $(O_b \setminus \mathbf{h}) \subset Z_{\mathbb{R}}(b \neq 0)$. Moreover, if $\mathbf{h} \not\subseteq Z_{\mathbb{R}}(b = 0)$ holds, one can choose O_b such that $O_b \subset Z_{\mathbb{R}}(b \neq 0)$ holds. Furthermore, in both cases, we can impose the following additional condition: the set $O_b \cap \mathbf{h}$ does not contain any singular points of $\mathbf{h}$.

Proof. We first prove the second claim and, thus, we assume that $\mathbf{h} \not\subseteq Z_{\mathbb{R}}(b = 0)$ holds. Let O be a witness ball for $\mathbf{h}$. Since $O \setminus \mathbf{h}$ is not connected, the dimension of $\mathbf{h} \cap O$ must be $d - 1$. From our assumption, it follows that $(\mathbf{h} \cap O) \setminus Z_{\mathbb{R}}(b = 0)$ is not empty. Let α be a point in $(\mathbf{h} \cap O) \setminus Z_{\mathbb{R}}(b = 0)$. Let r be the distance from α to $Z_{\mathbb{R}}(b = 0)$. Consider O_b an open ball centered at α and with radius less than r. Then we have

$O_b \subset Z_{\mathbb{R}}(b \neq 0)$. It is easy to verify that O_b is a witness ball for $\mathbf{h}$ as well. This proves our second claim. We now prove the first claim of the proposition. If $\mathbf{h} \not\subseteq Z_{\mathbb{R}}(b = 0)$ holds, the conclusion follows immediately from the second claim. So let us assume from now on that $\mathbf{h} \subseteq Z_{\mathbb{R}}(b = 0)$ holds, i.e. p divides b. Moreover, since b is squarefree, we have $\gcd(p, \frac{b}{p}) = 1$. Therefore, we have $\mathbf{h} \not\subseteq Z_{\mathbb{R}}(\frac{b}{p} = 0)$. Applying the second claim, we deduce that there exists a witness ball O_b for $\mathbf{h}$ such that $O_b \subset Z_{\mathbb{R}}(\frac{b}{p} \neq 0)$. It is obvious that $(O_b \setminus \mathbf{h}) \subset Z_{\mathbb{R}}(b \neq 0)$ holds as well. Finally, the third claim of the proposition follows from Sard's Theorem. □

3.1. Border polynomials and effective boundaries

The goal of this section is to show the notion of an effective boundary captures the topological properties of border polynomials. A first step in this construction is Theorem 1, which has also another important consequence: the effective boundaries of S are computable, which was not obvious from Definition 1. A second step is Theorem 2 which states that the effective boundaries of S play essentially the same role as its border polynomials in the context of parametric polynomial system solving. That is, in broad terms, the effective boundaries of S capture the locus of the points in the parameter space at which S is not $Z_{\mathbb{R}}$-continuous [4].

Before entering the results of this section, let us recall that the hypersurface defined by a border polynomial of S partitions the parameter space into regions, where the number of real solutions is locally invariant. One might imagine that the effective boundaries of S are strongly related to a border polynomial of S, thanks to this property. Indeed, we have the following proposition stating this relation.

Proposition 2. Let $b \in \mathbb{Q}[\mathbf{u}]$ be a border polynomial of S and let $p \in \mathbb{Q}[\mathbf{u}]$ be a polynomial defining an irreducible effective boundary of S. Then, the polynomial p divides b in $\mathbb{Q}[\mathbf{u}]$.

Proof. We proceed by contradiction. Let $\mathbf{h}$ be the irreducible effective boundary of S defined by p. Suppose $\mathbf{h} \not\subseteq Z_{\mathbb{R}}(b = 0)$ holds. Then by Proposition 1, we can choose a witness ball O of $\mathbf{h}$ such that $O \subset Z_{\mathbb{R}}(b \neq 0)$ holds. By Lemma 2, for any two points $\alpha_1, \alpha_2 \in O$, the equality $\#Z_{\mathbb{R}}(S(\alpha_1)) = \#Z_{\mathbb{R}}(S(\alpha_2))$ holds. That is a contradiction with the fact that O is a witness ball, see point (c) in Definition 1. □

Remark 1. The above Proposition 2 justifies the terminology of an irreducible effective border factor. This result also implies that there are only finitely many irreducible effective border factors, and thus only finitely many irreducible effective boundaries. Therefore, the set $\mathrm{eb}(S)$ itself is a hypersurface and $\mathrm{ebf}(S)$ is a finite set in $\mathbb{Q}[\mathbf{u}]$.

Remark 2. It follows from Proposition 2 that $\mathrm{eb}(S)$ is a subset of the intersection of all the hypersurfaces $Z_{\mathbb{R}}(b)$ where b is any border polynomial of S. In addition, the set $\mathrm{eb}(S)$ is also a subset of the minimal discriminant variety W of S. This follows from the fact if $\{g_1, \ldots, g_e\} \subset \mathbb{Q}[\mathbf{u}]$ is a basis of the ideal $\mathrm{I}(W) \subset \mathbb{Q}[\mathbf{u}]$, then each polynomial g_i is a border polynomial of S, see [15].

With Definition 1, one may wonder whether the set $\mathrm{eb}(S)$ is computable. The following theorem will lead to a positive answer.

Theorem 1. Let b be a border polynomial of S. An irreducible factor p of b is an irreducible effective border factor of S if and only if there exist two connected components C_1, C_2 of $Z_{\mathbb{R}}(b \neq 0)$ satisfying the following two properties:
(1) $\partial C_1 \cap \partial C_2 \cap Z_{\mathbb{R}}(p = 0)$ is of dimension $d - 1$,
(2) for each pair $(\alpha_1, \alpha_2) \in C_1 \times C_2$ we have $\#Z_{\mathbb{R}}(S(\alpha_1)) \neq \#Z_{\mathbb{R}}(S(\alpha_2))$.

[4] See Section 2 for that term.

Proof. "$\Rightarrow$". Assume that p is an effective border polynomial factor of S. Let O be a witness ball for p. According to Proposition 1, we can choose O such that $O \setminus Z_{\mathbb{R}}(p=0) \subset Z_{\mathbb{R}}(b \neq 0)$ holds. By definition of a witness ball, the set $O \setminus Z_{\mathbb{R}}(p=0)$ consists of two connected components, say O_1, O_2. Let C_1 and C_2 be the connected components of $Z_{\mathbb{R}}(b \neq 0)$ containing O_1 and O_2, respectively. Since $\partial C_1 \cap \partial C_2 \cap Z_{\mathbb{R}}(p=0)$ clearly contains $\partial O_1 \cap \partial O_2 \cap Z_{\mathbb{R}}(p=0)$, it follows from Lemma 3 that $\partial C_1 \cap \partial C_2 \cap Z_{\mathbb{R}}(p=0)$ is of dimension $d-1$. Now the above property (2) follows directly from the definition of a border polynomial.

"$\Leftarrow$". Suppose there exist two connected components C_1, C_2 of $Z_{\mathbb{R}}(b \neq 0)$ satisfying the above (1) and (2) in the theorem statement. Let u^* be a non-singular point of $Z_{\mathbb{R}}(p=0)$ belonging to $(\partial C_1 \cap \partial C_2 \cap Z_{\mathbb{R}}(p=0)) \setminus Z_{\mathbb{R}}(\frac{b}{p}=0)$. As in Proposition 1, we can choose an open ball O_{u^*} centered at u^* such that $O_{u^*} \subseteq Z_{\mathbb{R}}(\frac{b}{p} \neq 0)$. Moreover, since u^* is a non-singular, we can choose O_{u^*} such that the set $O_{u^*} \setminus Z_{\mathbb{R}}(p=0)$ consists of two connected components.

Let C be the connected component of $Z_{\mathbb{R}}(\frac{b}{p} \neq 0)$ containing u^*. Observe that $O_{u^*} \subset C$ holds, which implies

$$O_{u^*} \setminus Z_{\mathbb{R}}(p=0) = O_{u^*} \cap (C \setminus Z_{\mathbb{R}}(p=0)). \tag{1}$$

Observe that $C_1 \subset C$ and $C_2 \subset C$ hold as well. Indeed $O_{u^*} \cap C_1$, $O_{u^*} \cap C_2$ are both not empty and C_1, C_2 are connected components of $Z_{\mathbb{R}}(b \neq 0)$. Therefore, both C_1 and C_2 are two of the connected components of $C \setminus Z_{\mathbb{R}}(p=0)$ (which is contained in $Z_{\mathbb{R}}(\frac{b}{p} \neq 0)$). Observe now that we clearly have

$$O_{u^*} \cap C = (O_{u^*} \cap C_1) \cup (O_{u^*} \cap C_2) \cup (O_{u^*} \setminus (C_1 \cap C_2)). \tag{2}$$

Since C_1, C_2 are two connected components of $C \setminus Z_{\mathbb{R}}(p=0)$, Relations (1) and (2) lead to the following disjoint union

$$\begin{aligned} O_{u^*} \setminus Z_{\mathbb{R}}(p=0) &= O_{u^*} \cap (C \setminus Z_{\mathbb{R}}(p=0)) \\ &= (O_{u^*} \cap C_1) \ \uplus \ (O_{u^*} \cap C_2) \ \uplus \ (((C \setminus Z_{\mathbb{R}}(p=0)) \setminus (C_1 \cup C_2)) \cap O_{u^*}), \end{aligned} \tag{3}$$

Recall that $O_{u^*} \setminus Z_{\mathbb{R}}(p=0)$ has exactly two connected components. Since $(O_{u^*} \cap C_1)$ and $(O_{u^*} \cap C_2)$ are both not empty, we deduce

$$O_{u^*} \setminus Z_{\mathbb{R}}(p=0) = (O_{u^*} \cap C_1) \uplus (O_{u^*} \cap C_2).$$

Then, clearly O_{u^*} is a witness ball for $Z_{\mathbb{R}}(p=0)$. $\square$

Remark 3. The above theorem implies that the effective border factors are computable. This can be achieved with the adjacency information and sample points of a cylindrical algebraic decomposition of $Z_{\mathbb{R}}(b \neq 0)$. However, computing effective border factors is not the goal of this article and we will return to this question in a future paper.

The following theorem states that, given a border polynomial b of S, the complement of $\mathrm{eb}(S)$ and the complement of $Z_{\mathbb{R}}(b=0)$ have similar properties w.r.t. to the number of solutions of S.

Theorem 2. Let $b \in \mathbb{Q}[\mathbf{u}]$ be a border polynomial of S. If two points $\alpha_1, \alpha_2 \in Z_{\mathbb{R}}(b \neq 0)$ are in the same connected component of the complement of $\mathrm{eb}(S)$, then $\#Z_{\mathbb{R}}(S(\alpha_1)) = \#Z_{\mathbb{R}}(S(\alpha_2))$ holds.

Proof. Let $p \in \mathbb{Q}[\mathbf{u}]$ be a squarefree polynomial such that $Z_{\mathbb{R}}(p=0) = \mathrm{eb}(S)$ holds. Recall that p divides b. Let $g := \frac{b}{p}$. By definition of p, no irreducible factors of g is an effective border factor.

Let C be a connected component of $Z_{\mathbb{R}}(p \neq 0)$. Since p divides b, every connected component of $Z_{\mathbb{R}}(b \neq 0)$ is a subset of a connected component of $Z_{\mathbb{R}}(p \neq 0)$. Thus, each connected component of $C \cap Z_{\mathbb{R}}(b \neq 0)$ is a connected component of $Z_{\mathbb{R}}(b \neq 0)$. Consequently, we can consider $C_1, \ldots, C_e$ the connected components of

$Z_{\mathbb{R}}(b \neq 0)$ which are also all the connected components of $C \cap Z_{\mathbb{R}}(b \neq 0)$. We observe that, if $e = 1$ holds, then the conclusion of the theorem follows immediately from the definition of a border polynomial.

Assume from now on that $e > 1$ holds. We can write down a finite sequence $\mathcal{S} := C'_1, \ldots, C'_k$ $(k \geq e)$ such that: (1) each C'_i $(i = 1 \cdots k)$ is one of $C_1, \ldots, C_e$; (2) each C_i $(i = 1 \cdots e)$ is in $\mathcal{S}$; (3) for all $1 \leq j < k$, the dimension of $\partial C'_i \cap \partial C'_{i+1}$ is $d - 1$. For each $1 \leq i \leq e$, let n_i be the number of real solutions of S after specialization at a point of C'_i. The proof of this theorem will be complete as soon as we have established that all n_i's are equal. Proceeding by contradiction, we assume that the conclusion is false. Thus, there must exists $1 \leq j < e$ such that $n_j \neq n_{j+1}$ holds, which implies that there exists an irreducible factor of g which is also an irreducible effective border factor according to Theorem 1. This is a contradiction to the definition of g. $\square$

3.2. Properties of effective boundaries under splitting

The main objective of this section is to show that effective boundaries are well-behaved under splitting, which is an important property in view of efficient algorithm design. To be more precise, Theorem 4 states that if the solution set of a parametric polynomial system S decomposes generically (in some technical sense) into the union of the zero sets of two parametric polynomial systems S_1 and S_2, then every effective boundary of S is either an effective boundary of S_1 or an effective boundary of S_2. Moreover, if S_1 and S_2 have no common effective boundaries, then each of their effective boundaries is also an effective boundary of S. Neither border polynomials, nor discriminant varieties enjoy a similar property, as illustrated by Example 7.

A first step toward Theorem 4 is a notion of "genericity" defined through Definitions 2 and 3, illustrated by Examples 3, 4 and 5, then studied by means of Propositions 3, 4 and 5. In fact, those propositions provide fundamental examples that are used in our subsequent results. The proofs of those propositions are routine and not reported here.

A second step is Theorem 3 which states that the effective boundaries of S depend only on its main components.

Definition 2. Let I be the ideal associated with the parametric system S. Recall that we assume that S is well-determined. Let $\mathfrak{p}$ be an associated prime ideal of I. The ideal $\mathfrak{p}$ is called a *main prime component* of I (or S) if $\mathbf{u}$ is a $\subseteq$-maximal algebraically independent set modulo $\mathfrak{p}$.

Definition 3. Let $S_1, \ldots, S_e$ be finitely many well-determined parametric polynomial systems in $\mathbb{Q}[\mathbf{u}, \mathbf{x}]$ with parameters $\mathbf{u}$. We say that an assertion on $S_1, \ldots, S_e$ is *generically true* if there exists a polynomial $t \in \mathbb{Q}[\mathbf{u}]$ such that for each parameter value α with $t(\alpha) \neq 0$, the assertion is true after specializing the systems $S_1, \ldots, S_e$ at α.

Next we show some of the related concepts based on the "generically true" statements.

Example 3. A well-determined parametric polynomial system S is said to be *generically zero-dimensional* if there exists $b \in \mathbb{Q}[\mathbf{u}]$ such that for each parameter value α with $b(\alpha) \neq 0$, the ideal of $S(\alpha)$ is zero-dimensional.

Example 4. Let S_1, S_2 be well-determined parametric semi-algebraic systems in $\mathbb{Q}[\mathbf{u}, \mathbf{x}]$ with parameters $\mathbf{u}$. We say that the two systems S_1 and S_2 *generically have the same set of solutions*, or *are generically equivalent* if there exists $t \in \mathbb{Q}[\mathbf{u}]$ such that for each parameter value α, with $t(\alpha) \neq 0$, the solution set of $S_1(\alpha)$ equals that of $S_2(\alpha)$. One can check that this property is equivalent to S_1 and S_2 have the same set of main components. Similarly, one can check that S_1 and S_2 generically have no common solutions if and only S_1 and S_2 have no common main components.

Example 5. Let S_1, S_2 be well-determined parametric semi-algebraic systems in $\mathbb{Q}[\mathbf{u}, \mathbf{x}]$ with parameters $\mathbf{u}$. The solution set of S is generically a disjoint union of the solution set of S_1 and that of S_2 if and only if the following conditions hold:

- S_1 and S_2 have no common main component,
- the union of the sets of main components of S_1 and that of S_2 equals to the set of main components of S.

The proof of the following statements, namely Proposition 3, 4 and 5, is rather straight forward.

Proposition 3. Let $S_1 := [F_{1=}, N_{\geq}, H_{1\neq}, P_{>}]$ and $S_2 := [F_{2=}, N_{\geq}, H_{2\neq}, P_{>}]$ be two parametric well-determined semi-algebraic systems [5] of $\mathbb{Q}[\mathbf{u}, \mathbf{y}]$ such that we have

$$\langle F_1 \rangle : (\prod_{h \in P \cup H_1} h)^{\infty} = \langle F_2 \rangle : (\prod_{h \in P \cup H_2} h)^{\infty}$$

Then, the systems S_1 and S_2 are generically equivalent.

Proposition 4. Let $S_1 := [T_{1=}, N_{\geq}, H_{1\neq}, P_{>}]$ and $S_2 := [T_{2=}, N_{\geq}, H_{2\neq}, P_{>}]$ be two parametric well-determined semi-algebraic systems of $\mathbb{Q}[\mathbf{u}, \mathbf{y}]$ for which T_1, T_2 are squarefree regular chains with $\mathbf{u}$ as main variables and such that $\mathrm{sat}(T_1) = \mathrm{sat}(T_2)$ holds. Assume also that every polynomial in $N \cup H_1 \cup H_2 \cup P$ is regular w.r.t. $\mathrm{sat}(T)$. Then, the systems S_1 and S_2 are generically equivalent.

Proposition 5. Let $S_1 := [T_{=}, N_{\geq}, H_{\neq}, P_{>}]$, $S_1 := [T_{1=}, N_{\geq}, H_{1\neq}, P_{>}]$ and $S_2 := [T_{2=}, N_{\geq}, H_{2\neq}, P_{>}]$ be three parametric well-determined semi-algebraic systems of $\mathbb{Q}[\mathbf{u}, \mathbf{y}]$ for which T, T_1, T_2 are squarefree regular chains with $\mathbf{u}$ as main variables and such that $\mathrm{sat}(T) = \mathrm{sat}(T_1) \cap \mathrm{sat}(T_2)$ holds. Assume also that every polynomial in $N \cup H \cup H_1 \cup H_2 \cup P$ is regular w.r.t. $\mathrm{sat}(T_1)$. Then, the solution set of S is generically the union of the solution set of S_1 and that of S_2.

Theorem 3 states that the effective boundaries of S depend only on its main components. That is, if two parametric polynomial systems have the same solutions generically, then they share the same effective boundary.

Theorem 3. Given two well-determined parametric semi-algebraic systems S_1 and S_2, suppose S_1 and S_2 are generically equivalent. Then we have $\mathrm{eb}(S_1) = \mathrm{eb}(S_2)$.

Proof. Let $t \in \mathbb{Q}[\mathbf{u}]$ such that we have

$$\forall \alpha \in Z_{\mathbb{R}}(t \neq 0) \Longrightarrow Z_{\mathbb{R}}(S_1(\alpha)) = Z_{\mathbb{R}}(S_2(\alpha)).$$

Let $b_1 \in \mathbb{Q}[\mathbf{u}]$ and $b_2 \in \mathbb{Q}[\mathbf{u}]$ be border polynomials of S_1 and S_2, respectively. Let $\mathbf{h}$ be an irreducible effective boundary of S_1 and let O be a witness ball of $\mathbf{h}$ such that we have $O \setminus \mathbf{h} \subset Z_{\mathbb{R}}(t \neq 0)$. Clearly, O is also a witness ball for $\mathbf{h}$ as an irreducible effective boundary of S_2. By the same arguments, we can show that any irreducible effective boundary of S_2 is also an effective boundary of S_1. $\square$

Corollary 1 ([7])**.** For any two STSASes $R_1 = [T_1, H_{1>}, P_{>}]$ and $R_2 = [T_2, H_{2>}, P_{>}]$ satisfying $\mathrm{sat}(T_1) = \mathrm{sat}(T_2)$, we have $\mathrm{eb}(R_1) = \mathrm{eb}(R_2)$.

Remark 4. According to Theorem 3, the inequations are irrelevant to the effective boundaries of a parametric semi-algebraic system. However, the inequalities do contribute to effective boundaries (see Example 6).

[5] Please refer to Section 2 for these notations.

Example 6. The following two parametric semi-algebraic systems illustrate that the inequalities do contribute to effective boundaries:

$$S_1 := \{f = 0\}, \quad S_2 := \{f = 0, g > 0\},$$

where $f := x^2 - ax + b$, $g := x + a$. S_1 The set of the effective border factors of S_1 is $\{-4b + a^2\}$ while the set of the effective border factors of S_2 is $\{-4b + a^2,\ b + 2a^2\}$.

Since we aim at decomposing the input system (using triangular decomposition), an interesting problem to investigate is the relation between the effective boundaries of this input system and those of the subsystems in one of its (triangular) decomposition. The following theorem is dedicated to this problem.

Theorem 4. Given three well-determined parametric semi-algebraic systems S, S_1 and S_2. We assume that the solution set of S is generically a disjoint union of the solution sets of S_1 and S_2. Then we have $\mathrm{ebf}(S) \subseteq \mathrm{ebf}(S_1) \cup \mathrm{ebf}(S_2)$. Moreover, if $\mathrm{ebf}(S_1) \cap \mathrm{ebf}(S_2) = \emptyset$ holds, then we have $\mathrm{ebf}(S) = \mathrm{ebf}(S_1) \cup \mathrm{ebf}(S_2)$.

Proof. By assumption, there exists a polynomial $t \in \mathbb{Q}[\mathbf{u}]$ such that for all $\alpha \in Z_{\mathbb{R}}(t \neq 0)$ we have

$$Z_{\mathbb{R}}(S(\alpha)) = Z_{\mathbb{R}}(S_1(\alpha)) \uplus Z_{\mathbb{R}}(S_2(\alpha)).$$

Let us denote the effective border polynomial factor sets of S, S_1, S_2 respectively by E, E_1, E_2. Let b, b_1, b_2 be border polynomials of S, S_1, and S_2, respectively.

We first prove that $\mathrm{ebf}(S) \subseteq \mathrm{ebf}(S_1) \cup \mathrm{ebf}(S_2)$ holds. To do so, we proceed by contradiction and consider an irreducible polynomial $p \in \mathbb{Q}[\mathbf{u}]$ such that $p \in E$ and $p \notin E_1 \cup E_2$ both hold. Thus, there exists a witness ball O of $Z_{\mathbb{R}}(p = 0)$ such that

$$O \subset Z_{\mathbb{R}}(\prod_{f \in E_1} f \prod_{f \in E_2} f \neq 0) \ \text{ and } \ O \setminus Z_{\mathbb{R}}(p = 0) \subset Z_{\mathbb{R}}(t\, b_1\, b_2 \neq 0)$$

both hold, thanks to Proposition 1 (and its proof). Let O_1, O_2 be the two connected components of $O \setminus Z_{\mathbb{R}}(p = 0)$. Let $\beta_1 \in O_1$, $\beta_2 \in O_2$. Clearly, O is a subset of a connected component of $Z_{\mathbb{R}}(\prod_{f \in E_1} f \neq 0)$. According to Theorem 2, we have $\#Z_{\mathbb{R}}(S_1(\beta_1)) = \#Z_{\mathbb{R}}(S_1(\beta_2))$. Similarly, we have $\#Z_{\mathbb{R}}(S_2(\beta_1)) = \#Z_{\mathbb{R}}(S_2(\beta_2))$.

Since $Z_{\mathbb{R}}(S(\beta_i)) = Z_{\mathbb{R}}(S_1(\beta_i)) \uplus Z_{\mathbb{R}}(S_1(\beta_i))$ holds for $i \in \{1, 2\}$, we deduce that $\#Z_{\mathbb{R}}(S(\beta_1)) = \#Z_{\mathbb{R}}(S(\beta_2))$ holds, which is a contradiction to the fact that O is a witness ball of $Z_{\mathbb{R}}(p = 0)$ (see property (c) in Definition 1).

Now, we assume that $\mathrm{ebf}(S_1) \cap \mathrm{ebf}(S_2) = \emptyset$ holds and we prove that $\mathrm{ebf}(S) = \mathrm{ebf}(S_1) \cup \mathrm{ebf}(S_2)$ holds. In fact, we just need to show that we have $\mathrm{ebf}(S_1) \cup \mathrm{ebf}(S_2) \subseteq \mathrm{ebf}(S)$. We proceed again by contradiction. Without loss of genericity, we assume $E_1 \setminus E \neq \emptyset$ and consider a polynomial $p \in E_1 \setminus E$. Thus there exists a witness ball O of p (as an effective border factor of S_1) such that we have

$$O \subseteq Z_{\mathbb{R}}(\prod_{f \in E} f \prod_{f \in E_2} f \neq 0) \ \text{ and } \ O \setminus Z_{\mathbb{R}}(p = 0) \subset Z_{\mathbb{R}}(t\, b\, b_2 \neq 0).$$

Let O_1, O_2 be the two connected components of $O \setminus Z_{\mathbb{R}}(p = 0)$. Let $\beta_1 \in O_1$, $\beta_2 \in O_2$. Clearly, O is a subset of a connected component of $Z_{\mathbb{R}}(\prod_{f \in E} f \neq 0)$. According to Theorem 2, we have $\#Z_{\mathbb{R}}(S(\beta_1)) = \#Z_{\mathbb{R}}(S(\beta_2))$. Similarly, we have $\#Z_{\mathbb{R}}(S_2(\beta_1)) = \#Z_{\mathbb{R}}(S_2(\beta_2))$. Since $Z_{\mathbb{R}}(S(\beta_i)) = Z_{\mathbb{R}}(S_1(\beta_i)) \uplus Z_{\mathbb{R}}(S_1(\beta_i))$ holds for $i \in \{1, 2\}$ we deduce that $\#Z_{\mathbb{R}}(S_1(\beta_1)) = \#Z_{\mathbb{R}}(S_1(\beta_2))$ holds, which contradicts the fact that O is a witness ball of p as an effective border factor of S_1. $\square$

A decomposition property such as Theorem 4 does not hold for border polynomials or discriminant varieties [12]. Indeed, for border polynomials and discriminant varieties, the projection of the intersection of the "main components" count. Let's see a concrete but simple example illustrating this fact.

Example 7. Let $S := \{f_1 f_2 = 0\}$, $S_1 := \{f_1 = 0\}$ and $S_2 := \{f_2 = 0\}$ be three systems with parameters (a, b, c), where

$$f_1 := ax^2 + bx + c,$$
$$f_2 := bx^2 + cx + a.$$

Clearly, the solution set of S is generically a disjoint union of the solution sets of S_1 and S_2. For each of S, S_1 and S_2, its border polynomial factors define its minimal discriminant variety. The border polynomial set, B, of S is:

$$\{a, b, -4ab + c^2, -4ac + b^2, c + a + b, c^2 - ac + a^2 - cb - ab + b^2\}.$$

The border polynomial set, B_1, of S_1 is: $\{a, -4ac+b^2\}$. The border polynomial set, B_2, of S_2 is: $\{b, -4ab+c^2\}$. The difference $B \setminus (B_1 \cup B_2)$ is $\{c + a + b, c^2 - ac + a^2 - cb - ab + b^2\}$.

Corollary 2. Given three squarefree regular chains T, T_1, T_2 and a polynomial set P in $\mathbb{Q}[\mathbf{u}, \mathbf{y}]$. We consider the STSASes $R := [T, P_>]$, $R_1 := [T_1, P_>]$, $R_2 := [T_2, P_>]$ and assume the following properties

(i) $\mathrm{sat}(T) = \mathrm{sat}(T_1) \cap \mathrm{sat}(T_2)$ holds,

(ii) R_1, R_2 have no common effective boundary.

Then $\mathrm{ebf}(R_1) \cup \mathrm{ebf}(R_2) = \mathrm{ebf}(R)$ holds.

Theorem 5. Let $R = [B_{\neq}, T, P_>]$ be a pre-regular semi-algebraic system and $D = \mathrm{oaf}(\mathrm{ebf}(R))$. Then $D \cup B$ is an FPS of R.

Proof. By Theorem 3 in [6] (which states a property of the oaf operator) each realizable strict sign condition on D defines a connected component of $Z_{\mathbb{R}}(\prod_{f \in D} f \neq 0)$.

Let α_1, α_2 be any two points realizing the same strict sign condition on $D \cup B$, say $\mathcal{S}$. Let $C := Z_{\mathbb{R}}(\mathcal{S})$. Observe that C is contained in a connected component of $Z_{\mathbb{R}}(\prod_{f \in D} f \neq 0)$, thus contained in a connected component of $Z_{\mathbb{R}}(\prod_{f \in \mathrm{ebf}(R)} f \neq 0)$.

Since α_1, α_2 are in the same connected component of $Z_{\mathbb{R}}(\prod_{f \in \mathrm{ebf}(R)} f \neq 0)$ and since $\alpha_1, \alpha_2 \in Z_{\mathbb{R}}(\prod_{f \in B} f \neq 0)$, Theorem 2 implies that we have $\# Z_{\mathbb{R}}(R(\alpha_1)) = \# Z_{\mathbb{R}}(R(\alpha_2))$. Therefore, by definition, $D \cup B$ is an FPS of R. $\square$

Remark 5. Let $R = [B_{\neq}, T, P_>]$ be a pre-regular semi-algebraic system. Theorem 5 implies that we can rely on $\mathrm{ebf}(R)$, rather than B, in order to compute an FPS of R. Since B can be much larger than $\mathrm{ebf}(R)$, this provides an opportunity to speedup computations in practice.

Observe that Corollary 1 shows that, for an STSAS, its inequations do not participate to the effective boundary set. However inequalities do.

Corollary 2 indicates that we can recycle the computation of an FPS for S in order to compute an FPS of S_1 and an FPS of S_2, when S splits to S_1 and S_2.

3.3. Effective boundaries of squarefree triangular semi-algebraic systems

In this section, we shall discuss effective boundaries in the context of triangular decomposition. Example 2 shows that some of the factors of a (or even the minimal) border polynomial may not be effective. The main result is that the border polynomial factors related purely to initials are not effective border polynomial factors.

Theorem 6. Let x be a variable and $f \in \mathbb{Q}[\mathbf{u}, x]$ be non-constant and such that the term of degree zero in x is not zero. Denote by δ the discriminant of f w.r.t. x. Let S be the parametric polynomial system $\{f = 0\}$ with $\mathbf{u}$ as parameters. Let $p \in \mathbb{Q}[\mathbf{u}]$ be an irreducible effective border polynomial factor of S. Then p divides δ in $\mathbb{Q}[\mathbf{u}]$.

We will need the notation and two lemmas below in order to prove Theorem 6.

Notation 1. Let r be a rational number and let

$$p = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

be a polynomial in $\mathbb{Q}[\mathbf{u}][x]$, where x is a variable. We denote by $M(p, x, r)$ the polynomial

$$x^n \left((a_n(r + \frac{1}{x})^n + a_{n-1}(r + \frac{1}{x})^{n-1} + \cdots + a_1(r + \frac{1}{x}) + a_0 \right).$$

It is easy to check that the initial of $M(p, x, r)$ is $a_n r^n + a_{n-1}r^{n-1} + \cdots + a_1 r + a_0$, that is, $p(r)$.

Lemma 4. Let $F := \{f_0, f_1, \ldots, f_k\}$ be a set of polynomials in $\mathbb{Q}[\mathbf{u}]$ with $\gcd(f_0, f_1, \ldots, f_k) = 1$, then there exists only finitely many rational numbers r such that $\gcd(f_0, \sum_{i=1}^k r^i f_k) \neq 1$ holds.

Proof. We prove this by contradiction. Assume that there are infinitely many rational numbers r such that $\gcd(f_0, \sum_{i=1}^k r^i f_k) \neq 1$ holds. Then, by the pigeonhole principle, there must exist an irreducible factor g of f_0 such that for infinitely many rational numbers r we have $g \mid \gcd(f_0, \sum_{i=1}^k r^i f_i)$. Let r_j, for $j = 0 \cdots k$, be distinct rational numbers such that $g \mid \gcd(f_0, \sum_{i=1}^k r_j^i f_i)$ holds. Let $q_j = f_0 + \sum_{i=1}^k r_j^i f_i$, for $j = 0 \cdots k$. Then $g \mid q_j$ holds as well. Observe that we have

$$\begin{pmatrix} 1 & r_0 & r_0^2 & \cdots & r_0^k \\ 1 & r_1 & r_1^2 & \cdots & r_1^k \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & r_k & r_k^2 & \cdots & r_k^k \end{pmatrix} \begin{pmatrix} f_0 \\ f_1 \\ \vdots \\ f_k \end{pmatrix} = \begin{pmatrix} q_0 \\ q_1 \\ \vdots \\ q_k \end{pmatrix}.$$

Therefore $g \mid f_i$ holds, for $i = 0 \cdots k$, since each f_i can be represented as a linear combination of $q_0, \ldots, q_k$. Therefore g is a common divisor of $f_0, f_1, \cdots, f_k$, which contradicts the assumption that $\gcd(f_0, f_1, \ldots, f_k) = 1$. □

Lemma 5. Let $f := a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ and $g := b_m x^m + b_{m-1}x^{m-1} + \cdots + b_1 x + b_0$ be two polynomials of $\mathbb{Q}[\mathbf{u}][x]$. Let r be a rational number such that $f(r)g(r) \neq 0$ holds. Denote by f^* and g^* the polynomials $M(f, x, r)$ and $M(g, x, r)$, respectively. Then we have

(i) $\mathrm{res}(f, g, x) = \mathrm{res}(f^*, g^*, x)$ holds,

(ii) $\mathrm{discrim}(f, x) = \mathrm{discrim}(f^*, x)$ holds.

In particular, if f and g are primitive w.r.t. x over $\mathbb{Q}[\mathbf{u}]$, then one can choose r such that

$$\gcd(\mathrm{init}(f, x), \mathrm{init}(f^*, x)) = 1 \quad \text{and} \quad \gcd(\mathrm{init}(g, x), \mathrm{init}(g^*, x)) = 1$$

both hold.

Proof. Let $\alpha_1, \ldots, \alpha_n$ be the n roots of f w.r.t. x in $\overline{\mathbb{Q}[\mathbf{u}]}$, the algebraic closure of $\mathbb{Q}[\mathbf{u}]$. Similarly, let $\beta_1, \ldots, \beta_n$ be the m roots of g w.r.t. x in $\overline{\mathbb{Q}[\mathbf{u}]}$. We make the following observations.

(a) From the Sylvester matrix, it is clear that $\mathrm{res}(f, g, x) = \mathrm{res}(x^n f(\frac{1}{x}), x^m g(\frac{1}{x}), x)$ holds.

(b) From the Poisson formula, that is,

$$\mathrm{res}(f, g, x) = a_n^m b_m^n \prod_{1 \le i \le n, 1 \le j \le m} (\beta_j - \alpha_i),$$

it is easy to see that for any rational number r, we have $\mathrm{res}(f, g, x) = \mathrm{res}(f(x + r), g(x + r), x)$.

Hence, we deduce that $\text{res}(f, g, x) = \text{res}(x^n f(\frac{1}{x}), x^m g(\frac{1}{x}), x) = \text{res}(f^*, g^*, x)$ holds. Note the n roots of f^* are $\frac{1}{\alpha_i - r}$ (for $i = 1, \ldots, n$). Recall that we have $\text{init}(f^*, x) = f(r)$. Therefore, thanks again to the Poison formula, the discriminant $\text{discrim}(f^*, x)$ is given by

$$\begin{aligned}
\text{discrim}(f^*, x) &= f(r)^{2n-2} \prod_{1 \le i < j \le n} \left(\tfrac{1}{\alpha_i - r} - \tfrac{1}{\alpha_j - r}\right)^2 \\
&= f(r)^{2n-2} \prod_{1 \le i < j \le n} \left(\tfrac{\alpha_j - \alpha_i}{(\alpha_i - r)(\alpha_j - r)}\right)^2 \\
&= f(r)^{2n-2} \frac{\text{discrim}(f, x)}{a_n^{2n-2}} \left(\tfrac{a_n}{f(r)}\right)^{2n-2} \\
&= \text{discrim}(f, x)
\end{aligned}$$

The second part follows from Lemma 4. $\square$

Some results similar to the first part of the above Lemma 5 can be found in [3]. We are ready to give a proof of Theorem 6.

Proof. Let r be a rational number such that $f(r) \neq 0$ (regarding f as a univariate polynomial in x) and $\gcd(\text{init}(f), \text{init}(f^*)) = 1$ where $f^* := M(f, x, r)$.

It is obvious that for each parameter value α such that $\text{init}(f^*)(\alpha)\ a_n(\alpha)\ a_0(\alpha) \neq 0$, the number of real solutions of $f = 0$ equals to that of $f^* = 0$. Let $\mathbf{h}$ be an irreducible effective boundary of $f = 0$ and O be a witness ball such that $O \setminus \mathbf{h} \subset Z_{\mathbb{R}}(a_0 \neq 0)$. Then it is easy to verify that O is also a witness ball for $\mathbf{h}$ to be an irreducible effective boundary of $f^* = 0$. Therefore, the set of irreducible effective border polynomial factors of $f = 0$ and that of $f^* = 0$ are the same.

Note that $\text{init}(f, x)\, \delta$ (resp. $\text{init}(f^*, x)\, \delta$) is a border polynomial of f (resp. f^*). According to Proposition 2, if $p \nmid \text{discrim}(f, x)$, then p divides both $\text{init}(f, x)$ and $\text{init}(f^*, x)$, which is a contradiction to $\gcd(\text{init}(f, x), \text{init}(f^*, x)) = 1$. $\square$

Remark 6. Theorem 6 can be generalized to a regular chain instead of a single polynomial. This can be stated as follows. Let T be a squarefree regular chain and view T as a parametric semi-algebraic system with the free variables of T as parameters; let p be an irreducible effective border factor of the parametric semi-algebraic system T. Then, the polynomial p divides $B_{sep}(T)$ in $\mathbb{Q}[\mathbf{u}]$. Our current proof involves a fair amount of intermediate notions and results, thus it will be reported in a future paper.

3.4. Effective boundaries: algorithmic benefits

In this section, we show how the notion of an effective boundary and related results can speed up the computation of finger polynomial sets in two major scenarios.

We stress the fact that we are not computing effective boundaries explicitly. However, we take advantage of the fact that Theorem 6 and Corollary 2 (or more generally Theorem 4) provide criteria certifying that a given polynomial is not an effective border polynomial factor. We illustrate these criteria in the above examples. We note that obtaining an efficient algorithm for computing effective boundaries is work in process.

Given a pre-regular semi-algebraic system $R = [B_{\neq}, T, P_{>}]$, all known approaches for computing an FPS of R proceeds by searching an FPS (hopefully small in size) from a large "universal" FPS set, such as $\text{oaf}(B)$ (proposed in [6]) or $\text{oaf}(\text{ebf}(S)) \cup B$ (proposed in Theorem 5 and based the notion of an effective boundary). The following simple example shows the difference between these two "universal" FPSes and why the latter is computationally advantageous.

Example 8. Let $f_1 := x - b^2 - a\,c, f_2 := x^2 - a\,x - b\,c$. Let $b_1 := \text{res}(f_1, f_2, x) = (-c\,b - b^2\,a - a^2\,c + b^4 + 2\,c\,b^2\,a + a^2\,c^2)$ and $b_2 := \text{discrim}(f_2, x) = a^2 + 4\,b\,c$. Note $R := [\{b_1, b_2\}, [f_1\, f_2], \emptyset]$ is a pre-regular semi-algebraic system. It is easy to deduce that $\text{ebf}(R) = \{b_2\}$, since f_1 is of degree 1 and b_1 is not a effective border polynomial factor according to Theorem 4. With the projection order $a > b > c$, we have $\mathrm{oaf}(\{b_2\}) = \{a, b, c, 4\,c\,b + a^2\}$ (4 polynomials, three of them being variables) and $\mathrm{oaf}(\{b_1, b_2\}) = \{a, b, c, -1 + c, -c + b^3, c - \frac{1}{2}, 4\,c\,b + a^2, \frac{1}{2} - c + c^2, c^2 - c + \frac{1}{8}, 4\,c^3 - 4\,c^2 + b^3, c + b^3 - 4\,c^2 + 4\,c^3, -\frac{1}{2}\,b^2 - a\,c + c\,b^2 + c^2\,a, \frac{1}{16} - \frac{1}{4}\,c + \frac{5}{4}\,c^2 - 2\,c^3 + c^4, b^3\,c^2 - b^3\,c + \frac{1}{4}\,b^3 + 4\,c^3 - 8\,c^4 + 4\,c^5, -c\,b - b^2\,a - a^2\,c + b^4 + 2\,c\,b^2\,a + a^2\,c^2\}$ (15 polynomials in total, much more complicated).

In practice, in the process of computing an FPS of a pre-regular semi-algebraic system, one usually ends up producing a set with richer properties, that we call a *comprehensive finger polynomial set* and that we define as follows. Let S be a well-determined parametric polynomial system with parameters $\mathbf{u}$. Then a polynomial set $D \subset \mathbb{Q}[\mathbf{u}]$ is called a *comprehensive fingerprint polynomial set* of S if for all $\alpha, \beta \in Z_{\mathbb{R}}(D_{\neq})$ with $\alpha \neq \beta$ the following holds: if the signs of $p(\alpha)$ and $p(\beta)$ are the same for all $p \in D$, then the number of solutions of $S(\alpha)$ is finite and the same as the number of solutions of $S(\beta)$. It is easy to check that any comprehensive FPS of the pre-regular semi-algebraic system $R = [B_{\neq}, T, P_{>}]$ containing the B set is also an FPS of R.

Now, when computing triangular decompositions of semi-algebraic system, the following scenario often occurs: a pre-regular semi-algebraic system $R := [B_{=}, T, P_{>}]$ splits into two subsystems $R_1 := [B_{\neq}, T_1, P_{>}]$ and $R_2 := [B_{\neq}, T_2, P_{>}]$ such that $Z_{\mathbb{R}}(R) = Z_{\mathbb{R}}(R_1) \cup Z_{\mathbb{R}}(R_2)$ and $B_{sep}(T_1) \cap B_{sep}(T_2) = \emptyset$ both hold. When this happens, we have $\mathrm{ebf}(R_1) \cap \mathrm{ebf}(R_2) = \emptyset$ and, moreover, a comprehensive FPS of R will often be a (comprehensive) FPS of R_1 and R_2. The following example (where the set-theoretical difference of two semi-algebraic sets is wanted) illustrates this computationally advantageous scenario.

Example 9. Let $R := [\{a, b, a - b\}, [(x^2 - a)\ (x^2 - b)], \{\ \}]$ be a pre-regular semi-algebraic system with parameters a, b. Let $R_1 := [\{a, b, a - b\}, [(x^2 - a)], \{\ \}]$ be another such system, which encodes the solutions of $Z_{\mathbb{R}}(R) \setminus Z_{\mathbb{R}}(x^2 - b = 0)$. Suppose that one has already checked that $D := \{a, b, a - b\}$ was a comprehensive FPS. Suppose also that the current task is to compute an FPS of R_1. Then one can check that D is indeed an FPS of R_1 at little cost since the sample points of each connected components of $Z_{\mathbb{R}}(D_{\neq})$ have already been computed when checking that D was an FPS of R.

4. Relaxation techniques

Let S be a parametric polynomial system in $\mathbb{Q}[\mathbf{u}, \mathbf{y}]$, where $\mathbf{u}$ are the parameters and $\mathbf{y}$ are the unknowns. A fundamental question is the following.

Question 1. For which parameter values of $\mathbf{u}$, does S have a prescribed number, say k, of distinct real solutions?

Question 1 is a generalization of the existential quantifier elimination problem, which can be answered by classical techniques such as cylindrical algebraic decomposition. If S is a well-determined system, letting b be a border polynomial of S one may ask a weaker but strongly related question. In many applications, however, answering this latter question is sufficient.

Question 2. Assuming that $b(\mathbf{u}) \neq 0$ holds, give a necessary and sufficient condition on $\mathbf{u}$ for S to have k distinct real solutions?

Question 2 is usually addressed in the following way, via real root classification [20, 19].

(0) Set B to be the irreducible factors of b; initialize F to B;

(1) Check whether there is a disjunction of strict sign conditions of polynomials in F, providing the desired necessary and sufficient condition.

(2) If so, return the disjunction of sign conditions; if not, enlarge F in an appropriate manner and return to Step (1).

The above procedure terminates within finitely many steps. We stress the fact, however, that this produces an answer to a *variant* of Question 2. Indeed, when the above procedure terminates, a necessary and sufficient condition for S to have k distinct real solutions is determined by strict sign conditions of polynomials in F, rather than B. It is, therefore, natural to consider the parameter values $\mathbf{u}$ satisfying the following two conditions

(*i*) $\prod_{f\in B} f(\mathbf{u}) \neq 0$ holds and,

(*ii*) there exists $f \in F \setminus B$ such that $f(\mathbf{u}) = 0$ holds.

and determine which of those parameter values yield k distinct real solutions for S. This leads to a recursive process, which is potentially very costly.

One natural way to tackle this question is to check whether or not the strict sign conditions of the polynomials in $F \setminus B$ can be relaxed into non-strict ones. Such a technique is called *relaxation* and is the topic of this section. To help the reader focusing on the main idea, the results of this section are stated for k taking the value "at least one" instead of a positive integer value. But all the results of this section hold also for k being any positive integer or even any positive integer range.

4.1. Relaxation

The following notions of *sign condition* and *relaxation* appear in [2] in a more general setting. Throughout this section, we consider a finite set $F \subset \mathbb{Q}[\mathbf{x}]$ such that the polynomials of F are pairwise coprime.

Definition 4. We call any semi-algebraic system of the form

$$\bigwedge\nolimits_{f\in F} f \; \sigma_f \; 0, \tag{4}$$

where σ_f is one of $>, <, \geq, \leq$, a *sign condition on* F, or an *F-sign condition.* An F-sign condition is called *strict* if every σ_f involved belongs to $\{>, <\}$. An F-sign condition C is called *realizable* if C has at least one real solution. If clear from context, F-sign condition is abbreviated to sign condition.

Definition 5 (Relaxation of sign condition)**.** For an F-sign condition C given as in (4) and a subset E of F, the *(partial) relaxation* of C *w.r.t.* E, denoted by $\widetilde{C}^E$, is defined by

$$\bigwedge\nolimits_{p\in F} p \; \widetilde{\sigma_p} \; 0 \quad \text{where} \quad \widetilde{\sigma_p} = \begin{cases} \leq, \text{ if } p \in E \text{ and } \sigma_p \text{ is } <, \\ \geq, \text{ if } p \in E \text{ and } \sigma_p \text{ is } >, \\ \sigma_p, \text{ otherwise.} \end{cases}$$

Let $Q = \bigvee_{i=1}^{e} C_i$ be a quantifier-free formula, where each C_i is an F-sign condition. The *relaxation* of Q w.r.t. E, denoted by $\widetilde{Q}^E$, is defined as $\bigvee_{i=1}^{e} \widetilde{C_i}^E$. If E contains only one polynomial h, then we also denote the relaxation by $\widetilde{Q}^h$.

Remark 7. If every conjunction C_i of Q is a strict F-sign condition, Q defines an open set. It is natural to ask whether $\widetilde{Q}^E$, the relaxation of Q w.r.t. E, defines an open set or not. It turns out that this openness testing problem is strongly related to Question 2, see Theorem 7. We first supply two examples. Example 10 illustrates the case where the relaxed formula still defines an open set while Example 11 supplies a case where the formula no longer defines an open set after relaxation is applied.

Example 10. It is easy to see that $\{x \in \mathbb{R} \mid x^2 - 2 > 0, x > 0\}$ and $\{x \in \mathbb{R} \mid x^2 - 2 > 0, x \geq 0\}$ are equal and open.

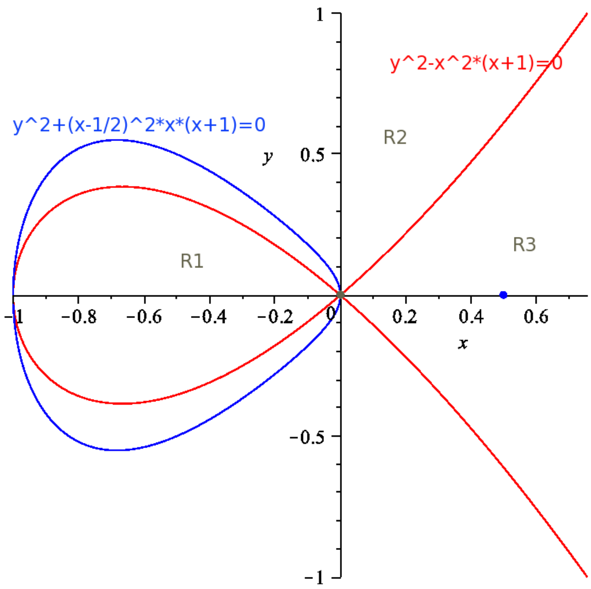


Fig. 2. The curves $b = 0$ and $p = 0$ in red and blue, respectively.

Example 11. Now consider two polynomials $b = y^2 - x^2(x+1)$ and $p = y^2 + (x-1/2)^2 x(x+1)$. The complement of the hypersurface $b = 0$ in $\mathbb{R}^2$, the (x, y)-plane, consists of three regions, namely R_1, R_2, R_3, as shown on Figure 11. A quantifier-free formula describing R_1 is $Q = (b < 0 \wedge p > 0)$. The relaxation of Q w.r.t. p is $\widetilde{Q}^p = (b < 0 \wedge p \geq 0)$, whose zero set contains the region R_1 and an isolated point $(x = \frac{1}{2}, y = 0)$ from the region R_3. Therefore, the zero set of $\widetilde{Q}^p$ is not open.

Remark 8. As we will see with Theorem 7, a question related to the relaxation of sign conditions is to decide whether a given semi-algebraic set, represented by a Tarski quantifier-free formula, is open or not. This problem can be solved via quantifier elimination. Given a semi-algebraic set S in $\mathbb{R}^n$ represented by a quantifier-free formula Q, consisting of constraints of polynomials in $\mathbb{Q}[\mathbf{x}]$, whether or not S is open can be characterized by the following quantified formula:

$$\forall (x_1, x_2, \ldots, x_n)\ \exists \epsilon\ \forall (a_1, a_2, \ldots, a_n)$$
$$\left(Q(x_1, x_2, \ldots, x_n) \wedge \epsilon > 0 \wedge \sum_{i=1}^{n} (x_i - a_i)^2 < \epsilon \right) \implies Q(a_1, a_2, \ldots, a_n).$$

By using a general purpose quantifier elimination routine, one can eliminate the quantified variables in the above formula. and obtain "true" (resp. "false") if S is open (resp. not open). This method will introduce $n+1$ extra variables and so is not of practical interest. However, the criterion for testing whether a semi-algebraic set (represented by a Tarski quantifier-free formula) is open or not, provided by Theorem 7, reduces to testing whether a semi-algebraic set is empty or not, without introducing extra variables. This is a much more attractive result in terms of algebraic complexity. Note that checking whether or not a semi-algebraic set is empty as well as computing the set-theoretic difference of two semi-algebraic set can be done by algorithms presented in [7] and the software tools in the `RegularChains` library [5].

Notation 2. Let us fix notation for the rest of Section 4. Let S, B, F be as in the discussion preceding Section 4.1, thus S is a parametric polynomial system in $\mathbb{Q}[\mathbf{u}, \mathbf{y}]$, with $\mathbf{u}$ as parameters and B, F are finite subsets of $\mathbb{Q}[\mathbf{u}]$ such that $B \subseteq F$ holds and B is the set of the irreducible factors of a border polynomial of S. In addition, let $D \subseteq \mathbb{Q}[\mathbf{u}]$ be such that we have $B \subseteq D \subseteq F$. Let Q_i $(i = 0, 1)$ be a quantifier-free formula in disjunctive form such that each conjunction clause C of it is in the following form: $C = \wedge_{f \in F} f \ \sigma_f \ 0$, where $\sigma_f \in \{>, <\}$ if $f \in D$ and $\sigma_f \in \{\geq, \leq\}$ if $f \in F \setminus D$. Moreover, assume that for any parameter value u such that $D(u) \neq 0$ holds, the system $S(u)$ has (resp. has no) real solutions if and only if $Q_1(u)$ (resp. $Q_0(u)$) is true. Let h be a polynomial in $D \setminus B$. Denote by D^h the set $D \setminus \{h\}$. Denote by ∂_i $(i = 0, 1)$ the frontier of the set $Z_{\mathbb{R}}(Q_i)$. Denote by G_i $(i = 0, 1)$ the set $Z_{\mathbb{R}}(\widetilde{Q_i}^h) \cap \overline{Z_{\mathbb{R}}(Q_i)}$. Finally, let S_i $(i = 0, 1)$ be the semi-algebraic set such that $Z_{\mathbb{R}}(\widetilde{Q_i}^h) = G_i \uplus S_i$ holds, where the symbol $\uplus$ denotes disjoint union.

4.2. Main Theorem

The following result implies a practical criterion for relaxation. Indeed Condition (i) can be tested algorithmically. When this condition holds, Condition (ii) is the desired conclusion. Examples 12 and 13 illustrate how to implement this technique.

Theorem 7. The following three statements are equivalent:

(i) we have $Z_{\mathbb{R}}(\widetilde{Q_1}^h) \cap Z_{\mathbb{R}}(\widetilde{Q_0}^h) = \emptyset$,

(ii) for any $u \in Z_{\mathbb{R}}(D^h_{\neq})$, the system $S(u)$ has real solutions (resp. no real solutions) if and only if $\widetilde{Q_1}^h(u)$ (resp. $\widetilde{Q_0}^h(u)$) is true,

(iii) both sets $Z_{\mathbb{R}}(\widetilde{Q_0}^h)$ and $Z_{\mathbb{R}}(\widetilde{Q_1}^h)$ are open.

Before providing the proof, we supply several properties of the objects involved in Theorem 4.2, through a series of lemmas.

Lemma 6. $Z_{\mathbb{R}}(Q_0)$ and $Z_{\mathbb{R}}(Q_1)$ are both open sets.

Proof. On one hand, $Z_{\mathbb{R}}(D_{\neq}) = Z_{\mathbb{R}}(Q_0) \uplus Z_{\mathbb{R}}(Q_1)$. On the other hand, there exists a finite set of connected open sets, $\mathcal{O} = \{C_1, \ldots, C_e\}$, such that $Z_{\mathbb{R}}(D_{\neq}) = \cup_{i=1}^{e} C_i$ holds. By Lemma 2, for each $C_i \in \mathcal{O}$, either $C_i \subseteq Z_{\mathbb{R}}(Q_0)$ or $C_i \subseteq Z_{\mathbb{R}}(Q_1)$ holds. Therefore, each of $Z_{\mathbb{R}}(Q_0)$ and $Z_{\mathbb{R}}(Q_1)$ is a union of finitely many elements of $\mathcal{O}$ and is are open. $\square$

Lemma 7. For each parameter value $u \in \overline{Z_{\mathbb{R}}(Q_0)} \cap Z_{\mathbb{R}}(B_{\neq})$, the specialized system $S(u)$ has no real solutions; for each parameter value $u \in \overline{Z_{\mathbb{R}}(Q_1)} \cap Z_{\mathbb{R}}(B_{\neq})$, the specialized system $S(u)$ has real solutions.

Proof. Let $u \in \overline{Z_{\mathbb{R}}(Q_0)} \cap Z_{\mathbb{R}}(B_{\neq})$. There exists a connected component C of $Z_{\mathbb{R}}(Q_0)$ and a connected component C' of $Z_{\mathbb{R}}(B_{\neq})$ such that $u \in \overline{C} \cap Z_{\mathbb{R}}(B_{\neq}) \subseteq C'$ holds. Since $C \subseteq Z_{\mathbb{R}}(Q_0) \subseteq Z_{\mathbb{R}}(B_{\neq})$, we have $C \subseteq C'$. Since the number of real solutions of R is constant above C' (by Lemma 2) and R has no real solutions above C, we conclude that $S(u)$ has no real solutions. The other part of the lemma is proved similarly. $\square$

Note that $G_i = Z_{\mathbb{R}}(\widetilde{Q_i}^h) \cap \overline{Z_{\mathbb{R}}(Q_i)} \subseteq Z_{\mathbb{R}}(B_{\neq}) \cap \overline{Z_{\mathbb{R}}(Q_i)}$ holds for $i = 0, 1$. We have the following proposition as a direct consequence of Lemma 7.

Proposition 6. For each parameter value $u \in G_0$, the specialized system $S(u)$ has no real solutions; for each parameter value $u \in G_1$, the specialized system $S(u)$ has real solutions.

Lemma 8. The following relations hold:

$$\partial_0 \cup \partial_1 = Z_{\mathbb{R}}(\prod_{f \in D} f) \ \text{ and } \ \partial_0 \cap \partial_1 \subseteq Z_{\mathbb{R}}(\prod_{f \in B} f).$$

Proof. By Lemma 6, both $Z_{\mathbb{R}}(Q_0)$ and $Z_{\mathbb{R}}(Q_1)$ are open sets. We have $\partial_0 \cup \partial_1 = \partial(Z_{\mathbb{R}}(Q_0) \cup Z_{\mathbb{R}}(Q_1))$, since $Z_{\mathbb{R}}(Q_0) \cap Z_{\mathbb{R}}(Q_1) = \emptyset$ holds. Therefore, we have

$$\begin{aligned} \partial_0 \cup \partial_1 &= \overline{Z_{\mathbb{R}}(Q_0) \cup Z_{\mathbb{R}}(Q_1)} \setminus (Z_{\mathbb{R}}(Q_0) \cup Z_{\mathbb{R}}(Q_1)) \\ &= \overline{Z_{\mathbb{R}}(D_{\neq})} \setminus (Z_{\mathbb{R}}(D_{\neq})) \\ &= Z_{\mathbb{R}}(\textstyle\prod_{f \in D} f). \end{aligned}$$

By Lemma 7, the sets $\overline{Z_{\mathbb{R}}(Q_0)} \cap Z_{\mathbb{R}}(B_{\neq})$ and $\overline{Z_{\mathbb{R}}(Q_1)} \cap Z_{\mathbb{R}}(B_{\neq})$ are necessarily disjoint. Therefore, the set $\overline{Z_{\mathbb{R}}(Q_0)} \cap \overline{Z_{\mathbb{R}}(Q_1)} \cap Z_{\mathbb{R}}(B_{\neq})$ is empty. Since $\partial_0 \subseteq \overline{Z_{\mathbb{R}}(Q_0)}$ and $\partial_1 \subseteq \overline{Z_{\mathbb{R}}(Q_1)}$ hold, the conclusion follows. □

Lemma 9. For $i = 0, 1$, the following relations hold:
(*a*) $\overline{Z_{\mathbb{R}}(Q_i)} \cap Z_{\mathbb{R}}(D^h_{\neq}) \subseteq Z_{\mathbb{R}}(\widetilde{Q_i}^h)$,
(*b*) $Z_{\mathbb{R}}(\widetilde{Q_0}^h) \cup Z_{\mathbb{R}}(\widetilde{Q_1}^h) = Z_{\mathbb{R}}(D^h_{\neq})$.

Proof. Since $Z_{\mathbb{R}}(\widetilde{Q_i}^D)$ is a closed set, we have $\overline{Z_{\mathbb{R}}(Q_i)} \subseteq Z_{\mathbb{R}}(\widetilde{Q_i}^D)$. Therefore, we have

$$Z_{\mathbb{R}}(D^h_{\neq}) \cap \overline{Z_{\mathbb{R}}(Q_i)} \subseteq Z_{\mathbb{R}}(D^h_{\neq}) \cap Z_{\mathbb{R}}(\widetilde{Q_i}^D) = Z_{\mathbb{R}}(\widetilde{Q_i}^h).$$

By (a), we have $Z_{\mathbb{R}}(D^h_{\neq}) \cap (\cup_{i=0,1} \overline{Z_{\mathbb{R}}(Q_i)}) \subseteq \cup_{i=0,1} Z_{\mathbb{R}}(\widetilde{Q_i}^h)$, which implies $Z_{\mathbb{R}}(D^h_{\neq}) \subseteq Z_{\mathbb{R}}(\widetilde{Q_0}^h) \cup Z_{\mathbb{R}}(\widetilde{Q_1}^h)$. And $Z_{\mathbb{R}}(\widetilde{Q_0}^h) \cup Z_{\mathbb{R}}(\widetilde{Q_1}^h) \subseteq Z_{\mathbb{R}}(D^h_{\neq})$ holds since all polynomials in D^h remain strict after relaxing h. □

Proposition 7. The inclusions $S_0 \subseteq Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(\widetilde{Q_0}^h)$ and $S_1 \subseteq Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(\widetilde{Q_1}^h)$ both hold.

Proof. Let $i \in \{0, 1\}$. Recall that we have $G_i = Z_{\mathbb{R}}(\widetilde{Q_i}^h) \cap \overline{Z_{\mathbb{R}}(Q_i)}$ and $Z_{\mathbb{R}}(\widetilde{Q_i}^h) = G_i \uplus S_i$. Therefore, we have $Z_{\mathbb{R}}(Q_i) \subseteq G_i$ and $S_i = Z_{\mathbb{R}}(\widetilde{Q_i}^h) \setminus G_i \subseteq Z_{\mathbb{R}}(\widetilde{Q_i}^h) \setminus Z_{\mathbb{R}}(Q_i) \subseteq Z_{\mathbb{R}}(h = 0)$. Hence, we deduce that $S_i \subseteq Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(\widetilde{Q_i}^h)$ holds. □

Lemma 10. Both $S_1 \subseteq G_0$ and $S_0 \subseteq G_1$ hold.

Proof. By Lemma 8, we have $\partial_0 \cup \partial_1 = Z_{\mathbb{R}}(\prod_{f \in D} f = 0)$. Since $h \in D$, we have $Z_{\mathbb{R}}(h = 0) \subseteq \partial_0 \cup \partial_1$, which implies that $Z_{\mathbb{R}}(h = 0)$ can be rewritten as

$$Z_{\mathbb{R}}(h = 0) \cap ((\partial_0 \setminus \partial_1) \uplus (\partial_1 \setminus \partial_0) \uplus (\partial_0 \cap \partial_1)).$$

By Lemma 8, we have $\partial_0 \cap \partial_1 \subseteq Z_{\mathbb{R}}(\prod_{f \in B} f = 0)$, which implies that $Z_{\mathbb{R}}(D^h_{\neq}) \cap \partial_0 \cap \partial_1 = \emptyset$ holds. Let S_h be $Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(D^h_{\neq})$. Then S_h can be rewritten as $(Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(D^h_{\neq}) \cap (\partial_0 \setminus \partial_1)) \uplus (Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(D^h_{\neq}) \cap (\partial_1 \setminus \partial_0))$.

Intersecting both sides of relation (a) of Lemma 9 with $\overline{Z_{\mathbb{R}}(Q_i)}$, we obtain $Z_{\mathbb{R}}(D^h_{\neq}) \cap \overline{Z_{\mathbb{R}}(Q_i)} \subseteq G_i$, which implies that $Z_{\mathbb{R}}(D^h_{\neq}) \cap \partial_i \subseteq G_i$ holds. Therefore, we have $S_h \subseteq G_0 \cup G_1$.

Since $Z_{\mathbb{R}}(\widetilde{Q_i}^h) \subseteq Z_{\mathbb{R}}(D^h_{\neq})$, we have $Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(\widetilde{Q_i}^h) \subseteq S_h$. By Proposition 7, we have $S_i \subseteq Z_{\mathbb{R}}(h = 0) \cap Z_{\mathbb{R}}(\widetilde{Q_i}^h)$. Therefore, we also have $S_i \subseteq S_h$. Finally, we deduce the desired conclusion by combining the inclusions $S_i \cap G_i = \emptyset$, $S_i \subseteq S_h$, and $S_h \subseteq G_0 \cup G_1$. □

Corollary 3. We have $Z_{\mathbb{R}}(\widetilde{Q_1}^h) \cap Z_{\mathbb{R}}(\widetilde{Q_0}^h) = S_0 \uplus S_1$.

Proof. We can rewrite $Z_{\mathbb{R}}(\widetilde{Q_1}^h) \cap Z_{\mathbb{R}}(\widetilde{Q_0}^h)$ as the disjoint union $(S_1 \cap G_0) \uplus (S_0 \cap G_1) \uplus (S_1 \cap S_0) \uplus (G_0 \cap G_1)$. By Proposition 6, we have $G_0 \cap G_1 = \emptyset$. Together with Lemma 10, we have $Z_{\mathbb{R}}(\widetilde{Q_1}^h) \cap Z_{\mathbb{R}}(\widetilde{Q_0}^h) = S_0 \uplus S_1$. $\square$

Next, we prove Theorem 7.

Proof. By Lemma 9, we first observe that $Z_{\mathbb{R}}(\widetilde{Q_0}^h) \cup Z_{\mathbb{R}}(\widetilde{Q_1}^h) = Z_{\mathbb{R}}(D^h_{\neq})$ holds. We now prove four implications, which establish the theorem.

$(i) \Rightarrow (ii)$**:** By Corollary 3, we have $S_0 = S_1 = \emptyset$ and $Z_{\mathbb{R}}(\widetilde{Q_i}^h) = G_i$ $(i = 0, 1)$. Then the conclusion follows from Proposition 6.

$(ii) \Rightarrow (i)$**:** We proceed by contradiction. Assume (i) does not hold. There exists $u \in Z_{\mathbb{R}}(D^h_{\neq})$, such that both $\widetilde{Q_0}^h(u)$ and $\widetilde{Q_1}^h(u)$ are true. This is a contradiction to (ii).

$(ii) \Rightarrow (iii)$**:** This follows from Lemma 6.

$(iii) \Rightarrow (ii)$**:** On one hand, each connected component of $Z_{\mathbb{R}}(\widetilde{Q_1}^h)$ is open; thus, it contains at least one connected component of $Z_{\mathbb{R}}(Q_1)$ as subset. On the other hand, each connected component of $Z_{\mathbb{R}}(\widetilde{Q_1}^h)$ is a subset of a connected component of $Z_{\mathbb{R}}(B_{\neq 0})$. So (ii) follows by Lemma 2.

$\square$

Example 12 (A simple example where relaxation enlarges the zero set)**.** Consider

$$Q_0 := (xy - 1 < 0 \wedge x \neq 0) \ \vee \ (xy - 1 > 0 \wedge x > 0), \quad Q_1 := (xy - 1 > 0 \wedge x < 0).$$

Let $D := \{xy - 1, x\}$. It is easy to verify that Q_0, Q_1, D satisfy the assumptions in Theorem 7. Observe that

$$\widetilde{Q_0}^x = (xy - 1 < 0) \ \vee \ (xy - 1 > 0 \wedge x \geq 0), \quad \widetilde{Q_1}^x = (xy - 1 > 0 \wedge x \leq 0)$$

and $Z_{\mathbb{R}}(\widetilde{Q_0}^x) \cap Z_{\mathbb{R}}(\widetilde{Q_1}^x) = \emptyset$ hold. Moreover, we have $Z_{\mathbb{R}}(\widetilde{Q_1}^x) = Z_{\mathbb{R}}(Q_1)$ and $Z_{\mathbb{R}}(\widetilde{Q_0}^x) \setminus Z_{\mathbb{R}}(Q_0) = Z_{\mathbb{R}}(xy-1 < 0 \wedge x = 0) \neq \emptyset$. Thus the "relaxed set" $Z_{\mathbb{R}}(\widetilde{Q_0}^x)$ is larger than $Z_{\mathbb{R}}(Q_0)$.

Example 13 (An example based on a parametric system study)**.** Consider the semi-algebraic system

$$S := \{x^2 + y - a = 0, y^2 + x - b = 0\}$$

with parameters a, b. The polynomial

$$p := 256a^3 - 288ab + 27 + 256b^3 - 256a^2b^2$$

is a border polynomial of S. Our goal is to compute the conditions on parameters (a, b) such that S will specialize to have real solutions, provided that $p \neq 0$ holds.

Denote by q the polynomial $ab - \frac{9}{16}$. Figure 3 illustrates the curves $p = 0$ and $q = 0$ in the (a, b)-plane. As a first step, one may obtain a "partial solution": provided that $q \neq 0$, $a \neq 0$, and $p \neq 0$ all hold, the system S has real solutions if and only if the quantifier free formula

$$Q_1 := ((a > 0 \wedge p > 0) \ \vee \ (p < 0 \wedge q > 0)) \ \wedge \ (a \neq 0 \wedge q \neq 0 \wedge p \neq 0)$$

holds. Let Q_0 be the quantifier free formula such that, provided that $q \neq 0$, $a \neq 0$, and $p \neq 0$ all hold, the system S has no real solutions if and only if Q_0 holds.

To get a complete answer, we need to consider the parameter values concealing q, p and a. The relaxation technique can help with studying the parameters annulling p and a here (for those that annul q, we can not use relaxation).

Consider the quantifier free formula $\widetilde{Q_1}^p$. Note that the set $Z_{\mathbb{R}}(\widetilde{Q_1}^p)$ is still open, since $Z_{\mathbb{R}}(\widetilde{Q_1}^p) \cap Z_{\mathbb{R}}(\widetilde{Q_0}^p) = \emptyset$ holds and Theorem 7 applies; note also that $Z_{\mathbb{R}}(\widetilde{Q_1}^p) \setminus Z_{\mathbb{R}}(Q_1)$ is not empty. Moreover,

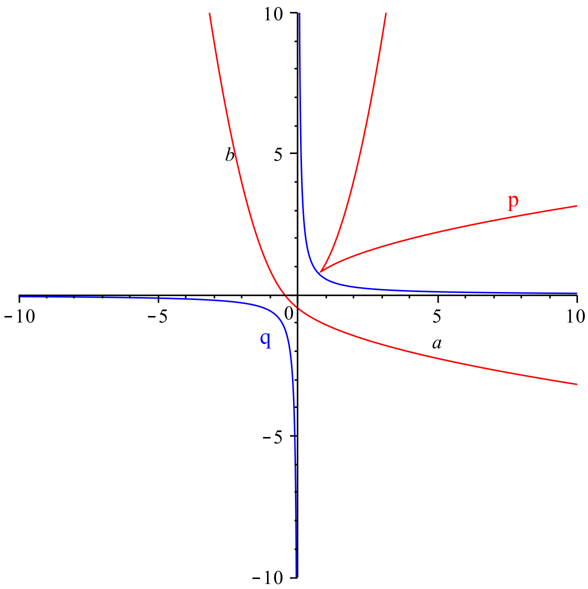


Fig. 3. The curves $p = 0$ and $q = 0$ in red and blue, respectively.

we deduce that provided that $q \neq 0$, $a \neq 0$, the system S has real solutions if and only if the quantifier free formula $\widetilde{Q_1}^p$ holds. Similar computation can be carried out for the parameter values canceling a.

Finally, we obtain the following complete answer: Provided that $q \neq 0$ holds, the system S has real solutions if and only if the following quantifier free formula holds:

$$(a \geq 0 \wedge p \geq 0 \wedge q \neq 0) \ \vee \ (p \leq 0 \wedge q > 0).$$

4.3. Experimentation

In this section, we report on experimenting with the relaxation technique implied by the results of Section 4.2. To this end, we have added an option to the RealTriangularize command in order to take advantage of Theorem 7. When a recursive call occurs in the conditions described in the discussion preceding Section 4.1, we test whether Condition (i) of Theorem 7 holds or not. If it does, we deduce from Condition (ii) that the recursive call can be avoided. With this enhancement, the RealTriangularize command possesses two options, both available in `Maple` 16.

- One controls the scheme of the algorithm which can be either *recursive* or *incremental*, as presented respectively in our papers [6] and [7].
- The other controls whether or not Theorem 7 is applied. Note that applying Theorem 7 implies testing Condition (i) which may have some non-negligible computational cost.

Our experimentation was conducted with `Maple` 16 on a machine with Intel Core 2 Quad CPU (2.40GHz) and 3.0Gb RAM memory. The time-out is set at 3600 seconds. The memory usage is limited to 60% of the RAM memory. In Table 1, RTD denotes RealTriangularize. The subscripts re and inc denote respectively the recursive and incremental schemes of RealTriangularize's algorithm. The suffixes $+relax$ and $-relax$ denote respectively applying and not applying relaxation techniques, that is, Theorem 7. For each algorithm option, the left column records the time (in seconds) while the right one records the number of components in the output. NA means the computation does not finish in the resource (time or memory) limit.

Table 1 The timing and number of components in the output with or without relaxation

sys	RTD$\vert_{re}$−relax		RTD$\vert_{re}$+relax		RTD$\vert_{inc}$−relax		RTD$\vert_{inc}$+relax	
8-3-config-Li	418.6	203	410.6	203	30.5	47	30.4	47
dgp6	65.17	20	64.37	20	47.73	19	47.43	19
Leykin-1	4.9	28	4.9	28	6.5	19	6.5	19
L	14.9	69	14.9	69	2.6	19	2.6	19
Mehta0	1294	21	713.6	15	1558	20	998.9	15
EdgeSquare	247.7	116	725.3	91	116.8	43	629.4	33
Enneper	6.1	18	5.4	13	4.9	17	4.9	12
IBVP	14.1	8	16.8	4	2.5	8	7.6	4
MPV89	2.7	6	2.4	5	2.1	7	2.1	6
SEIT	NA	NA	1411	1	NA	NA	NA	NA
Solotareff-4b	3223	3	3222	3	3424	3	3424	3
Xia	223.7	12	224.8	10	21.4	9	20.5	8
Lanconelli	1.1	7	1.1	7	1.0	7	1.0	7
MacLane	17.4	79	17.3	79	5.8	27	5.8	27
MontesS12	197.8	163	197.4	163	49.9	85	49.7	85
MontesS14	3.4	23	3.4	23	2.8	15	2.9	15
Pappus	750.5	409	748.2	409	29.1	119	29.0	119
Wang168	7.0	16	7.1	16	3.4	11	3.5	11
xia-issac07-1	2.7	13	4.4	11	2.2	12	4.2	10

For the system Metha0, with the relaxation techniques, both timing and the number of components in the output are reduced. For the system SEIT, with the help of the relaxation techniques, $RTD|_{re}$ can now solve it within half an hour, while RTD is not successful with its other options. For the systems Enneper, IBVP, MPV89, Xia and xia-issac07-1, relaxation techniques help reducing the number of components in the output with a slight time overhead.

5. Conclusion

The notion of an effective boundary provides a novel framework for solving parametric polynomial systems and semi-algebraic systems via triangular decomposition methods. For the purpose of designing decomposition algorithm, this new tool overcomes some of the weaknesses of the related notions of border polynomial and discriminant variety. We are currently working on taking advantage of effective boundaries within the algorithms presented in our recent papers [6, 7].

The technique of relaxation is another new tool for improving the practical efficiency of those decomposition algorithms and others, such as real root classification. Motivated by the idea of avoiding redundant computations, this technique achieves this goal on various problems presented in this paper, leading some time to more compact output or to reduced computing time or to both. We are currently investigating other criteria serving the same goal.

References

[1] S. Basu, A. Gabrielov, and N. Vorobjov. Monotone Functions and Maps. `http://arxiv.org/abs/1201.0491`.

[2] S. Basu, R. Pollack, and M-F. Roy. *Algorithms in real algebraic geometry.* Springer-Verlag, 2006.

[3] C. W. Brown. Improved projection for cylindrical algebraic decomposition. *J. Symb. Comput.*, 32(5):447–465, 2001.

[4] C. W. Brown and J. H. Davenport. The Complexity of Quantifier Elimination and Cylindrical Algebraic Decomposition. In C.W. Brown, editor, *Proceedings ISSAC 2007*, pages 54–60, 2007.

[5] C. Chen, J.H. Davenport, F. Lemaire, M. Moreno Maza, B. Xia and R. Xiao, Y. Xie. Computing the Real Solutions of Polynomial Systems with the RegularChains Library in MAPLE. Software demo at ISSAC 2011, June 8-11, 2011, San Jose, CA ACM Communications in Computer Algebra, Issue 117, pages 166–168.

[6] C. Chen, J.H. Davenport, J. May, M. Moreno Maza, B. Xia, and R. Xiao. Triangular decomposition of semi-algebraic systems. In *Proc. of ISSAC 2010*, pages 187–194, 2010.

[7] C. Chen, J. H. Davenport, M. Moreno Maza, B. Xia, and R. Xiao. Computing with semi-algebraic sets represented by triangular decomposition. In *Proc. of ISSAC 2011*, pages 75–82, 2011.

[8] C. Chen and M. Moreno Maza. Algorithms for Computing Triangular Decompositions of Polynomial Systems In *Proc. of ISSAC 2011*, ACM Press, 2011.

[9] C. Chen, M. Moreno Maza, B. Xia, and L. Yang. Computing cylindrical algebraic decomposition via triangular decomposition. In *ISSAC'09*, pages 95–102, 2009.

[10] G.E. Collins. Quantifier Elimination for Real Closed Fields by Cylindrical Algebraic Decomposition. In *Proc. of 2nd. GI Conference Automata Theory & Formal Languages*, pages 134–183, 1975.

[11] J.H. Davenport and J. Heintz. Real Quantifier Elimination is Doubly Exponential. *J. Symb. Comput.*, 5:29–35, 1988.

[12] D. Lazard and F. Rouillier. Solving parametric polynomial systems. *J. Symb. Comput.*, 42(6):636–667, 2007.

[13] E.L. Lima. The Jordan–Brouwer Separation Theorem for Smooth Hypersurfaces. *American Math. Monthly* 95:39–42, 1988.

[14] S. McCallum and C. W. Brown. On delineability of varieties in cad-based quantifier elimination with two equational constraints. In *Proc. of ISSAC 2009*, pages 71–78, 2010.

[15] M. Moreno Maza, B. Xia and R. Xiao. On solving parametric polynomial systems. *Proc. of MACIS 2011*, S. Ratschau Ed., pages 205-215, 2011.

[16] J.R. Munkres. *Topology.* Second edition, Prentice Hall, New Jersey, USA, 2000.

[17] A. Strzeboński. Solving systems of strict polynomial inequalities. *J. Symb. Comput.*, 29(3):471–480, 2000.

[18] D. M. Wang. *Elimination Methods.* Springer, New York, 2000.

[19] R. Xiao. *Parametric Polynomial System Solving.* PhD thesis, Peking University, Beijing, 2009.

[20] L. Yang, X. Hou, and B. Xia. A complete algorithm for automated discovering of a class of inequality-type theorems. *Science in China, Series* **F**, 44(6):33–49, 2001.

[21] L. Yang and B. Xia. Real solution classifications of a class of parametric semi-algebraic systems. In *Proc. of the A3L'05*, pages 281–289, 2005.